CONCEIVING SYSTEMS

Derek K Hitchins

Thesis submitted for the

Degree of

Doctor of Philosophy

Department of Systems Science
City University
London

January 1990

Conceiving Systems

Table of Contents

CONCEIVING SYSTEMS

by

Professor Derek K HITCHINS

ABSTRACT

The thesis is concerned with the development of innovative, robust design concepts for a class of systems called Information Decision Action (IDA) Systems. IDA systems are typified by Command and Control (C^2) and Command, Control, Communications and Intelligence (C^3I) systems as used by police, emergency services and the military - the two titles refer respectively to the human activity and the technological systems. The class of systems is much wider, however, and includes, financial, traffic control, business and even governmental systems where information is gathered, used as a basis for human decision-forming, and results in action, all in real, or near-real time. IDA system complexity stems largely from the dominance of robust human activity systems within the overall system, and also from the employment of often-rigid, technology-based, decision support systems which are unable to adapt as swiftly as the humans they serve.

The thesis is in two parts. In the first part, the author presents a perspective on "hard" and "soft" systems and the gradual move by so-called "hard" systems engineers towards softer concepts in the search for more satisfactory IDA systems. This progression is presented partly by anecdote, supported by some of the author's papers showing the development of his contribution to understanding of, and partly by an exposition of the essential themes inherent in, IDA systems. Keynote papers in the first part are: *MOSAIC: Concepts for the Deployment of Air Power in Europe* and *The Human Element in C^3I*: The first of these presents a highly-survivable alternative to the present force and C^2 deployment approaches which have evolved little since World War II; the second considers the human and his social behaviour as keys to understanding IDA systems. Other papers develop the themes and show their application to systems in which the author has had major involvement

The second part is concerned with the process of conceiving and creating IDA systems and it too draws on published papers as direct support for the thesis. Keynote papers here are *A General Theory of Command and Control*, a unique recent paper which proposes a set of design axioms for an idealized IDA system, the award-winning *Managing Systems Creation* which presents

an engineering framework for Creating Systems, and *SEAMS (Systems Engineering, Analysis and Management Support)* which signals a major design initiative to develop engineering frameworks into company-wide IT environments. The second part also introduces a complete Conceiving System, called the Seven-Step Continuum (SSC), describes some prototype tools developed by the author to perform some of the tasks of design conception and - in Chapter 9, which is a paper within the thesis - shows results from using the SSC, its methods and tools, in practice. The second part closes with a look forward to the building of flexible future systems which can adapt to their environment.

Part A

Information - Decision - Action Systems

1. INTRODUCTION

1.1. Coming to Terms with the Human Dimension

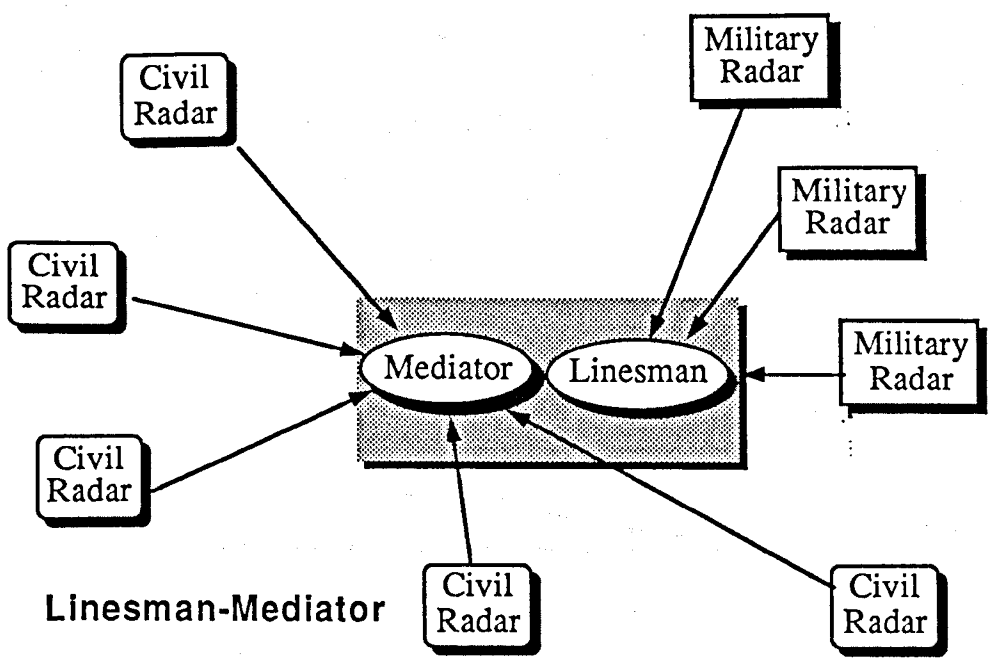
Early Training

Like all young engineers of my era, and like almost all engineers today, I was trained largely in the science of design. Unlike many of today's young engineers, however, I was trained in a school which believed that future careers for young engineers would change often and radically; in consequence, the training was both broad and deep. Training covered mechanical, electrical, electronic and aeronautical engineering, with digital and analogue computing, materials science, thermodynamics, hydraulics, pneumatics and many others thrown in. Not forgotten were the humanities of life, English, literature, leadership, history, some economics and even a little politics and international affairs. Consistently, I was taught that an answer existed to any design issue and that it was undoubtedly a quantitative answer.

The training was expensive, prolonged and instilled great confidence in the recipients; as time was to tell, much of the training was to prove inappropriate to any of the jobs which followed. In my case, I was in an organization dedicated to operations where design was not exercised, merely observed and frequently despaired over.

Ambitious Design Concepts

My first job of significance was concerned with a major undertaking - Linesman / Mediator. The system concept was concerned with correlating air defence of the UK with civil air traffic control. A number of military radars was to be installed around the coast of the UK to detect intruding aircraft and to help control our interceptor aircraft on to their potential targets. Civil air traffic control radar pictures were also to become involved. To harmonize civil and military air traffic, and to provide central control of air operations, all the radar data were to be transferred by radio links to a central point near London, where civil and military activities could combine. Linesman was the military part; Mediator was the civil part.



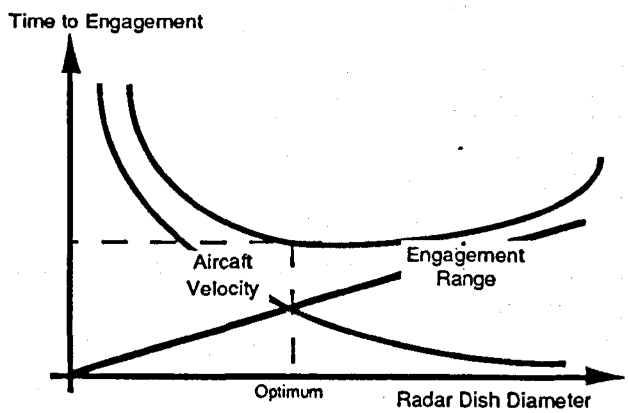
Linesman-Mediator

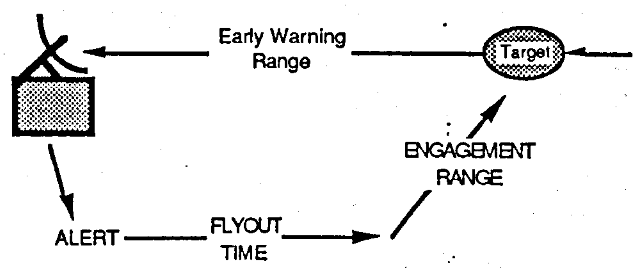
The design concept was years ahead of the technology. The architectural concept was vulnerable to attack, since it presented a spectacular node, the disabling of which would bring chaos to both civil and military air activities. The expense was unacceptable, too. Not all was bad, however. The military radars from that era are still working today. Out of the project's ashes is emerging IUKADGE, the Improved UK Air Defence Ground Environment, some 25 years later. And, on a personal note, I learned more lessons from observing failure than from success. One lesson was concerned with architecture, another with optimization.

Optimizing through Compromise

My particular role in Linesman Mediator was concerned with the Lightning interceptor. The Lightning was developed from a very high speed prototype; the operational aircraft had to be fitted with radar and missiles which, while making it into a useful fighting machine, would also impair its speed. The optimization process was designed to minimize the incursion by enemy aircraft into UK airspace. The system therefore comprised the ground radars for long range warning, the ground fighter controllers who directed the interception to the general area of the incoming target, the Lighting pilot with his aircraft, its radar and its missiles - and of course the hostile.

The principal area of design compromise concerned the aircraft radar and missiles. If the radar were to detect the target and home-in, then it needed long range. But longer range meant a wider radar dish, which widened the aircraft fuselage, which increased the drag, which slowed the aircraft, used more fuel and reduced the radius of action. Similarly, longer range in the radar would be of value only if the missile could use the extra target detection range, which meant larger missile motors, which also increased drag and weight. Increasing the range of the ground surveillance radars was difficult, as they were at the edge of the then-technology.





Compromise was reached by single-minded dedication to one measure of effectiveness - shortest enemy penetration distance. Penetration distance could be minimized by a variety of approaches:

Reducing the sum of the following times:

* Time taken by ground (or airborne patrol) surveillance to alert the fighters

* Time for the chosen fighter to transit to the target area

* Time to engage the enemy

These times could, in turn, be reduced not only by technical optimization, but also by locating the surveillance radars and the fighter bases well forward in the direction of expected incursion.

With the need to see even beyond ground surveillance radar horizon, the potential for an Airborne Early Warning aircraft became apparent.

The results were impressive, as those associated with UK Air Defence can testify. From that point, I was dedicated to the systems approach, whatever that was - it gave real results to real problems.

Territorial Imperatives

Later I became the Tri-Service Consultant on Automatic Test Equipment (ATE) to all three services, Army, RN and RAF. This MOD post gave considerable insight into management issues between the various companies competing for defence contracts. It was the fashion at the time to replace a plethora of test facilities serving the hundred or more black boxes to be found on any modern aircraft with one ATE. In principle, ATE is as hard a systems problem as one might expect to meet. As more black boxes are taken off the list of manually tested boxes and brought into the queue for the one ATE, the latter's queue becomes progressively longer, following an exponential growth pattern predictable by simple queuing theory. By producing a cost model it is simple enough to show that, beyond a certain level of automation, the cost of filling the ATE queue with expensive black boxes exceeds the savings to be achieved by the introduction of the ATE.

Unexpectedly, this tidy concept presented a number of problems we might call "soft", in that they were not concerned with the engineering and costs at all; difficult though the engineering was for these Wurlitzers of the avionics world, the management problems proved even more intractable.

The contractor for the single ATE invariably was also a manufacturer of some of the boxes it was intended to test, and in this arena he was in competition with other box manufacturers. They were required to give the ATE contractor details of their boxes so that he could write test programs, but he was a competitor. Result? Stalemate. The box manufacturers did not actually refuse to give the necessary test data, they just did not hand it over. This outcome was, of course, predictable; the procurement agents had not considered the idea of competition. As it was, an apparently considerable economy from the use of ATE was effectively denied.The lesson to be learned was concerned with territorial imperatives, boundary marking and the host of simian behavioural characteristics we humans display in our everyday lives

Image versus Human Needs

As a last example of appreciating the human condition, consider the plight of the aircrew of a Tornado fighter aircraft as they cruise at altitude over a featureless North Sea for seven hours at a stretch, day in, day out with rarely another aircraft in sight. Their radar produces a synthetic image - there is no screen "noise", so with no target the screen is plain black. As the Tornado Avionics System Design Manager, this time in industry, I co-operated with the Institute of Aviation Medicine in devising a scheme for the pilot and the navigator, who sit in tandem, to be able to play GO or draughts between their respective screens. This had a number of advantages: it kept their eyes on the radar screens; it improved manual dexterity with the somewhat cumbersome hand controllers; and it stopped them from becoming too bored and losing concentration. The MOD turned down the no-cost proposal because "none of our aircrew officers is prone to loss of concentration". A triumph of image over necessity? Yes, but the real lesson was concerned with tribal loyalties, taboos and boundaries.

1.2. DAMASCUS.....

Feeling through Fog

These experiences have been recounted to show developing awareness that all is not well with the progress of hard systems engineering in which I, along with many others, have been immersed for many years. The number of major system fiascoes is almost countless; it is rather hard, in fact, to find a successful major systems engineering project, particularly in the defence field. Nimrod, the Type 22 Frigate and the Sergeant York Missile are just three of the contemporary representatives of a long heritage.

Like other system designers in the late 70s, I recognized the symptoms but found solutions less easy to come by. Until, that is, the U.S. company, TRW, sent Robert J. Lano to Europe in 1979 to give a series of two day lectures on C3I - Command, Control, Communications and Intelligence.

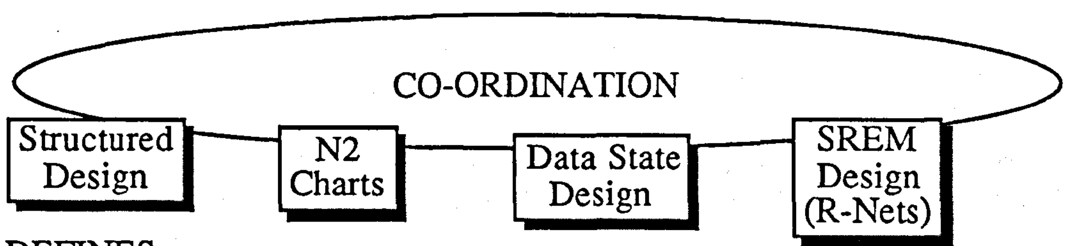
Robert J Lano

Robert Lano is a software and systems engineer who was specializing in methodologies at the end of the 1970s. Regrettably, but like many employees of companies, he has published little in the way of papers or books - see Lano, (1979) and Lano, (1980). His lectures were excellent however. He showed how, using four.essentially simple, hand-operated techniques, it was possible to unravel, understand, redesign and re-integrate complex systems. His insight was concerned with levels of abstraction, viewpoints and, although he never mentioned the word, architecture. His toolset has found wide use in many companies throughout the western world's defence industry; many of the papers included in this document use the methods (as I have come to apply them - I have modified their application from, principally, software analysis to general system analysis.)

The Lano Toolset and the Viewpoint

The four tools introduced by Lano were: N^2 ("N-Squared") Charts, Data State Design, R-Nets and Structured Analysis. N^2 Charts are explained in the enclosed papers "Managing Systems Creation" and "The Human Element in C^3I", Data State Design is best seen at work in "Air Command and Control", R-Nets and Structured Analysis are used in "Managing Systems Creation"

Lano did not invent all of these tools. R-Nets, or Requirement Nets are part of a US Programme called SREM, Software Requirements Engineering Methodology, see Alford (1985). The R-Net technique is finding application in Manufacturing Systems Design, Hughes (1986). Data State Design is said by Lano to have been produced by Peterson of IBM. Lano was responsible for bringing them together, however, and for using them together to develop differing viewpoints of the system in question. Structured Analysis is well known and documented, Ross, (1977), Ross and Schoman (1977), Gane and Sarson (1979). Lano seems to have been responsible for recognizing the true potential of the N^2 chart and, of course, for bringing the tools together in a matched set.



DEFINES

* FUNCTIONS	*INTERFACES	*DATA FILES	*TRANSACTIONS
* HIERARCHY	*RELATIONSHIPS	* DATARELATION	*DATA FLOWS

DOES NOT DEFINE

*DATA FILES	*HIERARCHY	*FUNCTIONS	*HIERARCHY
*INTERFACES	*DATA FILES	*HIERARCHY	*INTERFACES
*DATA FLOWS	*TRANSACTIONS	*DATA FLOWS	*DATA RELATIONS
*TRANSACTIONS	*DATA RELATIONS	*INTERFACES	*RELATIONSHIPS

Lano's Toolset

The viewpoints, advantages and limitations of the four tools are shown in the figure and it can be seen that all four tools had to be used concurrently to address any problem at system level

Lano's insight unlocked the door to understanding for many designers and systems engineers; he was seminal and inspirational, but seems singularly unrecognized in the world of systems science.

1.3. Aims and Objectives

This introduction has been anecdotal in nature to show some of the developments which have been taking place in the systems engineering world and most particularly in the major defence systems engineering arena where systems engineering is most active due to the large, complex and general one-off nature of the projects. Systems engineering technique has been developing, albeit slowly, but its method of development has to a large extent been bottom-up, that is driven by component and subsystem influences, notably those of software engineering. The objectives, then, of this document are as follows:-

*** To review the present Systems Engineering Scene**

*** To identify shortcomings in that Scene**

*** To examine the contribution which softer methods might make to Systems Engineering**

*** To develop the essentially human and anthropomorphic nature of human-designed IDA systems**

*** To introduce the concept of Conceiving Systems as a bridge between Enquiring Systems and Creating, or Design & Development, Systems**

*** To present a Conceiving System for the conception of Information-Decision-Action Systems designs.**

*** To explore the role of the user / operator in Conceiving Systems**

*** To achieve these objectives, in part, by presenting work undertaken in the last 10 years and, in so doing, to present a view of the confluence of the so-called Hard and Soft Schools of systems thinking.**

1.4. The Structure of the Document

Chapter 2 will review the developments in systems thinking, both in the so-called "hard" practice of systems engineering and in the "soft" systems school as it relates to conceiving systems.

Chapter 3 will develop the underlying themes to be found in the author's papers, enclosed at the end of the document, in respect of Information-Decision-Action systems

Chapter 4, looks at the human aspects of IDA systems, and presents a view, unlike SSM, of the predictability of human activities when human activities occur in groups under the influence of territorial imperatives and "tribal" influences

Chapter 5 considers systems for creating systems, metasystems, as a pre-cursor to understanding conceiving systems

Chapter 6 presents the basic elements of, and arguments for, a system for developing design concepts - a conceiving system, also a metasystem.

Chapter 7 presents and explains the Seven-Step Continuum, a conceiving system

Chapter 8 reviews the tools and techniques needed to support the Seven-Step Continuum

Chapter 9, effectively a paper within the body of the thesis, applies the Seven-Step Continuum to a variety of problems to demonstrate its use in action

Chapter 10 looks ahead to the use of conceiving systems in building future systems

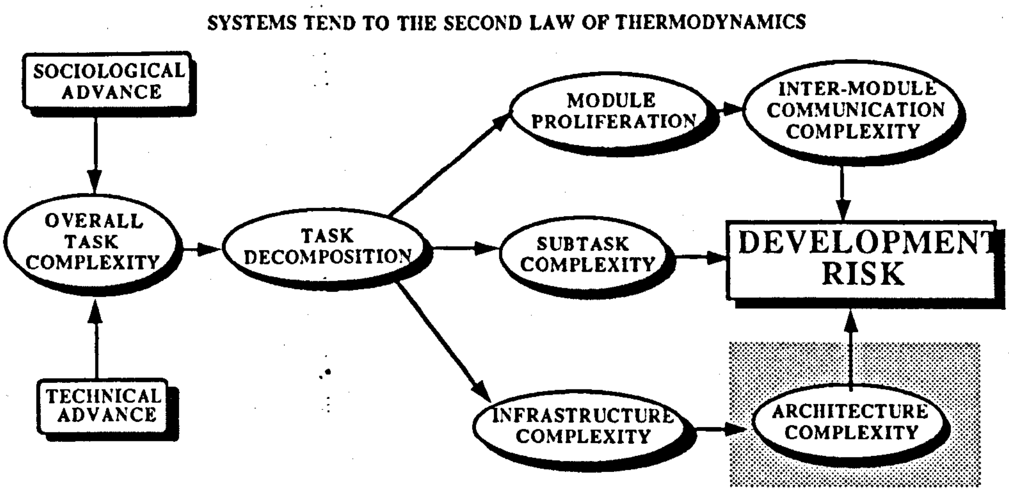
2. DEVELOPMENTS IN SYSTEMS THINKING

2.1. Continuing Development in Systems Engineering

Systems engineering lacks a definition with which practitioners would generally agree. Academia proposes holism, emergent properties, management of complexity and other criteria. In industry, viewpoints are more pragmatic and focused, sometimes even myopic. To some, systems engineering is little more than progress-chasing to ensure that all engineering aspects of a project are considered and progressed according to plan. To others, many others, systems engineering is making the software and the hardware of a computer-based system work together. To a few, systems engineering encompasses all aspects of analysis, design, development, manufacture, maintenance and operation - but excluding project management, which is generally felt to be separate.

Industry's better approaches to systems engineering are perhaps epitomized in the conception, design and development of satellites, where the constraints of launcher space and payload invoke a severe optimization ethic on each of the many subsystems, which must be optimized not only in physical, but in temporal and financial dimensions, simultaneously and continually throughout a project. The one-shot opportunity presented to many satellite projects requires an intimate relationship between all participants, be they project managers, systems engineers, scientists, or single discipline engineers in order to meet the launch window. And, of course, the system design skills needed for deep-space probes, with their very long, unattended lifetimes in a hostile environment, has been tempered in the most demanding of furnaces. Unmanned satellites are, perhaps, the present-day pinnacle of hard systems engineering achievement

The advent of distributed computing has indelibly changed the engineering task on a much wider front than satellite engineering, however. As more computers have been introduced into human activities, supposedly to ease the burden of human effort (or to remove the human from the activity altogether), the interconnection of computing elements has burgeoned.



Architecture Generation

As more is demanded of systems and as they become more complex and interwoven, there is a continuing need to reduce them to understandable elements, or subtasks in the figure. While systems may be evolving, man is not changing perceptibly and the size of subtask which he can accommodate mentally has changed little either. Thus many more "man-sized" subtasks become necessary to constitute more complex systems, and infrastructure expands accordingly.This interweaving of modules, links and channels is referred to as architecture. "Architecture" generally does not seem to include the humans as part of the system, however.

The architectural complexity introduced by distributed computing and distributed operating systems, of which the best known is UNIX, has grown explosively. As little as ten years ago, architectures for computer based systems comprised a central processor, star-connected to sets of peripherals, some via modems to permit distant operation. Now the trend is towards distributed, replicated databases and away from server systems to sets of desktop processors, each of individual power greater than yesterday's central processor. Integrated services digital networks promise intelligent communications, with the network containing stored knowledge of the location of data, so that users may simply ask for what they want and the system will conduct the necessary transactions with a variety of remote databases, collate the results and present the answer, translated if need be.

To accommodate this wealth of complexity, a variety of robust standards and protocols has emerged. Standards for communication include the International Standards Organization's Open Systems Interconnection (ISO OSI), a seven-layered protocol for interconnecting heterogeneous processors by telecommunications, see Bird (1987). In the manufacturing industry in the US, the car manufacturer General Motors has developed a widely-used, 3-

layer protocol called MAP (Manufacturing Automation Protocol - see MAP, (1985) - which is much faster than the OSI protocol due to the smaller number of layers, and is used for the "tails" or stubs which attach small, dedicated systems to larger, networked systems.

At a lower level, communications protocols have been introduced for packet switching , circuit switching and message switching systems and the military, not to be outdone, have developed their own, generally to deal with the much harsher environments that they must anticipate. Processor protocols are less in evidence as each vendor has, until recently, vied to be as different as possible in order to preserve his or her market area, a classic example of territorial marking.

The biggest single influence on the systems engineering front has undoubtedly come from advances made in software engineering. These advances have come of necessity, as debacle followed debacle with timescale and cost overruns often in the several hundred percent. Initially, the defence industry lead the way in developing solutions because they tended to be involved with more advanced, more complex project developments. Latterly the burden has perceptibly shifted toward commercial applications as these become more complex, more widely interwoven and, in particular because they tend to be so short lived compared with their military counterparts.

A significant shift is observable in the technology which supports this vast new industry; whereas the military called the tune on new devices and materials in the seventies, now the military are following industry and using industry-proven devices, protocols and applications. This shift is inevitable as computer-based technology reaches the vast market presented by medium and even small enterprises throughout the western world, which market greatly exceeds the slow-turnover, specialist military market.

Software engineering has seen the concomitant development of an array of tools principally designed to provide:-

*A much needed discipline in the capture and decomposition of software requirements - see Jackson (1975) and DeMarco (1978)

*Configuration management, made essential by the large number of system modules in simultaneous development at various build standards

*Project management over the software analysis, design and development life cycle

That these tools and techniques development have greatly benefitted software engineering is unquestioned; later topics will , however, examine the risk to creativity that is endemic in their use.

Advanced Networked Systems Architecture

One current enterprise is worthy of mention under the present heading - the Advanced, Networked System Architecture, ANSA. Readers should note the title carefully; it refers, not to a network, but to a networked system. ANSA originated in the UK's Alvey Project and is sponsored by:-

British Telecom
Digital Equipment Company Ltd
GEC Marconi Ltd
GEC Plessey Telecommunications Ltd
Hewlett-Packard Ltd, Information Technology Ltd
RACAL Group Services Ltd
Standard Telephone and Cable Laboratories Ltd.

The development of ANSA is being conducted, post Alvey, by Architecture Projects Management Ltd of Cambridge, with continuing sponsorship from the participant companies. ANSA is an ambitious concept An extract from the current ANSA Reference Manual, Volume A, will indicate the point

"ANSA supports the design, implementation, operation and evolution of distributed information processing systems where the different components that make up the system, such as application packages, operating systems, computers and networks, come from different vendors. The complexity, and consequent cost, that arises from this heterogeneity of hardware and software can be significantly reduced if information technology vendors adopt a common approach to the design and interconnection of the components they offer

"Common design principles based on a consistent model of information processing across the system ease the task of integrating diverse applications into a coherent, extendible system.........

".......ANSA has five projections called enterprise, information, computational, engineering and technology.......The enterprise projection is concerned with modelling the relationships between the system owners, operators and users............"

ANSA concerns itself also with:-

*Financial Security	*Responsibility	
*Mistrust		
*Ownership	*Value	*Privacy
*Security	*Fear of Failure	*Ethics
*Fear of Injury and Damage	*Fear of Alienation	
*Frustration		
*Fear of Loss	*Reasoning, Creativity and Decision Making	
*Aesthetics and Presentation	*Avoidance of Tedious Work	

ANSA is state-of-the-art distributed computer system design, at least in concept - much of it is still on the drawing board; yet it serves to show the trend in commercial systems engineering, which quite clearly embraces - as it must - the human dimension

The ANSA, and other system engineers', viewpoints are inconsistent with the comments of some soft systems practitioners who express the view that systems engineering is concerned with mathematically determinable aspects of design, the so-called hard features and who further suggest that soft systems have "taken over" as it were from the RAND school of OR-and science-based analytical, reductionist hard systems approach. On the contrary, the hard practitioners have been moving steadily towards the "firm' centre for over 15 years and now embrace some - although far from all - of the soft concepts, albeit using markedly different approaches appropriate to their environment.

2.2. Present Limitations in Systems Engineering Methods

The Systems Theology

The lack of consensus concerning the nature of systems engineering was referred to above. It is, perhaps, unwise to pin the definition of such a subject down too narrowly in any event.There is within the minds of most practitioners a wider acceptance of the meaning of "the Systems Approach" (see the enclosed paper *"Systems Creativity")*. In broad terms, there is kernel Top Down approach to system design, which supposedly pursues the following path:-

Start at the highest abstract level of system requirement description

B. Functionally decompose the requirement

C. Map the decomposed functions to the elements of a physical architecture

D. Develop, and progressively integrate the physical elements into a system

Hall (1962) provides a more detailed, classical view of the steps in systems engineering. Unfortunately, there is no generally accepted way to achieve the first three steps, although many systems engineers have patent methods of their own; it is perhaps for this reason that some observers regard systems engineering more as an arcane art-form than a science-based engineering discipline

Indeed, there is no consensus on the meaning of "function" in the Systems Approach. Many experienced systems engineers cannot establish the range of functions in, say, a radar system. The usual first decomposition is into sensors, communications, processing and display. Unfortunately, that turns out to be most unhelpful since these so-called functions apply to almost every system of significance and hence the division offers no insight into a particular system

The point can best be understood by attempting to functionally decompose a human being. To start at sensors (all five), nerves and brain is not really helpful: not only is it true for all animals (and therefore offers no discrimination or differentiation), the decomposition misses out the interesting parts of the human design. A better approach is needed, and will be described later.

Similarly, functional-to-physical mapping is obscure in practice, with most practitioners unwilling to be drawn on their methods, which are often therefore declared to be "obvious". These methods generally recognize the complexity of interfaces between groups of functions and attempt to group functions so as to reduce the residual interface complexity between the resulting groups. The methods tend to be difficult to justify under pressure, the more so since they are generally based on an uncertain definition of functions to be grouped.

Part of the systems engineering ethic includes the development of design options, their modelling and subsequent tradeoff to select the most cost-effective solution. Choosing the appropriate range of design options from which to trade is, at the best, a crude art. Modelling of systems is becoming increasingly expensive, although there is a growing tendency to prototype information and Command & Control systems. Trading between options often employs dubious weighting-and-scoring techniques. And there is, surprisingly, no agreed way of defining what the term "cost-effective" means.

All in all, it has to be said that the theology on which systems engineering is supposed to be based has some dubious foundations. Systems engineering survives, even thrives, despite this lack of scientific underpinning, principally because it is a theology, a way of approaching problems that is axiomatically sound. It must be better to approach a problem top-down, to view the whole system rather than simply its parts, to proceed from function or purpose toward realization, and so on.

Software Engineering Threats to Systems Creativity

There is a risk, presently significant and increasing rapidly, that the application of requirements capture tools to higher levels of system design will erode creativity at those levels. The tools are seductively simple to use, being generally graphically based, but it is broadly their task to decompose established design concepts into their component parts, ensuring consistency and completeness in the process. The present tendency to use these tools to formulate the design concepts themselves, and particularly those of the higher, parent system rather than the computer-dominated sub-system, is not encouraging. .

Examples of this problem area are security-sensitive in some instances. One recent paper on the application of Yourdon (Real Time), as defined by Ward and Mellor (one of the many software requirements analysis tools) concerned a complex but un-named modern defence platform. The tool had been used to perform a functional decomposition of the complete platform requirement, after a period of knowledge elicitation. The result was incomprehensible to users and to peer group designers and was not followed up by functional-to-physical mapping. This last point arose for several reasons. First, the time taken to perform the analysis had exceeded expectation; second, the team could find no way to integrate the decomposed functions; third, not surprisingly, the MOD customer was becoming nervous about both cost and the validity of the approach

"Off-the-Shelf" Syndrome

Before moving on to softer issues, it is worthwhile to consider perhaps the biggest element which militates against top-down design in industry. Companies which design, develop and manufacture components and systems, invest intellectual effort as well as money in the process. Faced with the requirement for a new system, the pressure to simply adapt their presently available, "off-the-shelf" system to a new requirement is almost irresistible

Unfortunately, there is ample evidence to show that complex systems rarely repeat themselves and that the off-the-shelf solution is an illusion. That this must be so, can be seen by considering just one aspect, the timescale of contemporary projects. An off-the-shelf solution must have been designed several years before to be on the shelf now. With the pace of technological development so high, the off-the-shelf solution is out of date. Perhaps it is for this reason that more developed countries attempt to unload their present solutions on to third world countries. Third world countries are no longer naive, however, and are often smarting from having been "sold a pup" in the past. The true ethic of the systems engineer has always been to do the best by his customer. Current claims being made for "Total Quality Management" and similar vogue concepts look suspiciously like the standards which were applied assiduously in the past by systems engineers

2.3. Enquiring Systems

Since the end of the seventies the systems science community has seen an upsurge of the so-called "soft" systems approach as opposed to "hard" methods which, according to Checkland (1981), are characterized by assumptions that "problems can be formulated as making a choice between alternative means of achieving a known end". It is not the purpose of this document to provide an authoritative assessment of these methods, but some are of considerable interest , particularly in the field of management, and their broad approach may offer succour to systems engineers who, as has been demonstrated, are moving steadily into less firm territory in an attempt to solve mounting human problems. These problems include the understanding of the requirement, an area where soft methods promise capability.

SSM and Peter Checkland

The doyen of the soft academics is undoubtedly Professor Peter Checkland from Lancaster University who produced his seminal book "Systems Thinking, Systems Practice" in 1981, some two years after Robert Lano et al were showing their methods to the world. Checkland's Soft Systems Methodology, see Checkland (1972) and Checkland (1981), conceives of hierarchies of systems including natural systems whose origin is in the origin of the universe, designed physical systems which man has made, ranging "from hammers via tram cars to space rockets", designed abstract systems such as mathematics, poetry, books and human activity systems. In earlier work, Checkland (1971), had closely linked some human activity systems with natural and designed systems, but in "Systems Thinking..." he took the view that definition of human activity systems should be confined to the activities themselves.

Unfortunately, this definition tends to exclude the Information Decision Action system if, as later chapters will strongly suggest, there is advantage in perceiving an IDA system as containing both the humans and their technology in one inseparable set. Happily, at least one of Checkland's acolytes, Brian Wilson (Wilson,1984), takes a broader view by describing human activity systems as "undertaking purposeful activity such as man-machine systems, industrial activity, political systems, etc.

Both Checkland and Wilson follow a classically simple route in exploring their problem domains. In essence, they appreciate a real problem situation, analyse and express the problem situation, formulate a variety of viewpoints concerning the real problem situation, form idealized conceptual models of the problem situation, compare characteristic features of the idealized model with the real world, and hence identify any feasible and desirable change. By choosing a variety of viewpoints they hope to bring robustness to the process and avoid the pitfalls of pre-conceptions.(I hope that precis of their approach is not so brief as to misrepresent).

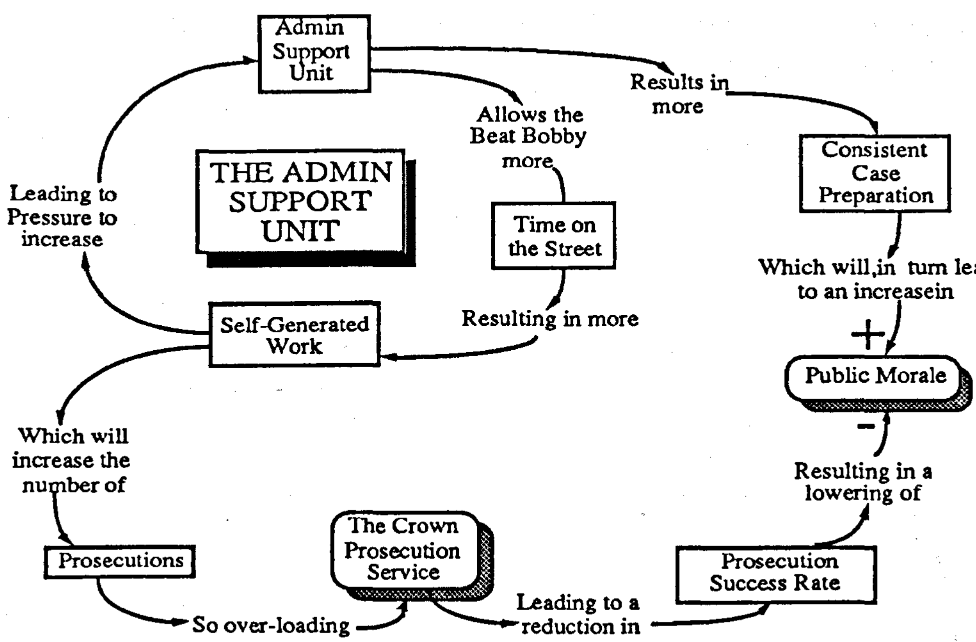
Critiques of SSM

Critiques of Checkland's work are not difficult to find.- see Boulding, (1982) and Burrell, (1983). Burrell , for example, observes that the method is not objective since it depends upon soliciting views from those immersed in the problem situation. Further, the introduction of a variety of *Weltanschauungen* "raises the issue of incommensurability of world view". Both observations combine in Checkland's choice of client, which is invariably from among those sitting high in an organization. In effect, Checkland seems to preach the value of *Weltanschauung*, but practise from only a narrow selection of world views.

This point concerning world views is important to pursue in the present context of IDA systems, since the differing world views are supposed, via different root definitions, to result in different idealized systems. One of two outcomes must follow: either a single world view is pursued, or several world views are explored. If only one, then there has to be a pre-exploration selection process; if more than one, then there has to be either post-exploration selection, or post-exploration reconciliation to produce a combined basis for deductions and recommendations about the problem situation. It has to be said that this selection / reconciliation process is not strongly in evidence. Similarly, bounding of the System of Interest (SOI), essential to sound analysis, is not well supported or explained by Checkland. An example from my own work will illustrate the point.

Exploring *Weltanschauungen*

I have been exploring, with several police forces, the implications of Home Office Circulars 104 and 105, concerned with civilianisation and with efficiency / effectiveness within the police forces. The exploration consisted in the first instance of appreciating the problem situation, which was elaborated in concert with a group of policemen of varying ranks. I simply modelled their views using the very high - level modelling approach favoured by Checkland. The first topic concerned Administrative Support Units (ASUs); these are small groups of policemen formed within police stations from personnel who, for one reason or another (often health) are not on the street. An analysis of ASUs might appear as follows.

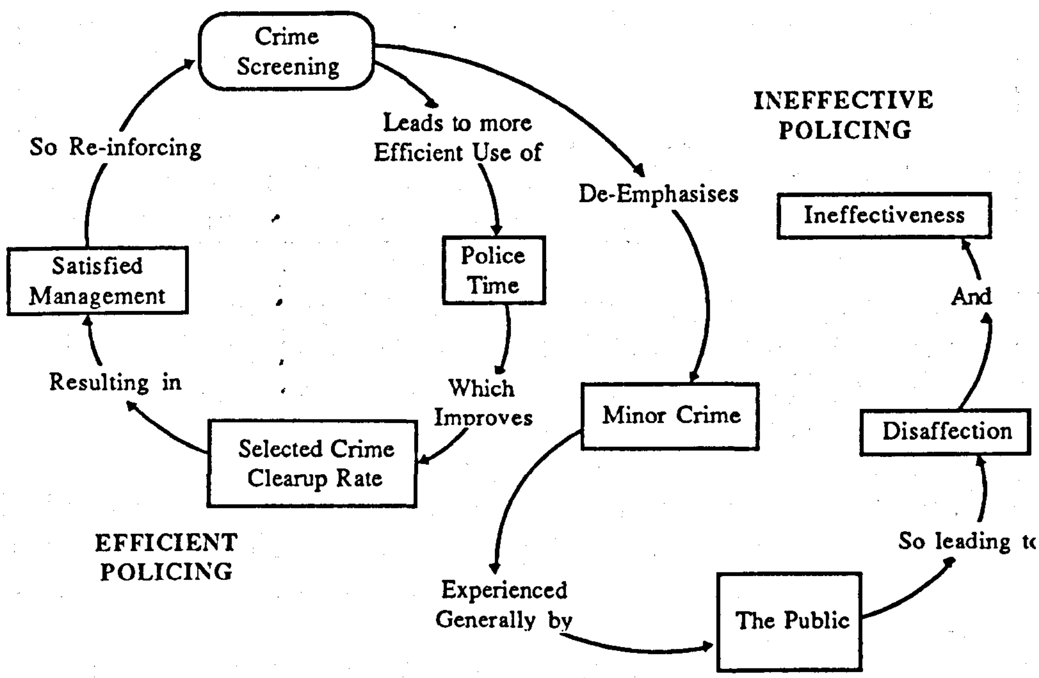


The anticipated result of introducing ASUs is that patrolling policemen will have more time on the beat because they will not spend so much time in the station doing paperwork. Instead, the paperwork aspect will be covered more efficiently by the ASUs who, because they will become practised, will be faster, make less mistakes and should develop a consistency of approach which will result in enhanced public confidence that "like crimes will receive like treatment".

The view of the police involved with ASUs over a period of time was that other effects soon took over after a promising start, which initially supported expectations. The beat copper, with more street time and no negative feedback (previously provided by the dread of "more paperwork"), generated much more work for the ASUs. In consequence, there was pressure to increase their staffing. More seriously, perhaps, the prosecution rate impacted the Crown Prosecution Service (CPS); as queues increased for prosecution, delays meant that prosecutions had to be dropped and, as the model indicates, police felt that this would have quite the opposite effect on public morale from that desired.

The analysis is simple, persuasive and easily understood - but whose view does it represent? Not the general police view, since I talked with only a small selection, chosen for their broad, intellectual approach. The particular group with which I was interacting were, moreover, representing not only their view of the ASUs but their view of the public's attitude to the police. Further investigation with my local police at senior level revealed concern that separating the beat copper from his paperwork and hence from continual contact with his suspect would result in bad policing because it could lead to irresponsibility. (In this officer's view, the measure of policing was likely to be the number of cups of tea that patrolling policemen were offered in their area - an interesting and provocative thought!) Interestingly, the analysis did show to the policemen concerned why it was that the ASUs were less effective than forecast, and they were quick to see that treating the police force as the System of Interest (SOI) while excluding the CPS was not likely to work

A second example from the same study shows a different angle



This example concerns police crime screening, in which a SOCO or Scene of

Crime Officer, assesses whether there is sufficient evidence to justify expenditure of resources. The analysis conducted with the police, shown in the above diagram, clearly showed a dichotomy in the mind of the police , with efficient use of police resources militating against effective policing in terms of satisfying the public. The analysis showed the police engaged in the analysis to be concerned and caring, prepared to see the views of others, indeed sensitive to those views. This contrasts rather sharply with a political view of crime screening as presented by Brian Sedgemore, M.P. who writes in the Hackney Gazette of 28th July 1989 as follows:-

> "I was astonished to learn recently that the Met Police don't bother to investigate 70% of the crime committed in places like Hackney. A local shortage of police officers, for which the Home Secretary is responsible, means that most calls for help from the public are not met.
>
> Except for very serious crime, local police operate a points system which in the jargon they use "screens out" most crime as being impossible to solve.
>
> We're supposed to praise this policy as being sensible because without clues it's obvious that crimes cannot be solved.
>
> Only a fool would fall for that argument. In the case of Hackney the clues don't exist because there aren't enough police searching for them, acting quickly enough, or on the spot to prevent them.
>
> My own view is that we need more resources, new structures that make local policemen more accountable to local people and a complete rethink of police policies.........."

In Sedgemore's view, the police are understaffed, yet he says that they "don't bother". He condemns crime screening, rightly or wrongly but then goes on to propose that the problem would in some way be solved by making policemen more accountable locally. Since this last view, be it reasonable or otherwise, is not traceable to his foregoing argument, it appears to be politically motivated, so presenting quite a different world view of police resource shortages.Indeed, a little thought suggests that his motivations could be radically different from his overt, expressed view.

The two examples support Burrell's view amply. Observing students attempting to employ Checkland's methodology, one is struck by the excellence with which they develop the so-called "rich picture" compared with the relative paucity of subsequent traceable and rational analysis of that picture. Could it be that Checkland's framework, straightforward as it seems to be, is a vehicle principally for those already expert in the domain of the problem situation to be operated within?

The SSM approach is applied by Checkland to management issues in the main; clearly he is intellectually powerful and, for him at least, and for students under his direction, the method shows results. At a conference held under the auspices of the Institution of Measurement and Control on 6th March 1986, Checkland expounded the following (paraphrased) view.

> Systems are not, as hard systems thinkers might suppose, like marbles in a bag which can be individually extracted, examined and replaced. Instead, systems resemble more the branches in a privet hedge: removing any one branch is impracticable, since it will be damaged in the process and cannot be replaced.

Again, Checkland's views are intellectually appealing, but is the insight he provides helpful in the context of developing IDA system design concepts?

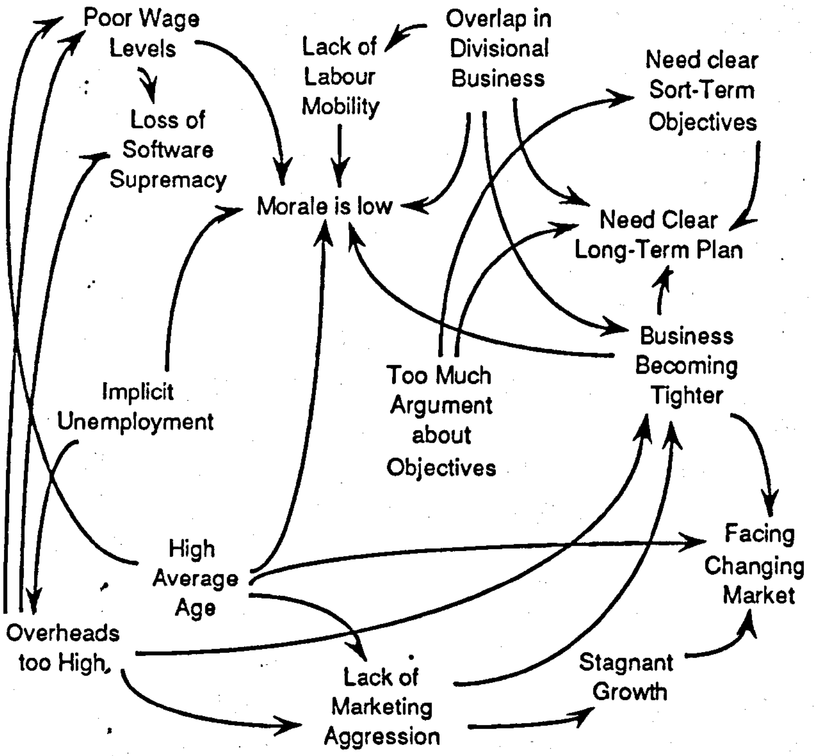
Wilson's SSM

Wilson is much more concerned with the hurly-burly of practical operations and engineering systems, to the point that his book smacks more of conventional, "hard"systems analysis than of pure "SSM according to Checkland". In both cases, there is a predisposition, born I suspect of the nature of SSM, to apply themselves to extant systems, rather than to conceive systems ab initio. Wilson is clearly an expert systems engineer, and his work encourages the view that bridges can be built from soft to hard systems concepts; Wilson, however, does not explain the intellectual jumps he frequently makes from soft to hard, and so his approach, worthy though it may be, lacks traceability and justification - two essentials in the real world, where system design concepts are formulated by sizeable groups of engineers, perhaps twenty or more - with some being less experience than others - and where many millions of pounds may ride on "getting it right"

Colin Eden

Another UK practitioner of soft methods is Colin Eden, lately of Bath University, presently at the University of Strathclyde. He is continuing to develop his methods, and he tends to adopt the facilitator role, acting as transparent "oil" between the members of a client team so that they may reach their own group consensus view without his views being imprinted.

The diagram shows an application of Eden's early graphic approach by the author to an organization in 1986. An analysis was undertaken by interviewing a selection of middle managers representing their views and linking them by discussion. Although this is not precisely the way Eden tackles the problem, the work was stimulated by his approach and it proved valuable. The connecting arrows were seen to point toward "sinks" and away from 'sources". So, morale was a symptom of the problem situation, not a source. Overlap in Divisional business was seen as a root problem, not a symptom. The diagram has been simplified from the tangled original, but it serves to demonstrate the process. Eden's graphical approach, simple though it may appear, can reveal to the user a system's latent characteristics, and in this respect it has a capability not implicit in Checkland's approach where such revelation is due to the analysts intellect rather than to the methodology. (Note: I use the term methodology, which surely means "the study of method" only because it is the current practice. Checkland's and Eden's work present methods, and are not methodologies per se)



COMPANY SELF-APPRAISAL

IDA Systems are designed and purposeful systems. One aspect of the two approaches illustrated so far is that neither method is implicitly goal-seeking. They are simply ways of unravelling some aspects of a human problem, usually a group problem, in a subjective way If, as a result, that problem is in some way alleviated, then the problem owners, the facilitator/or and the analysts have provided the creative components, not the methodology per se.

Ross Janes

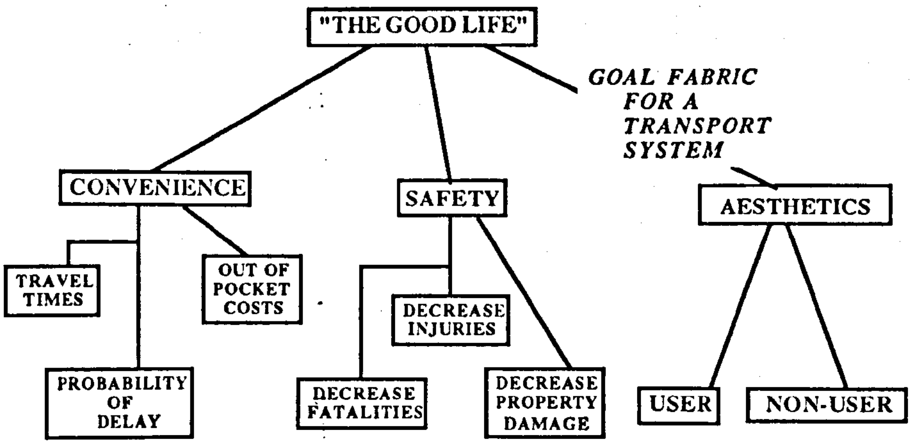
Ross Janes, in the Department of Systems Science at City University, London, adopts the facilitator role, not unlike Colin Eden, and his methods, although essentially subjective like Eden's, are quite different. Janes favours Interpretive Structural Modelling, a system of developing structure within a set of situation objectives, aims or factors by successive pair-wise comparison. The potential combinatorial explosion implicit in large-scale problems is contained using processor based tools to eliminate redundant comparisons and to manipulate and draw the resulting hierarchical networks. Janes' approach is interesting in the IDA systems context for two reasons:-

It is a group exercise, in which participation by the owners of the problem are of necessity involved, with Janes acting as an expert facilitator but, in principle, introducing no problem-related expertise which he might possess. Under his guidance, the group generate ideas about, and develop their own understanding of, the issues they face. They prioritize and rank the issues. And in the process of participation, a group consensus, and a group identity emerges. It is this last, more than any other aspect, which is the benefit of the Janes approach

The process reveals emergent properties in the problem which are not necessarily visible in the individual factors which are generated under Janes' guidance by the participating group. The structuring and grouping of the issue factors reveals architecture. See Toda and Sugiyama (1983) for a comparison of Q-Analysis, Interpretive Structural Modelling and Visual Q-Analysis. See Warfield (1974) for Interpretive Structural Modelling

Goal-Seeking

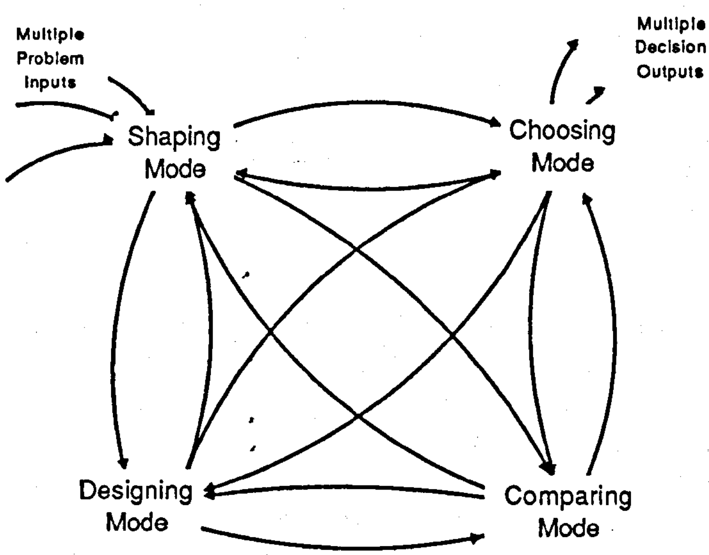
A method that purports to be goal-seeking is that attributed to Manheim and Hall (1967), and similar approaches to Multi-Attribute Utility Functions are noteworthy from Churchman and Ackoff(1957) and many others. Manheim and Hall develop a goal fabric, a hierarchical decomposition. This method has characteristics of visibility, traceability and - as will be shown later - also possesses the beginnings of some formality in that the decomposition is developed with Necessary and Sufficient Set concepts seemingly in mind



MANHEIM AND HALL'S APPROACH

The figure, which is an extract from "Systems Analysis for Engineers and Managers (de Neufville and Stafford, 1971), relates to the discussion surrounding the proposed purchase of Supersonic Transports at the time of Concorde. Evidently the goal fabric relates to value judgements and strategic decision formulation and once again, although it might well assist in that arena, its contribution to the ab initio design of IDA systems is not evident.

Decision-Based Approaches



John Friend, of the Institute for Operational Research and the Tavistock Research Institute, presents yet another soft approach to managing complexity. Complexity, he states, is not systemic, so it is better approached from a decision perspective than from a systems perspective. He draws together concepts, approaches and methods from both OR and from social science into a framework which, as the figure shows, operates in one of four modes - shaping, choosing, comparing and designing. His ideas are particularly appealing because they show an orderly progression toward a decision. Typical applications of his methods have been: County Council planning; LPG storage and distribution for the Dutch government; pollution control on the Rhine; and community health services.

System Dynamics

Last in the series of soft method overview is System Dynamics -. see Roberts, N.,(1983), Forrester (1961) Coyle (1977) and Lammers (1987)- a technique which is viewed with the gravest suspicion in some defence industry circles, owing to its potentially imprecise approach to modelling - it is, of course, that very imprecision which makes systems dynamics potentially useful for addressing softer issues. In use, the formulation of so-called "Influence Diagrams" precedes any numerical modelling, and it is my view, based on evidence that will be presented in subsequent chapters, that Influence Diagrams have much to offer in the development of IDA system design concept formulation. Judgement on the efficacy of the mathematical sections is reserved, although I have used the technique myself for modelling many-on-many helicopter-helicopter combat (see *Air Command and Control,* enclosed). Indeed, experience to date suggests that System Dynamics may be an effective approach to the thorny issue of bounding systems. Since Systems Dynamics is used extensively in one of the enclosed papers, *A General Theory of Command Control*, it will not be enlarged upon further here.

Postscript

An interesting parallel can be seen emerging between the developments of soft attitudes in Systems Science and in Anthropology, Horgan, J., (1989). Writing in Scientific American about Clifford Geertz,the noted U.S. social anthropologist, Horgan records Geertz as saying:"...the way out of the impasse is not to saddle anthropology with a specific ideological or political purpose, or to ape the rigorous methodology of the "hard" sciences......the solution lies in accepting cultural anthropology as, for better or worse, a literary enterprise - 'imaginative writing about real people in real places at real times'". Geertz has borrowed the term "thick" from the philosopher Gilbert Ryle, and has produced an essay entitled"Thick Description: Towards an Interpretive Theory of Culture". "Thick" invokes concepts, not of recording culture, but of interpreting it , to extract meanings from it that are, ideally, as complex and richly imagined as the culture itself.

Geertz sounds a warning to anthropologists which might well have been penned for systems scientists, judging by the gap separating systems science from systems practice. He insists that anthropologists should not let "epistemological hypochondria drive them into a formal academicism in which anthropology will become only for anthropologists"

The parallel between social anthropology and systems science in the context of IDA systems is far from illusory. Human behaviour may not be predictable at the individual level with any precision, but modern tribal and territorial influences predominate in the way groups of humans conceive, design, procure and employ IDA systems.

Summary

Systems Engineering has been developing steadily, if unspectacularly, over the last 10-15 years, driven by ever-more complex system design needs. A major influence has been the emergence of large scale computer-based systems and in particular geographically-distributed information and processing systems. These have necessitated a rapid development in software engineering methods, bringing a much-needed discipline to bear, but with the associated risk of misapplying software design tools at wider system level, to the prejudice of creativity at that level.

Systems engineering continues to be more of a faith or theology than a method based on sound, scientific principles. Its principles are treated as axioms, either not requiring, or not capable of, proof.

Recent years have seen the emergence of a variety of challengers to the conventional systems engineering approach, the so-called soft systems practitioners.
Their concentration on understanding the issues, though laudable, is underpinned neither by scientific proof - which they generally suggest to be impracticable - nor by a sufficiency of "track record". The arrival of soft methods has, however, considerably livened the systems debate and promises to produce more robust approaches to the conception and design of future systems.

3. CONSISTENT, UNDERLYING THEMES

3.1. Architectures

Structure, Balance and Flow

There are features about any good design which seem somehow "right". We are familiar with aircraft and ships being the "right shape", for example. So it is with systems, and IDA systems in particular. A well-designed system exhibits three characteristics: structure, balance and flow. This thesis is substantially about realizing the substance of those three words. First, structure.

Pervasive Architectures

Architecture is widely used term, implying many things to many people.Consider Bronowski's viewpoint, Bronowski,(1973)

> "I am making a basic separation between architecture as moulding and architecture as the assembly of parts. That seems a very simple distinction: the mud hut, the stone masonry. But in fact it represents a fundamental intellectual difference, not just a technical one. And I believe it to be one of the most important steps that man has taken, whenever and wherever he did so: the distinction between the moulding action of the hand, and the splitting, or analytical action of the hand
>
> "...The notion of discovering an underlying order in matter is man's basic concept for exploring nature. The architecture of things reveals a structure below the surface, a hidden grain which, when it is laid bare, makes it possible to take natural formations apart and assemble them in new formations. For me this is the step in the ascent of man at which theoretical science begins. And it is as native to the way man conceives his own communities as it is to his conception of nature
>
> "We human beings are joined in families, the families are joined in kinship groups, the kinship groups in clans, the clans in tribes, and the tribes in nations. And that sense of hierarchy, of a pyramid in which layer is imposed on layer,runs through all the ways we look at nature"

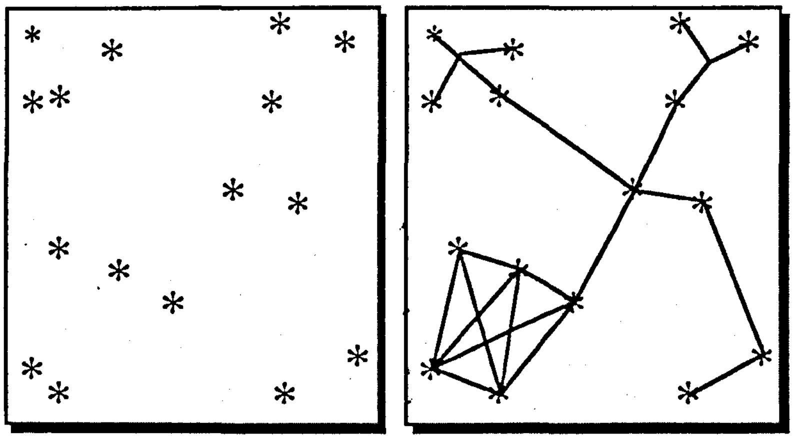
A view of balance, at a different level, is expressed by Edward Rubenstein (1989), Associate Dean at Stanford University School of Medicine:-

"How does nature encompass and mould a billion galaxies, a billion billion stars - and also the earth, teeming and exuberant with life? New insights into how nature operates come from parallel advances in particle physics and in molecular biology, advances that make it possible to examine the fundamental physical and biological processes side by side. The resulting stereoscopic views reveals a previously hidden, unifying logic in nature; its paradigm for construction.

"What nature does, in essence, is to make assemblies. It relies on the same template of programmed actions in each step of assembly along the way. Continuing sequences of assembly are the veins of evolution. Biological evolution is the result of natural selection operating on random variations. Physical evolution is a similar process of construction: a chain of chance associations from which new structures arise. Whether these objects survive or vanish depends on their environment.

"......Physical evolution and biological evolution are both characterized by common descent, natural selection and the eventual - and apparently inevitable - expression of symmetry,"

My view of structure and balance as essentials of good architecture embrace both Bronowski's and Rubenstein's views, as later chapters will reveal. But first, what is architecture in a visual sense?



What Is Architecture?

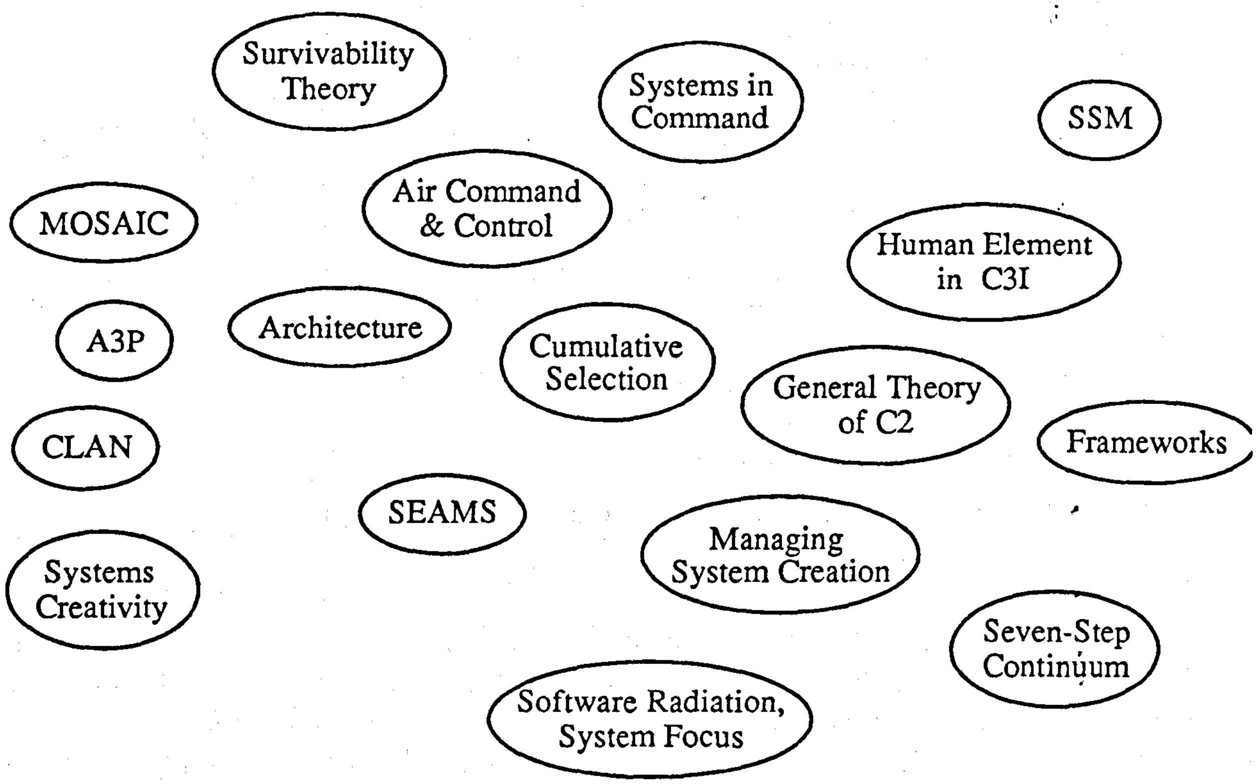
In the figure on the left is a set of entities. There is no interconnection. There is no architecture. On the right, the same entities have been connected in patterns which reveal two features: clusters of entities, and both cluster intra-connections and inter-connections. Now there is architecture. Architecture at its most elemental, then is clustering and linking. For artifacts, clustering and linking are purposeful, where purposeful includes aesthetics. Entities, at this level of definition, may be - physical chunks, activities, people, ideas even.

Architecture can be comprised of heterogeneous entities, so long as purposeful clustering can take place.

Clusters of papers

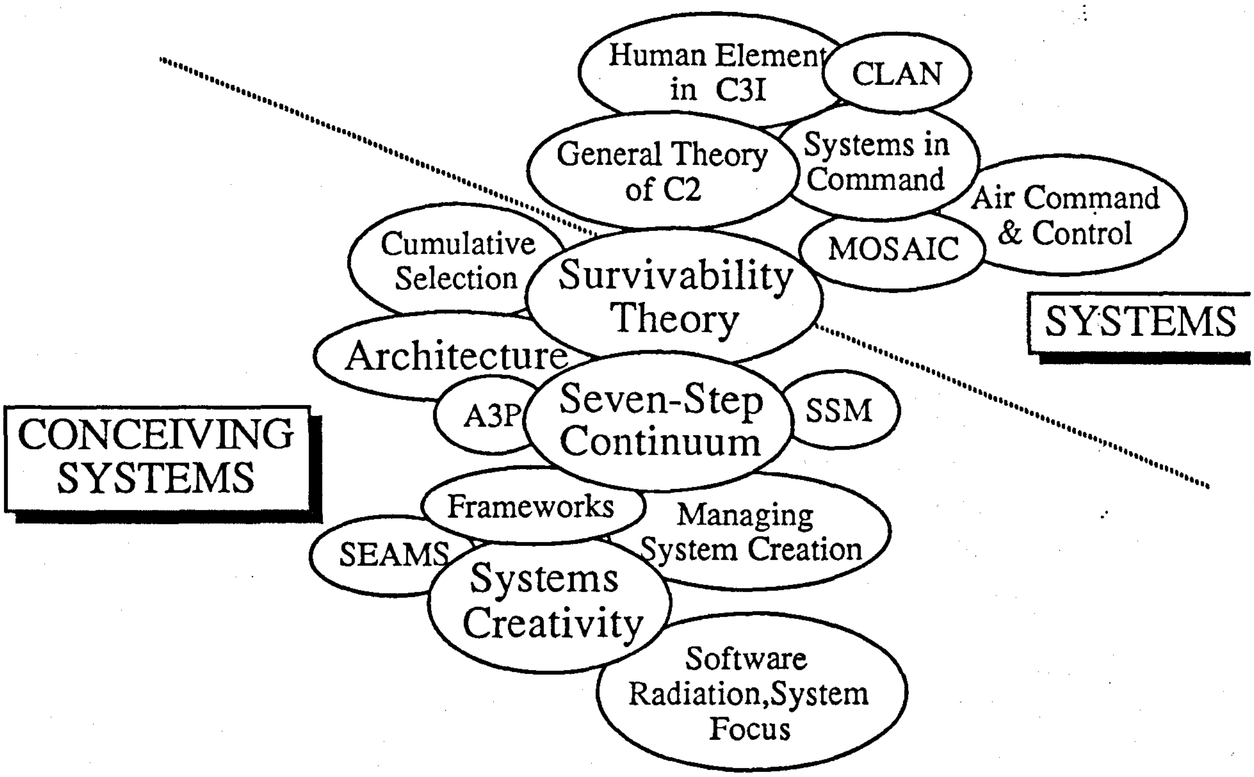
It would give a quite false impression to suggest that only physical entities can be clustered. As an exercise, I examined the relationships between the various papers which I have written over the last 10 - 12 years, with a view to clustering them At first glance, they appear to overlap and interconnect in a confusing fashion and so I clustered them, and some predominant themes, using the CADRAT tool which will be explained later (for now, it is simply a way of grouping entities according to the characteristics, degrees and strengths of their relationships, using a processor to reduce the very considerable workload)

The unclustered set of papers and themes may be viewed as follows:-



Document Organisation - Unclustered

Interesting although this set of disconnected titles might be, they certainly lack structure. As the author, and knowing the contents, of course I am able to connect the papers containing like ideas, but unfortunately many of them map to many of the others in a confusing mess. A more organized way is to construct a simple hierarchy of strengths of association so that, for example, *The Human Element in C3I* is strongly associated with *Systems in Command,* moderately with *A General Theory of Command and Control* and hardly at all with *SEAMS*. Executing this idea for all papers to all papers produces a 17x17 matrix of relationships, which may be clustered as follows:-



Document Organisation - Structured

The clustering process revealed that the papers were in two main groups,, linked by the underlying theme of Survivability, and that the two groups were about IDA Systems on the one hand, and about metasystems for conceiving and designing IDA Systems on the other hand. Until the application of the tool, these emergent properties of the papers - for that is what they are - were not so clear to me, the author although I knew, of course,that they were thematically connected in several ways. It is this ability of some methods to reveal underlying structure in seemingly-featureless complexity which makes them particularly valuable for the analysis of systems and the development of design concepts.

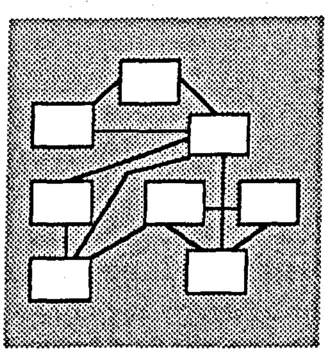
Binding and Coupling

Robert Lano developed concepts of coupling and binding, which provide another way to view structure. Entities bind together when they share common processes, interfaces, communication links, physical features, etc.Functionally-bound entities are candidates for physical grouping in order to reduce the end-system interface complexity.

Coupling is the degree of interaction between entities in different groups, and it is desirable to reduce coupling (of this variety) also to reduce complexity of interface and communications infrastructure. Lano was concerned principally with software and computer systems. I have developed his concepts and applied them to higher levels of architecture, including grouping of humans into organizations

The following diagrams illustrate the N^2 chart which is associated with Lano, but is now used world-wide. The N^2 chart is an incidence matrix comprising N rows and N columns - hence N-squared squares. Entities are recorded on the leading diagonal and interfaces between them occupy the other squares in the matrix. By convention, all outputs from an entity are on the row containing that entity, while all inputs to an entity are on the columns containing that entity. Each square (other than those occupied by entities) thus stands at the coincidence of an output from, and an input to, two entities; it thus represents the one-way relationship between those entities.

Lano observed that relationships, or interfaces in the software jargon, formed patterns when mapped on to the N^2 chart. Consider the following:-

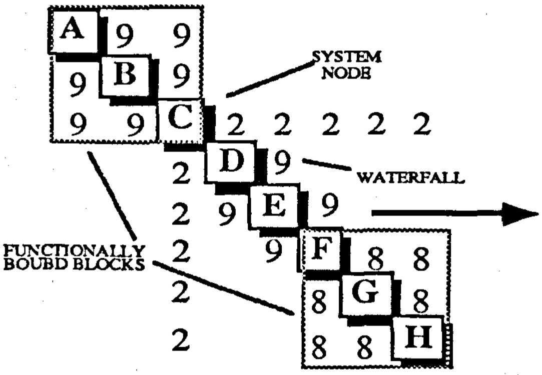


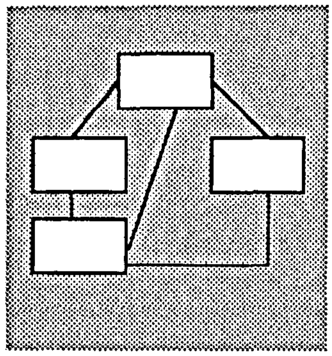
UNCLUSTERED

A	9		9				
9	C	2	9	2	2	2	2
	2	D				9	
9	9		B				
	2			H		8	8
	2	9			E		9
	2			8	8	G	8
	2			8	9	8	F

The N2 Incidence Matrix

The entities at the left have been mapped into the chart at the right, using an arbitrary numbering scheme. There is some pattern: evidently, entity C has some relationships with all the other entities and is some form of system node, but otherwise there is little to see





CLUSTERED

The Clustered N2 Matrix

The clustered matrix reveals much more pattern, however, with the emergence in particular of two tightly-bound functional blocks, detectable by the completely filled-in pattern of interfaces. These are candidates for physical aggregation, as shown on the right.

So, the unstructured first pattern has led to the structured, simplified, second structure. This is a powerful technique, which becomes very much more powerful when coupled with the speed of a computer, since there are some N! / 2 (factorial N divided by two) permutations of entities which, for a large matrix can become extremely unwieldy. I have applied the N^2 technique in most of the accompanying papers and to a variety of subjects. In particular, I have developed techniques for identifying poorly-formed patterns which reveal design weaknesses, missing links, vulnerable nodes,parallel nodes, source and sink nodes, over-tightly bound systems, different types of clusters, and so on. The enclosed paper *A General Theory of Command and Control* shows some of the variety of application, and more will emerge in Chapter 9

Push and Pull ideas

It is the practice at present to send information in IDA systems at the behest of the sender as opposed to the needs of the receiver. This approach has arisen for two reasons:

> First, it decouples the sender and receiver since the sender has no need to consult the receiver before sending provided, of course, that the receiver has some mechanism for accepting and temporarily storing the messages.
>
> Second, there is a need to ensure that the receiver's information is "fresh", that is that the information is up to date when needed. One way

to achieve this in dynamic situations is to send updates either periodically or whenever change occurs.

Seductive though both arguments appear, their practical consequence is a profligate consumption of communication bandwidth and data processing memory. The receiver may be undertaking a variety of activities, and he or she may not refer to the information received from a particular source for some time. During that time, the information may have been sent many times, on each occasion overwriting the previous one, and on each occasion wasting communication bandwidth. The simple alternative in many, but not all, situations is for the receiver to either call for, or enable the transmission of, information when interested in it. (This is how pain signals are accommodated in the human body - see *The Human Element in C^3I,* enclosed.)

I refer to these two philosophies as "push" and "pull" respectively. In *The Human Element...*presentation, I calculated that a typical IDA system could reduce its communication bandwidth by up to forty times and its data storage by a lesser, but still impressive, amount.

The signal entropy dilemma

A dilemma facing all designers of IDA systems concerns signal entropy. The Linesman-Mediator project referred to in the first chapter illustrates the point ideally. Consider an information-receiving and decision-forming system in which the sensors are remote; radar sensors are typical. The sensed information can either be reduced at the sensor, or conveyed to the IDA system. Reducing the sensor data at source saves on transmission bandwidth, but it is possible that some of the source data may be inadvertently lost in the reduction process.

For example, the source data may be sent as:-

Raw video, faithfully reproducing the originally received signals

Extracted plots, showing all the real target plots and the plots generated by supposed noise

Extracted tracks, made from a series of coherent plots, with the bonus perhaps of attached identity, height and vector data

Data reaching the IDA system may come from various locations and sources. It may be possible, by correlating data from, say, a radar source and an infra-red source - neither of which gives a positive target signal on its own - to deduce the presence of a target. Clearly, that would not work if the radar had passed tracks instead of raw video, but might work on the basis of transient plot data.

So, the designer is faced with a serious problem: to reduce signal entropy at source, so simplifying and reducing cost; or to maintain signal entropy to improve correlation and extraction prospects at sink. This dilemma is at the root of current concerns over so-called data fusion projects which seek to correlate (the word fusion is used very loosely) data from a large number of air, surface and sub-surface sensors with human intelligence

Russian Dolls and Fractal Thoughts

Peter Checkland is attributed with the statement that there are no such things as sub-systems. Whether the thought is useful to the would-be analyst, or not - I believe not - I agree with the statement. Systems do, however have parent systems, and systems do serve other systems, their siblings if you will. The reason for making the distinction is that, at each level in a hierarchy of systems, the current level contains systems; move up or down a level and there are still systems. I call this phenomenon Russian Dolls, since each system sits inside a larger systems which sits inside a larger system until we reach the Universe.

This concept is far from fanciful. In the enclosed paper *A General Theory of Command and Control* I attempt to apply classic control theory to the Command and Control process; the attempt is a failure. In an organizational hierarchy, each level exists in a "sandwich"; each person has both an immediate superior, or superiors, and immediate subordinates. At each level in the hierarchy an individual is at one and the same time a boss or manager and subordinate or worker. If one attempts to apply classic control theory at any particular level, that which is controlled is the subordinate. But the subordinate has exactly the same view of the world; he or she too controls his or her subordinates, and so on, not quite ad infinitum. Classic control theory has no means of dealing with this continual, hierarchical iteration, although The Law of Requisite Variety, Ashby (1985), does cast helpful light. The Russian Dolls problem is one of the reasons why analysts find it very difficult to analyse organizations and management structures and to produce sensible models.

The mathematics of hierarchies is also expounded in *A General Theory of Command and Control.* If we consider an N^2 chart, and exclude the entities themselves on the diagonal, then there remain $N*N - N$, or $N(N-1)$ squares or interfaces. As N increases, then, it appears that the number of interfaces / relationships rises approximately as the square of N. If we now impose on the N^2 chart the concept of hierarchy, many of the squares are left blank owing to the limitation imposed by the pyramid organization, which favours communication within the local group and upwards and downwards, but not between groups at the same hierarchy level. Thus the number of interfaces does not rise as the square of N in such cases

Consider, for example, the case where each superior had three subordinates. Starting with the top person, the number of people in the organization at each successive level is 1, 3, 9, 27, 81.......and so on. Consider now an N^2 chart for those first three levels of hierarchy. Drawing this out on an N^2 chart results in 13 entities (1+3+9), giving 13x13=169 squares from which we subtract the 13 entities themselves to leave 156 interface, or relationship opportunities. If we assume a strict, pyramid hierarchy,then number of occupied squares is only 48 (see figure 10 of *A General Theory of Command and Control),* even allowing for full peer communication at each management level. So, instead of a convenient square-law rule concerning the growth of relationships with hierarchy levels, we have a law which lies somewhere between unity and two

(For the mathematical record. If we define the full N2 chart as containing NxN relationships, by the simple expedient of allowing entities to have relationships with themselves, then the organizational pyramid described above contains 61 (48+13).interfaces rather than the full 169 (13E2) This number 61, corresponds to 13E(1.602715). The index , 1.602715, lying between one and two, is fractional.)

Taking the concept, in which the expansion index for relationships lies between the linear and the square, together with the Russian Doll concept of iteration, raises thoughts that human organizations may tend to be fractal in nature. (Fractal geometry displays similar characteristics, of course - hence the term fractal, indicating fractional indices of expansion). For the purist mathematician, this would be difficult to accept because of the indeterminate nature of fractal structures; evidently, human organizational pyramids do not go on for ever in greater and greater detail. Accepting the limitation, one is none-the-less struck by the similarity and the potential richness of the resulting patterns which hierarchy can exhibit.

3.2. Information-Decision-Action Systems

Classification and Membership

IDA systems are human activity systems distinguished by the following features:-

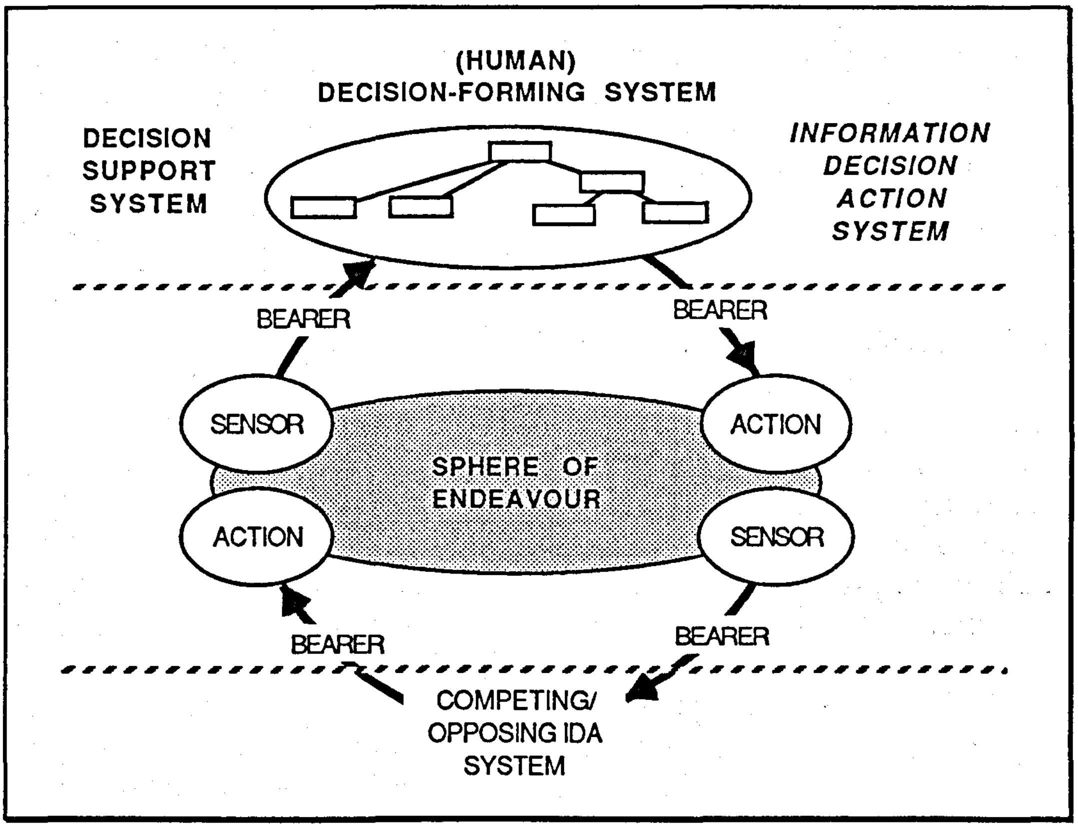
Existence with competing or hostile systems and addressing spheres of human endeavour

Close association with sensor and action systems, together with which, IDA systems form a closed loop acting within the sphere of endeavour

Sensed information providing an input for rational judgement

Intelligent decision taking

Decision-based action instructions



Generally, IDA systems comprise human organizations grouped to undertake tasks supporting human decision-taking and supported by information collection, analysis, handling, presentation and communicating systems. An IDA system need not, essentially, contain technological devices for these supporting purposes

Process control systems are not IDA systems; there is no human decision-taking. Process control systems may, however, form subordinate systems within an IDA system.

Management information systems are not, of themselves, IDA systems since they lack the action-ordering element. Management information systems may form part of an IDA system when integrated with decision-taking and action-ordering elements

The Nature of IDA Systems - Human Decisions, Information Support

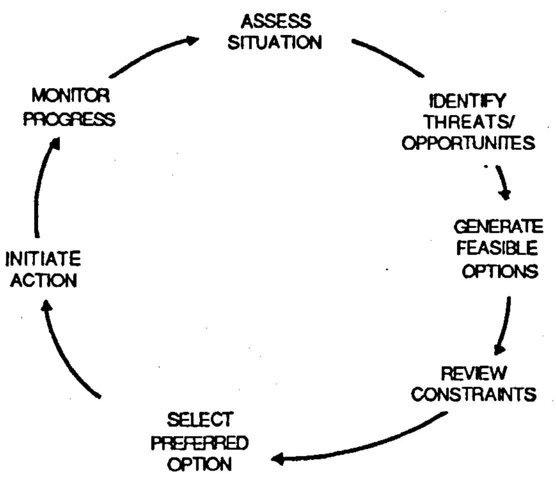
In IDA systems, then, a key characteristic is that decisions are taken by the humans; they may, if they wish, ignore all the advice and sensor indications presented to them and make their own decisions. A man driving a car is the focus of an IDA system when a dog runs out into the road; he senses the situation, decides on his course of action, initiates the action and he and his car change vector (or not) according to the chosen decision.

In the kind of IDA system of particular interest, groups of people act in concert to arrive at decisions. In some cases, the amount of information collected, handled, analysed and presented requires the support of processors, sometimes of very large processors, with sophisticated control and display systems. None-the-less, the basic rules apply; humans making the decision, with decision-support in an advisory role.

IDA systems can become very difficult to conceive, design, develop and use. One reason for this is because the human activity system, in which the machinery is intended to form an intimate part, is extremely robust. Humans respond to changes in situation and environment extremely well - indeed it is their special capability above all other species. A decision support machine that is designed to satisfy particular situations and environments cannot adapt as can humans under such dynamic conditions. Many of our present-day command and control systems fall into this category of rigid design; as a result, they are obsolescent by the time they are delivered to operate in a peace-time environment, and would be white elephants in conflict.The need is for systems that can support the robust human activity system, be a part of that system, and yet not inhibit its essential flexibility.

The Decision Circle

Core to IDA systems is the concept of decision-making. I contend that all rational decisions can be represented in the following cyclic manner



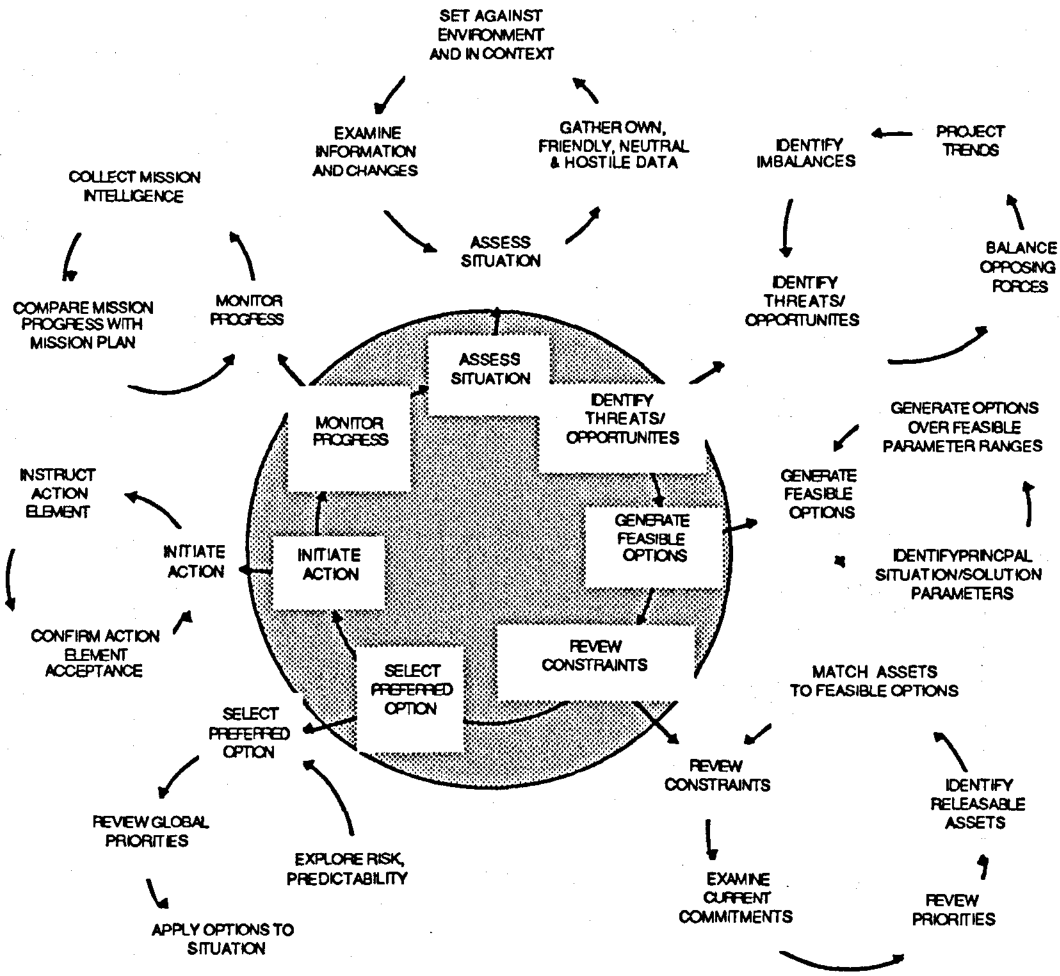
The decision circle is, to the best of my knowledge, my own invention although similar concepts do exist. Wohl (1981) offers the SHORE paradigm, which stands for Stimulus,Hypothesis, Option, REsponse. Clearly SHORE is an abstraction of my Decision Circle to some degree but excludes the essentially cyclic nature that I have observed in IDA systems in operation

A variety of other pundits has produced versions of the decision process, generally in the field of conflict; Mayk and Rubin's (1988) *Paradigms for Understanding C^3,Anyone?* produces twenty variants. The title "decision circle" has also been used by others, including Coe and Dockery (1988), but in the substantially different context of conflict, where the circle is in two halves, each representing the decision processes of their relative opposition.. The beauty of my decision circle illustrated above is its simplicity; it applies not only to a car-driver faced with a dog running out into the road,but equally to a football manager with a choice of substitutes and ten minutes to the final whistle, a police chief facing an unruly crowd, and so on.

IDA systems often assign different people, or groups of people, to undertake different elements of the decision process represented by the Decision Circle. A typical distribution of activities might be as follows:-

Assess Situation	Intelligence & Operations
Identify Threats	Intelligence
Identify Opportunities	Intelligence and Operations
Generate Feasible Options	Operations and Intelligence
Review Constraints	Logistics & Engineering
Select Preferred Option	Commander
Initiate Action	Operations
Monitor Progress	Intelligence & Operations
Assess Situation	Intelligence and Operations
Etc	

It is then possible to take each one of the above activities forming the Decision Circle and use it as a nucleus to develop its own activity circle



Whereas the N^2 chart looked at structure, the Decision Circle develops flow in IDA systems design; groups of people are purposefully organized around the centre circle, and around each of the supporting circles. The various support circles may be thought of as cog wheels, meshing with the central decision circle. Following the analogy, note that the direction of arrows in the supporting circles must be anti-clockwise to maintain the rotation of activity and that the speed of rotation of the central circle will be no faster than that of the slowest supporting circle. The coupling of decision circles operating at different levels in a hierarchy, and the mapping of decision circles on to N^2 charts, are dealt with in the enclosed papers, *Systems Creativity* and *A General Theory of Command and Control.*

I discussed earlier the subject of civilianization of police posts. That topic can now be reviewed in the light of the Decision Circle above. The police force would be ill-advised to prejudice the inner circle by civilianizing any of the inner circle elements, since these are at the first level of decision-making; if serving police officers do not make the first level of decision for normal operations, then they no longer constitute a police force in the general sense of those words, since they have no inalienable control of over their own decision-making. Their degree of concern over civilianization diminishes as they consider first the outer ring of support circles and - not shown - the rings beyond that, since the support circle concept can stretch to several layers of support circles.

Boundaries to IDA Systems

Bounding an IDA system, like any other system, is one of the most difficult tasks facing the would-be designer. *A General Theory of Command and Control* shows a series of system dynamics influence diagrams at successively increasing resolution, which shows the following inter-linking systems:-

A.An Information-Collecting and Analysing System
B.An Action-Monitoring System
C.An Option Generating System
D. A Decision-Forming System
E.An Order-Issuing System
F.An Action System
G. A. Logistic System
H. A Higher-Authority (Political?) System
I. A Competition System

Within the above systems are contained systems. For example, the action system has its own IDA system for local management of operations and its own replenishment, servicing and engineering systems; the logistic system, too, has its own IDA system. The extent of the IDA system - how many of the listed systems are contained - is subject to convention. Usually, the IDA system does not contain the action elements under its control. Thus, items A and B on the list might well be included in an IDA system under the sub-division of Intelligence, items C,D and E might be included under the heading of Operations and Operations Plans, but items F to I would be conventionally be excluded. In other cases, however, the action elements are so intimately bound up with their control that the action system becomes a part of the overall Information-Decision-Action System - a fighter aircraft may present an example of this latter situation.

3.3. Creativity

A third consistent underlying theme to my work has been the subject of creativity. This thesis is essentially concerned with that process, and so are all of the enclosed papers in one sense or another. Explicitly on the subject are *Managing Systems Creation* which will be discussed later and *Systems Creativity*. More than any other papers, these two papers have generated widespread interest with copies being called for particularly from eastern Europe. Systems Creativity, which amongst a variety of concepts, including decision circles, creating creative environments and state transitions within organizations, develops the so-called *Principles of Creativity*. These Principles are repeated below

Highest Level of Abstraction Breadth before Depth Level at a Time Disciplined Anarchy Decomposition before Integration Functional before Physical Tight Functional Binding Loose Functional Coupling Functional Migrates to Physical

The principles, collected over the years from a variety of sources - including "doing it wrong" - are worth a closer look.

Highest Level of Abstraction This is the cardinal rule when approaching a new problem situation or design concept. It is essential to gain a panoramic view of the situation which removes the confusion of detail. Only by actively and assiduously pursuing this principle, is it possible for the analyst to see his own prejudice and rise above it.

Breadth Before Depth and **Level at a Time** Similar concepts aimed at successively reducing the level of abstraction in an orderly and coherent fashion. These principles oppose the tendency to concentrate prematurely on parts of the design problem in depth, to the exclusion of other parts and of the overall design. Breadth-before-Depth implies covering the whole problem "in the round", while Level-at-a-Time implies that each level of abstraction / decomposition should be completed before descending to the next. Together, they provide an ethic of orderly progress

Disciplined Anarchy. It is essential to create an environment in which new ideas may flourish, rather than "hit the cutting room floor" before being given a real chance. Ideas are generally the province of the young - in mind if not in body - and are incompatible with strict, authoritarian control. Timescales and budgets have to be met, however, and so the principle of Disciplined Anarchy emerges - set periods when creative juices are encouraged and negative thoughts are banned. Organized brain storming and idea writing are among the many approaches to achieving the objective

Decomposition before Integration. This principle, taken with the first, proposes that creativity requires the examination of the component parts of a solution prior to their being grouped and assembled. Without such decomposition, often a major task, it is not possible to see how best to combine the various elements of a solution

Functional before Physical This principle has already been mentioned. It is essential to be concerned with the purpose of the solution before becoming embroiled in its form.

Tight Functional Binding and **Loose Functional Coupling.** Both principles have already been discussed, and are attributable to Robert Lano

Functional Migrates to Physical. This principle is self-evident, but can be overlooked in the heat of concept formulation. Eventually, all the functions are going to be realized in some physical form. The grouping of functions for functions' sake is not the end of the matter; functions become bedfellows for physical reasons, too. The archetypal example is, perhaps, the I/O (input/output) for a processing system: functionally, I and O are at opposite ends of the process; physically, their construction is generally very similar and hence they are often grouped together.

It is interesting to note that the Principles of Creativity find echoes in other spheres, notably art and music. Discussions with painters in oils and with composers indicates that their approach to composition is very similar, suggesting strongly that creativity in the systems arena is linked to creativity in other spheres.

3.4. Survivability

Survivability (of Performance) as a topic is to be found throughout the enclosed papers. I believe it to be a fundamental design concept for all IDA systems and for many other systems and artifacts, too. "An animal that is adapted to its environment is first and foremost organized with a view to its own survival", Guilbaud (1959) states in discussion about Ashby's cybernetic"homeostat".

Why is survivability so important? It forms a basic set with Performance and Availability (of Performance). Together, Availability and Survivability underwrite or guarantee Performance, the one against internal failures, the other against external threats. Availability is well-developed as both an engineering and a scientific discipline. Survivability, however, is not. There is no established science or mathematics of survivability, and yet our various IDA systems are threatened continually by terrorist attack, by electro-magnetic attack, by enemy missiles or groundfire, and so on. Those who have operated in situations of danger are generally agreed that survivability is paramount, even at times beyond performance itself. I address the subject in the enclosed brief *System Survivability Science* which is one of the very few attempts to address many of the issues.

Damage Tolerant Architectures

For IDA systems, as usually bounded, there is a wide variety of strategies and mechanisms that can be employed to achieve requisite levels of survivability - requisite, that is, in relation to some supposed threat environment.

At the visible, physical level, the choice is generally concerned with hardening (armour, for example), mobility (which sits uncomfortably alongside armour due to the latter's weight) and replication such that damage to one part does not destroy the whole.

At the human level, much more important in practice than physical aspects and yet often poorly addressed, there arises the concept of autonomy. Survivability could, perhaps, be exemplified by the World War II Japanese soldiers who emerged from the jungle up to twenty years after the war had finished. These soldiers survived because they were self-sufficient and because they had clear instructions and the will to follow them. Misguided? Perhaps. Survivable? Certainly.

In IDA systems, however, there is a tendency to control from the centre; central control is comforting to the controlling party, who then knows all that is going on. Should the central control be eliminated, or simply disconnected, the system will no longer operate, despite the capability of the action elements which may be in complete working order. Additionally, central control requires communication and takes time - the concept results in slow, authoritarian behaviour.

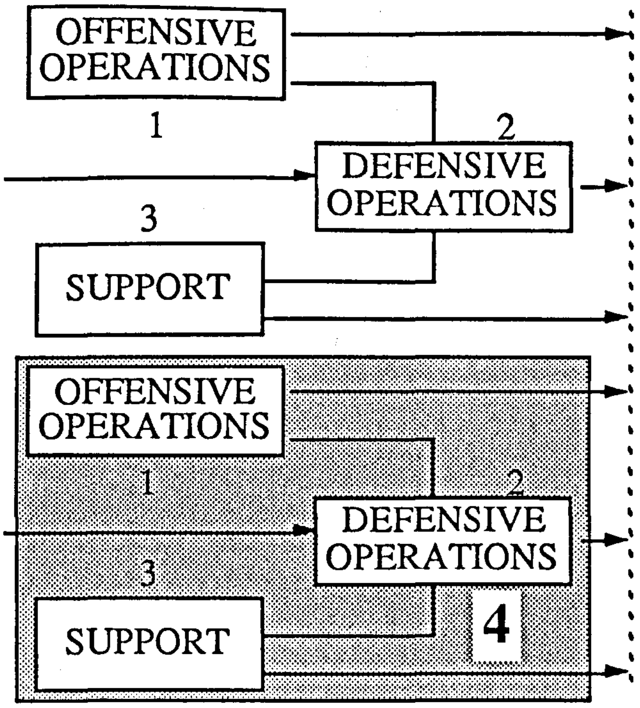
Happily, we humans are designed differently; when treading on a thorn, we do not convey the pain sensation from the foot to the brain and then send a motor signal from the brain to the foot. If we did, the thorn would have penetrated by the time we had responded. Instead, we have a sensor-motor crossover in base of the spine - the spinal reflex arc - which short-circuits the communication network to the brain, and enables us to lift the foot quickly. Of course, the brain also receives the pain sensation, often after the reflex has removed the foot. And so it should be in good IDA system design. There are many responses which should be delegated to lower levels of authority both to speed overall responsiveness and to enhance survivability. The enclosed paper *The Human Element in C3I* addresses these issues in some depth

Systems Survivability Science is a short brief, prepared for an international audience, concerned with a very large, classified international defence project. The brief contains mathematically robust approaches to the quantification of survivability in C^2 and C^3I systems, and is part of a continuing programme of research in this area

3.5. Optimal Design

Optimizing Air Command and Control

A further consistent theme is that of optimal design, a Holy Grail for system designers. The enclosed paper *Air Command and Control* presents the following diagram which illustrates the issue:-



AXIOM: "A system comprised of elements which have been separately optimised will not itself be optimised"

COROLLARY: "An optimised system will contain elements which are not, of themselves, optimal"

REASONS: A. "4" is a new function. Its optimisation has not been attempted

B.The interfaces between the elements 1,2 and 3 have not been considered in their context of 4

SYSTEM DESIGN AXIOM AND COROLLARY

The conclusion drawn in the paper, apparently obvious, was as follows:" Improving the effectiveness of air power requires a total system approach rather than a function-by-function development". Strangely, this concept was - and is - avante garde in the air power business. The idea that three "tribes", Offensive Operations, Defensive Operations and Transport Operations, should be treated as one is difficult to accept.At such high level, the application of such self-evident concepts are difficult, if not impossible. Optimizing approaches that do not cross tribal boundaries, or better still that appear to preferentially benefit an already-dominant tribe, have a much better prospect of working.

Tradeoffs

Optimal design supposes an objective, or set of objectives, against which the design may be optimized. In practice, there are many, mutually-incompatible objectives for any complex system and a means of trading between them is necessary

The usual practice is "weighting and scoring", in which optional designs are compared according to a set of criteria, or measures of effectiveness, by scoring numerically. The criteria are weighted according to perceived relative importance, weighted scores are totalled and the highest weighted score sum decides the winner.

Weighting and scoring is invalid, for two main reasons;-

The measures of effectiveness are generally different dimensions, such as "operational performance" and "ease of maintenance", or "utility" and "cost". It is clearly invalid to add scores algebraically for criteria of different dimensions

Scoring and, particularly, weighting, are subjective. The view on which criterion to weight and to what degree.depends on the perspective. An engineer will weight ease of repair, an operator will weight utility as he perceives the need, and so on.

There are many valid ways of providing valid multi-variate analysis. The method I have developed and used successfully is Rank Matrix Analysis (RMA). RMA also works on a matrix format, with design options as column headings and criteria as row headings. For each criterion in turn, the design options are ranked in order of preference. It is generally easy to state whether A is better than B, provided there is no need to state by how much. Genuine inability to rank results in allocation of equal ranks. The resulting table of ranks is assessed statistically to establish whether the pattern of ranks differs significantly from a random set, and a Coefficient of Concordance is produced If the probability of being random is, say, less than 5% then the table is meaningful and the column (design option) showing the minimum rank sum is selected. The process is illustrated in the enclosed paper *Managing System Creation..*

RMA cannot be said to be entirely objective, because it still uses data solicited from system designers and operators. The manner of the data categorization and the manner of its processing are, however, reasonably resistant to challenge.

Comparing Weighting and Scoring with RMA is interesting. Generally clear winners and clear losers emerge the same by either method, while the group of design options in the centre change in order of preference. For close-run choices, using both methods is preferable. Users invariably feel happier with the (invalid) Weighting and Scoring approach; they are happy enough with ranking but not with the concept that criteria are not weighted in RMA.

Summary

The chapter has presented the principal underlying themes in the last 10-15 years of work.These have been:-

• Architecture, in all its systems-related aspects, including decomposition, synthesis, binding and coupling, push and pull, signal entropy, and many other aspects, viewpoints, methods, tools and techniques.

• Recursion in architecture, as exemplified by the Russian Dolls paradigm, with its fractal connotation

• Information-Decision-Action Systems and, in particular, Command and Control Systems, with their human activities of structured organisation of information and of decision formulation.

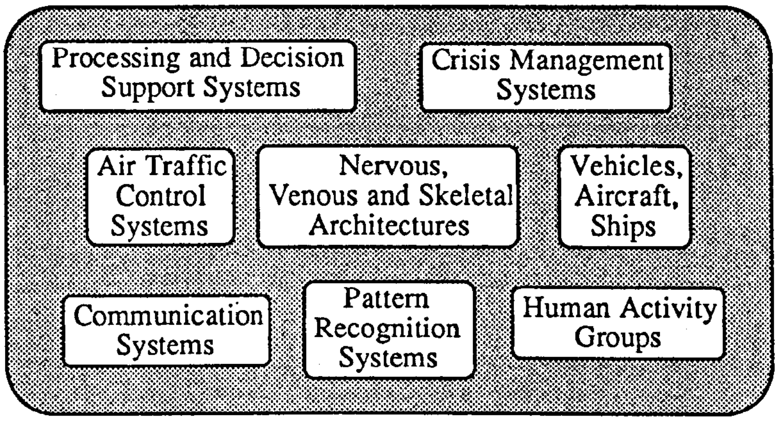
• Creativity, how to promote and nurture it

• Survivability, the overlooked but fundamental characteristic required of every system

4. THE NATURE OF THE ANIMAL

4.1. Isomorphisms

The isomorphism relating humans and the systems that they design has been mentioned and is enlarged upon in the enclosed paper *The Human Element in* C^3I. The degree with which we replicate ourselves in our designs seems to surprise us, perhaps because we are often unaware that we do it. But do it, we do.



Architecture Isomorphisms

The figure shows just a few of the systems in which isomorphisms, in the broadest sense perhaps, can be envisaged. Isomorphs are becoming particularly evident with the advent of neural networks, for example, but all of the examples in the figure share architectural features, with central nervous systems, sensors of external and internal activities, energy distribution and so on. It is, of course, possible to see isomorphs where none exists - that would be a human thing to do, too, and we humans would perhaps be unable to detect our own fallibility. On the other hand, if the process of detecting isomorphisms is employed as part of a cross-development of ideas, then only good can come. We need a master pattern to ensure that comparisons are fruitful; that master pattern, by virtue of the elegance and success of its evolution to success, is our human frame and our human society

IDA systems never work fully when first employed; purchasers of technological solutions are inevitably disappointed. But most IDA systems, intimate mixtures of human activity with machine support, evolve towards being valuable in use. Evolution is a much more powerful stimulus to performance than a design team's cerebral power, at least in this respect. The human frame appears to be the most evolved on the planet; it is certainly the most successful. It seems eminently reasonable to use it as a working master plan for IDA system design.

The first observation to be made about human design is that it does not excel at anything. It is not the fastest, strongest, biggest, smallest, most agile...... No, instead it is a master of compromise. And the compromise in the individual has often been in favour of the group as the following table, taken from the presentation which accompanied *The Human Element in C³I*, will show;-

FEATURE	COMPROMISE	LOSE	GAIN
Bipedalism	Reduced Running Speed, Manoeuverability	• Survivability (Flight) • Energy • Olfactory Tracking	Free Hands, Early Warning from Raised Remote Sensors, Better Intra-Group Communications
10 - 15 year Nurture	Limited Pre-Birth Imprint; Infant Helplessness	Group Mobility and Survivability (Flight)	Adaptability, Flexibility, Social Evolution
Hardening	Limited Cranial, Spinal and Thoracic Protection	Survivability (Fight)	Mobility, Survivability (Flight)
Distributed Processing (Spinal Reflex)	Some Processing Occurs Outside of the Spine	Central Control	Speed of Response
Weight / Size / Volume	Not the Largest / Fastest / Strongest, etc, of animals	Individual Survivability (Fight)	Larger Group Size, Greater Group Survivability
Sensors :- • Small Eyes	• Limited Night Vision	• Ability to operate at night	• Simpler, more Survivable Sensor by Day
• Eyes and Ears	• Colocated and Passive	• Direct Range Information	• Ease of Correlation / Avoidance of Detection
• Cone of Foveal Resolution	• Narrow Cone Supported by aural and peripheral flicker clues	• Wide Field of View	• Faster Processing / Speed of Response

Human Design Compromise

So, we may look inside the human frame for design concepts, perhaps of a harder nature, but it is to human groups that we must look to determine the softer issues. Since IDA systems deal largely with the more primitive aspects of human behaviour - warfare, terrorism and the like - it is reasonable to look to the more primitive side of human behaviour for insights into IDA system design concepts

4.2. Anthropomorphic Designs

The Human Element in IDA Systems

The enclosed paper has been mentioned several times; it was, when written, something of a venture since it was not, and is not, usual for practising system designers to present such papers in front of a largely military and industrial engineering audience. The paper starts by viewing the human body through a command and control system designer's eyes and observing the built-in redundancy and survivability features in the design, from the hardening of the skull to protect the dual-redundant central processor to the splendidly designed rib-cage, with its dual functions of flexible protection of the organs and lifting of the lungs to draw in air.

The paper employs Professor Patrick Wall's Pain Gate Theory to design a digital processing system which would unquestionably operate, but for which there is no precedent in conventional design. The paper picks up on the eye-brain combination and sensor correlation between eye and ear, which concept the author had previously used in the *Heuristic Intelligent Threat Assessment System (HITAS)* design - see below. The paper ends also raises the topics of tribal behaviour and territorial imperatives, which will now be enlarged upon.

Simian Social Behaviour

In "The Quest for a C^3 Theory; Dreams and Realities", Levis and Athens (1988) present the view that a theory of C^3 (Command, Control and Communication - IDA but with the communication aspect raised in perspective) is unlikely to be attained owing to system complexity.

> "A major source of the complexity is that many human decision makers are integrated in the C^2 process. These humans - intelligent, active persons - are not just components or users of C^3 systems, but an integrated part of the process itself - on both sides, ours and the adversary's"

Clearly, the statement has substance, but it is not helpful. I believe that it is possible to predict and anticipate human behaviour to a degree if we understand our primitive responses. Why, for example, do we arrange our offices with our desk opposite the door and our backs against the far wall? Territorial imperative. We wish to create an area which is ours and into which others may not encroach, at least without our knowledge. Bosses and management scientists know about dominance and territorial imperatives, the one perhaps intuitively and the other through study. Social scientists and ethologists in particular have developed an understanding of the way we behave at the social interaction level. Subjects of direct relevance to IDA systems include limb discipline, touching and dominance, triad (three-person groups) instability, boundary marking, the instability of larger groups of people and non-verbal communication

Consider and IDA system operations room with, say, ten personnel and their relative information sources. How should the room be laid out? Should the operators face outwards, with their backs to the centre of the room, or inwards so that they can communicate directly with each other? This is a more interesting example, and not simply resolved. If there is a real-time controlling operation in which the personnel communicate principally through their screens and controllers, then they have to face them. Facing outwards will inhibit direct person-to-person communication and presents a submissive rear aspect to a supervisor, making the position of the operators subservient. But if they are required to offer advice, to come to consensus views, then they should be able to communicate both visually and aurally, by voice, by body language and, yes, even by pheromone.In such situations, which are very common, the attempt has been made to have screen operators supplied with low-level screens, so that they may face inwards for direct personal communication and see their own screens at the same time. Such solutions are limited; body language is screened by the equipments and there is little room at the centre for a controller to occupy.

This seemingly simple consideration of which way to face significantly characterizes IDA systems. Power station control rooms and air traffic control rooms almost invariably are of the "glued to the screen" variety, such as we see on TV for NASA space shots. Strategic command and control, on the other hand, is more likely to be of the "round-the-table" variety.

In all organizations of more than a few people, there develops hierarchy and division. The quotation from Bronowski given at the beginning of Chapter 3 illustrates my view admirably. We develop into family groups, clans, tribes and so on. This tribal behaviour is very apparent once looked for. The Royal Air Force, of which I was once a proud member, and for which I have an abiding affection is, like all three services, highly tribal. First there are the clans. The principal clans are:-

Operators, or commissioned aircrew
Engineers
Suppliers
Ground Controllers
Administrators

Then, of course there are family groups within each clan: the Operators have pilots, navigators, air electronic officers and others as families, with the pilots occupying senior family status. In the other direction, the clans aggregate into tribes. RAF tribes include the following:-

Bombers
Fighters
Flying Trainers
Groundcrew Trainers
Etc

Each tribal group demands tribal loyalties, which are reinforced by tribal-specific custom (the open, top uniform button for the fighter pilot?) and inter-tribal competition. The tribes, of course, have a variety of clans within them. A member of the Engineer Clan within the Fighter Tribe, as I was for some time, forms binding relationships within his tribe and his clan, and to a lesser extent with clan members in other tribes. Efforts are always being made to promote splinter clans. These often take the form of clubs or groups to which membership is limited by some clear demarcation which excludes rather than includes. Masonry is an overtly primitive clan and tribal activity, for example, making "free masonry" something of an oxymoron.

Should the reader feel this to be fanciful, there is ample evidence of the fiercest internal competition within almost any organization of significance, with members of different tribes often more intent on maintaining tribal position than in addressing the external threat / competition - many businesses operate in this manner. The behaviour of unions in seeking solidarity against the bosses is tribal. The divisions between different unions in the same industry is more clan-like.

Once groups of humans are viewed in this old, rather than new, light, their group behaviour is both more understandable and predictable. While I do not suggest that an individual's response is predictable, I most certainly do suggest that groups of people do behave predictably, by jealously guarding territory and by competing with other clans and tribal groups. So deep is the instinct that clan loyalties can be observed in people who left the clan over half a century previously.

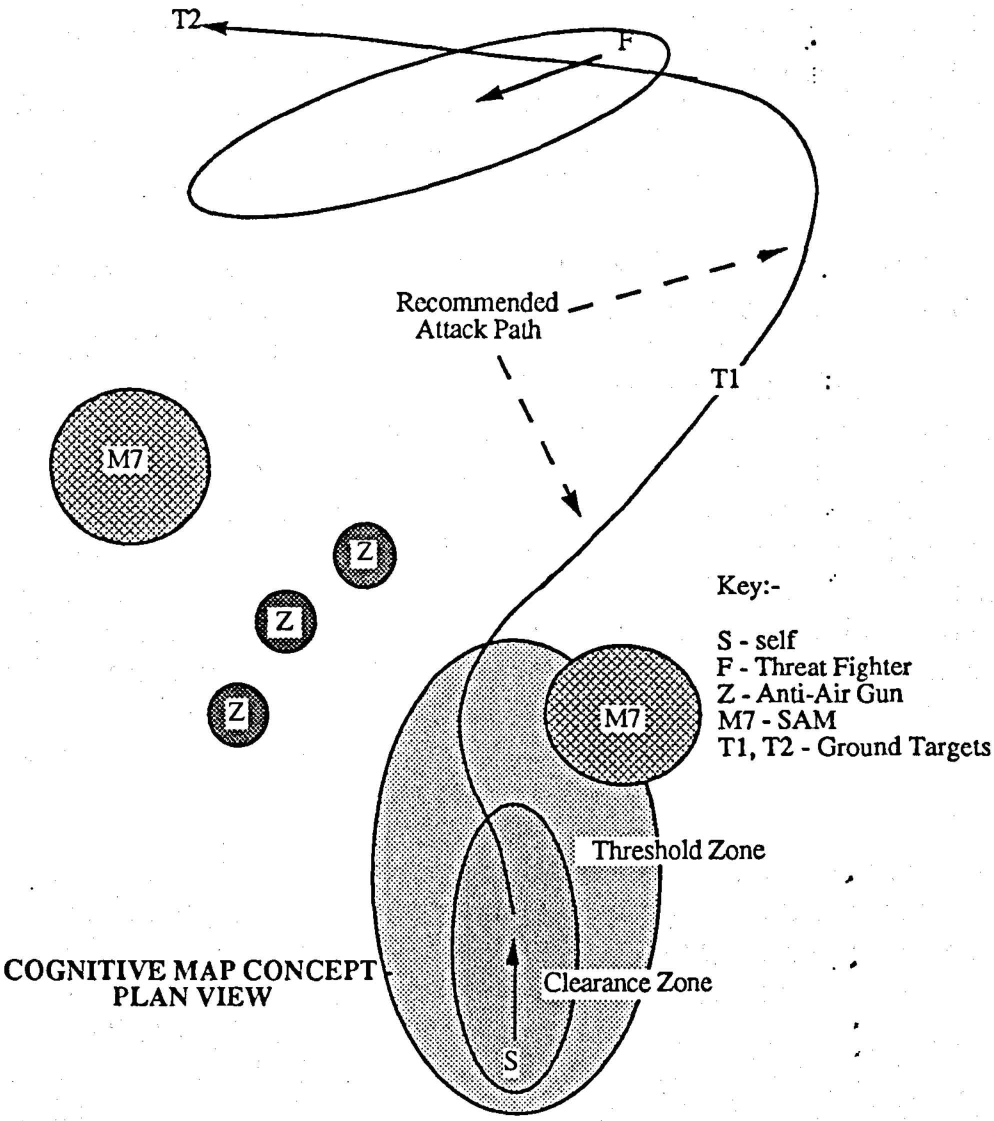
An understanding of territorial imperatives and tribal culture is not academic; it is an essential element of IDA design concept formulation. An IDA system which cuts arbitrarily across tribal or clan boundaries *will not work,* because the highly robust human systems will revert to tribal boundaries. The military are equipped with many such systems; security inhibits description.

The last paragraph presented the negative viewpoint; the positive approach to design concept formulation is to recognize the strong human system drivers and to design systems compatible with these influences. In fact, this approach seems in practice to offer a double advantage. Systems which are designed with subsystems which relate to those in the human body seem to be more readily understood and operated by new users. The following two systems were designed by me according to these rules. *HITAS* was designed to be human-like; *CLAN* was based on territorial imperatives and clan loyalties

The Heuristic, Intelligent, Threat-Assessment System (HITAS)

In 1981 British Aerospace were engaged in a design for a new fighter, the Agile Combat Aircraft. That project is now defunct but, in the way of these things, the design work translated into the successor, the European Fighter Aircraft (EFA).

BAe designers felt that the crew workload dictated a two-man crew, but the aircraft was intended to be small, agile and inexpensive, all of which favoured a one-man aircraft. I was asked to conceive an "automatic observer", an electronic second crew-member with the particular role of detecting enemy ground threats and responding to them



The resulting design study is in the enclosed paper *Automatic Airborne Threat Assessment*, but the design became better known as HITAS. The study was very well received, both by British Aerospace and by the Royal Aircraft Establishment , Farnborough, and is still referred to within the company of origin, EASAMS Ltd, as the first study which produced a complete avionics architecture from first principles. The design has the following subsystems:-

Non-imaging sensors
Thermal Imager
Radar
Multi-sensor correlation
Hazard Sensor
Ranking
Response

The non-imaging sensors gave very low directivity but wide scan, and were used as peripheral vision sensors to direct the infra-red and radar which equated to foveal vision in their resolution. Thermal imager and radar gave scene interpretation on their own, but their outputs were correlated to give a combined scene which was set against stored multisensor threat models representing the known threats.

Recognized threats were ranked using a 3-D cognitive map approach, with the aircraft at one focus of an ovoid, or solid ellipse - see the figure - the major axis of which extended in front of the aircraft by an amount related to aircraft speed; the faster the aircraft, the longer the ovoid. Around recognized threats were generated, in the systems processors, spheres of lethality corresponding to the ranges of the recognized threat weapons. The volume of the solid intersection between these threat spheres and the zone of safety set by the ovoid represented the degree of risk facing the aircraft. Threats were ranked according to degree of risk and imminence.The degrees of risk were used via an autopilot to steer the aircraft by the minimum risk path via the threats in precisely the same way that pedestrians avoid each other when hurrying along a busy street - the model I used, based on research work into human cognition being undertaken at Newcastle University at the time.

HITAS had to "learn" about threats. The basic idea was to present the sensors first with a series of physical threat models, suitably scaled, so that the processors could learn what a threat looked like. Next, the system was to be taken airborne and the same process repeated against real targets. It was intended also that HITAS should be able to gather intelligence by recording real threats that did not fit its models and that a degree of latitude should be provided in the design such that HITAS could modify the polynomials which formed the basis of its threat models. Hence, the system was heuristic, and intelligent, using that last term in a limited sense that it was to be able to learn from its environment and to modify its behaviour sensibly according to that learning.

A list of the analogous, human-like features embedded in the HITAS design concept include:-

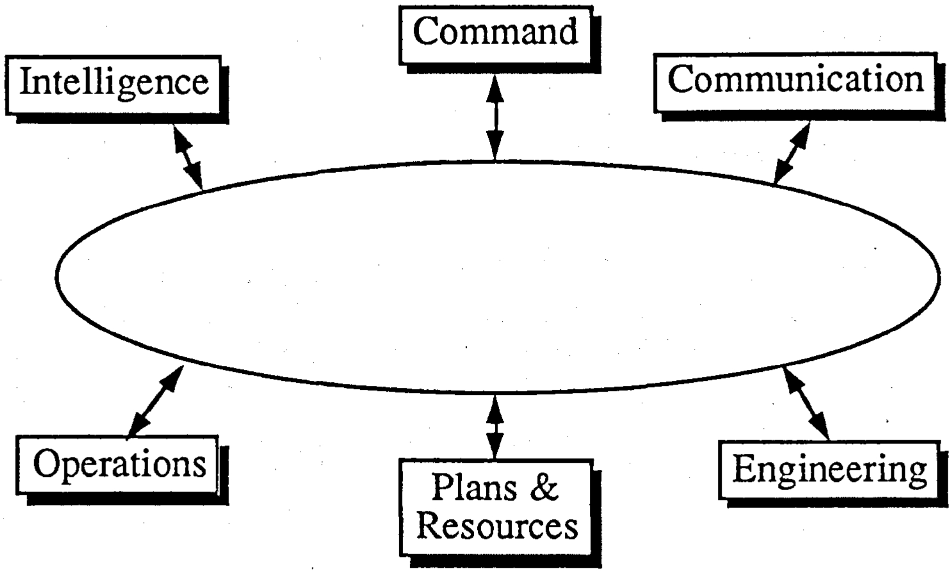
Peripheral Flicker Vision
Foveal Vision
Heterogeneous Remote Sensors
Sensor Correlation
Image Learning
Image Recognition
Cognitive Mapping
Route Finding
Reflex
Learning
Adaptive Behaviour
Judgement / Prioritization

The time was 1981. AI had yet to become popular. Expert systems did not exist, and my solutions were Bayesian. My calculations showed that the amount of processing needed was such that, while a slow-flying helicopter might be equipped, it was impracticable to fit a high-speed fighter with the then technology. However, in the design, I developed a basic three-bus avionics architecture which, I am lead to believe (I can say no more, as a subcontractor at the time) became the basis for the bus architecture for EFA. It is noteworthy that maintaining the analogy with the human, both anatomically and socially, had enabled a complete avionics design concept to be developed

CLAN

CLAN, Command of Land, Air and Naval (Forces) is not described in an enclosed paper; the company concerned did not encourage professional papers for reasons of commercial security. The design concept was avante garde and still is some six years later; it had a profound effect on the subsequent approach to system design within the company, and on competing companies who saw it at exhibitions

Instead of the usual central processor supporting sets of terminals arranged in cells, Clan used a UNIX-based approach to connect a number of workstations via a Cambridge Ring. Each workstation corresponded to an operational group, or clan. There were stations for intelligence, communications, operations, engineering, plans & resources and, of course, the commander. Each workstation could support a set of terminals working through it, so that a section of staff could simultaneously contribute to the task in hand; these supporters could, but normally would not, communicate over the Cambridge Ring - each section was clan-based, in the human societal sense.



Clan - Function-Mapped Architecture

A strict code of data ownership was conceived (now popular, but novel at that time, 1983). So, the intelligence desk officer was responsible for all intelligence data in the HQ; others might read it, only he could authorize its update.This approach maintained territorial imperatives

Each HQ communicated with other HQs via bridges between the rings. Intelligence communicated with intelligence, operations with operations, engineering with engineering, and so on, so maintaining clan protocols

Subordinate formations also stuck to the clan code. A logistic depot, for example, communicated stock levels exclusively to the Plans & Resources Officer, who then authorized the data for dissemination. This last concept invoked a level of sophistication in the protocol; the P&L Officer could disagree, with the depot, or wish to propagate different data. The concept of Data and Information Ownership arose, with the subordinate formation, who could actually count items physically being declared the data owners, while the HQ officers, who aggregated and occasionally "massaged" that data became information owners. Thus pecking order dictates were satisfied.

The effects of CLAN on customers was quite staggering; generally military operators themselves, but often foreign and lacking in English, they understood CLAN immediately. Each workstation was "stand alone" and they could perceive a clear, physical, territorial boundary. The separation between the purposes and activities of different workstations was also instantly understood; it mapped directly on to their internal military organizations. The physical design of each workstation employed a low desk, with inset graphics display screens and controls, so operators could literally sit in a circle, as in the diagram, facing inwards and backed by their supporting sections which could be visible or concealed, but which could in either case communicate verbally and through the various section terminals. The commanders desk was similar to the others except that it was a little higher and had a larger graphic screen, consistent with his position of dominance.

Regrettably, like many firsts-in-the-field, CLAN suffered development problems; it employed a software communications package called Newcastle Connection, named for the university of origin, which lacked security features and was less than robust. Nonetheless, CLAN's design concept influenced many other companies and its goals are still as valid today as they were. ANSA, mentioned earlier, seeks to provide a similar nodeless, integrated networked system.

4.3. Models of Human Organization in Command & Control

The Air Command and Control System (ACCS)

At the time of writing, the NATO Air Command and Control System is rising, somewhat like Lazarus, from a moribund five-year slumber. I first wrote about this mammoth international venture in the enclosed paper *Air Command and Control*, a paper presented at Hendon Air Museum in 1981 as part of the Marconi 81 Symposium. I spent four years subsequently as the UK Technical Director for Airspace Management Systems, one of two giant consortia bidding for the ACCS business in Brussels from 1982 until today. During that process the following concept, *MOSAIC*, sprang to life, unbidden but irresistable in its simplicity and logic.

MOSAIC

MOSAIC is a major work, addressing at the most fundamental level the manner in which NATO deploys its air power in Europe. Presented first in Brussels at the 1984 European Symposium of AFCEA, the Armed Forces Communications and Electronics Association, it was presented again at the IEE's First International Conference on Advances in Command, Control and Communications Systems in 1985, and was subsequently published, along with other selected papers in book form.

MOSAIC appears to preach heresy, at least when regarded from a conventional airman's viewpoint. The present approach, which has changed little over the years, is to defend a land area, using air defence aircraft, surface-to-air missiles, etc, and to use the secure area as a base from which to launch bombing attacks on enemy ground forces and into enemy territory for interdiction. The concept, as developed in western Europe, results in a layered defensive screen facing east, behind which the attack forces shelter. There are limitations with this approach; it requires the enemy to attack from the east. The advent of surface-to-surface missiles, ground and submarine launched, together with aircraft carriers in the various seas surrounding northern Europe, render the assumption suspect.

The essence of *MOSAIC - Moveable, Semi-Autonomous Integrated Cells* - is Survivability, and the approach is in the title. Instead of dividing air power conventionally into Offence and Defence, with the Offence traditionally hiding behind a defensive screen (provided by the Defence) and sallying forth on missions, why not form cells comprising Offence, Defence and Transport all in one such that each cell could operate autonomously, could fight, retreat, or jump over and was, indeed an independent unit? Further, since autonomy is not the normal practice, why not band these autonomous cells together into groups, giving them superior command and control to coordinate their activities? They would thus be autonomous only in extremis but, like the Japanese soldier, would be highly survivable.

It might seem that *MOSAIC* operates across the tribal boundaries and to an extent it does. However it is in essence little different from an aircraft carrier which has transport, offensive air and defensive air. Imagine a series of land-based aircraft carriers, mobile, potentially autonomous, but under the control of a command team which groups them according to threat and environment so that coordinated air attacks and defence can be undertaken. Essential clan loyalties remain intact.

The value of *MOSAIC* is not simply that it offers an alternative approach; *MOSAIC* enables a fresh look to be taken at the existing approach in which a vast amount of effort and money has been invested, and to see it differently. In other words, it presents a different model, not unlike Checkland's concept, from which recommendations about the real-world system might flow.

Today, *MOSAIC* stands as the only alternative to conventional air power deployment in Europe. *MOSAIC* has created, and still creates, a considerable amount of interest not only because of the basic idea, but also because of the spin-offs. These include the concept of a universal language, TADIL-ACCS (Tactical Data Interchange Language for ACCS) and the concept of an airborne action group, analogous to surface action groups, comprised of a variety of aircraft types, including C2 aircraft.

The General Theory of Command & Control

A General Theory of Command and Control is a bold attempt to provide a structured approach to C^2 design by developing an underpinning theory. The paper arose in response to a gauntlet thrown down by Professor C J Harris,the chairman of the IEE's organizing committee for the Second International Conference on Command, Control, Communications and Management Information Systems, in April 1987. He deplored the lack of any coherent C^2 theory. It took some two years to conceive one.

Why is *A General Theory of Command and Control* bold? Some experts believe that a theory of command and control is not a practicable proposition. Levis and Athans (1988) take the view that a theory is impracticable because of the high degree of complexity and because of the unpredictability of the humans who are embedded as decision makers in both of the opposing forces' command and control organizations.

A General Theory of Command and Control postulates an ideal world of human activities, and seeks fundamental axioms in that world. This approach has unimpeachable credentials; it is the method of Galileo, Newton and Faraday. Newton's three laws of motion apply in a perfect world, for example, where mass is equated to inertia - a reasonable, but unprovable supposition. *A General Theory of Command and Control* produces ten so-called Laws or Postulates of Command and Control which apply to a similarly-ideal world and so set standards for the real world, in much the same way as Shannon's Laws of Information and Communication present useful, but unattainable, benchmarks. *A General Theory* certainly brings to the fore relationships that C^2 system designers presently leave unaccounted and it offers some guidance in such difficult areas as the prediction of database sizes and communication capacities - and yet technology is deliberately not discussed.

The method adopted in *A General Theory of Command and Control* is to model the subject using a wide variety of modelling methods, each effectively forcing a different viewpoint. Surprisingly, neither classic control theory not information theory gave much insight, while Systems Dynamics and Simple Queuing Theory proved much more helpful. Most insights came simply from examining pyramidal command structures, their clusters and linkages. Some of the laws suggest that conventional approaches to designing IDA and C^2 systems are extremely profligate with technological resources.

A General Theory of
Command & Control

Vertical Data Compression
Diminishing Lateral Cooperation
Equipartition
Infrastructure Expansion
Action/Order-Rate Equality
Action/Reporting-Rate Equality
Infrastructure Expansion
Decision Rate Invariance
Decision Scale Invariance
Survivability

Command and Control "Laws"

The ten laws presented in *A General Theory of Command and Control* are principally about human inter-relationships. The Law of Diminishing Lateral Cooperation states that peers in a pyramid organization will have less success in cooperating the further they are separated horizontally within the structure. This phenomenon, well known to us all, arises because the greater their lateral separation, the larger the number of superiors through which sanction for cooperation must pass. With each additional superior, the probability of approval diminishes (Occam's Razor or contingent probabilities) and hence the law.

The Law of Vertical Data Compression operates up and down a hierarchy, as the name suggests, rather than laterally. The Law observes that humans have approximate parity of intellectual capacity. A superior with, say, five subordinates (span of control equals five) can absorb the same amount of data as each of his subordinates, individually. If he is to supervise their activities then he can at most absorb only one fifth of each person's output, on average. Hence the Law, which states that the mean rate of vertical data compression equals the span of control (in a pure pyramid organization). The remaining laws are along similar lines and concern themselves with decision dynamics, human system inertia, and infrastructure in terms of the numbers of relationships generated by different organizations.

A General Theory of Command and Control has raised considerable interest amongst military specialists, varying from the view that the laws are little more than statements of the blindingly obvious, to requests for experimental validation. If the laws are obvious, then they probably do qualify as axioms which have, surprisingly, never been stated. As for experimental validation, that will take rather longer - even Newton's First Law lacks that seal of respectability today.

Summary

The chapter concerns itself with humans, their nature and characteristics both as individuals and socially. The value of understanding both humans and groups of humans, as organisms is presented in a variety of papers, as follows:-

- HITAS, an intelligent, learning system based on the abilities of people to weave their way along a crowded pavement without colliding with others similarly occupied.

- CLAN, an IDA system which took advantage of clan and tribal loyalties in its design rather than, as so often happens, seeking to override such boundaries in the interests of technological optimization.

- MOSAIC, the only alternative to the current approach to the deployment of air power in European NATO, emphasising survivability of human and function rather than of physical assets.

- A General Theory of Command and Control, which develops a set of axioms for an idealised human hierarchy.

PART B

CONCEIVING IDA SYSTEMS

5. CREATING SYSTEMS

5.1. Incentives

In one small organization where I was the line-manager for all engineering activities, the organization was proposing to supply quite complex, custom intelligence systems for overseas customers at the rate of one every two weeks.The engineers working with me were bright, intelligent, but inexperienced and I found myself executing most of the design tasks in an effort to meet the demand. It is a familiar trap to most managers; in the time it would take to explain the job, you could do the job yourself, so you do. Result? Overworked manager, under-worked and demotivated staff.

It became evident that I was repeating the same groups of tasks in the same order with each new project. Clearly there was a well-founded pattern to the process, which involved providing customers with a design solution to a fixed price and timescale. To establish both of these it was necessary to plan the whole of the requirements analysis, system design, system development, test, integration, shipping, setting up, commissioning, handing over, training of user staff, after-sales service, and so on. Although each project was different, there was a substantial core of commonality between them.

I developed a standard framework of tasks which was given to the engineers as the basis for each new project. The framework evolved with experience, becoming more abstract in some areas, more detailed in others. The resulting framework is enclosed as a set of figures under the heading *Initial IDA Engineering Framework.*

The engineering framework proved highly successful within its limitations. It was applicable to all the projects the organization undertook, even for quite different customers; while it never fitted any project precisely, it was a sound basis on to which to graft project-specific needs and similarly from which to eliminate inappropriate tasks. The framework was, however, too specific in two respects. All of the organization's design solutions presupposed a particular range of technology, CLAN (q.v.), and these specifics had become embedded in sections of the framework. Further, Middle East customers were quite uncommunicative about their needs, it being considered a loss of face for them to admit lack of knowledge on the one hand, and for security reasons on the other. As a result, the framework took little account of human issues and concentrated on "turning the handle" to produce technological solutions as quickly and effectively as was practicable.

5.2. Managing Systems Creation

The enclosed paper, *Managing System Creation*, first presented at the IEE in 1985, grew out of the initial framework. This paper, which was awarded the IEE Management and Design Division's 1986 Premium, covered a much wider field, a general approach to design, development and implementation of IDA systems - in fact, the term IDA System was coined in this paper. For one company, the paper has become a "bible", a handbook given to each engineer in the company to ensure a common, coherent approach across the board for the design and development of systems.

Managing Systems Creation takes a high-level view of systems, which are presented as having seven ages, by analogy with the seven ages of man.:-

Conception
Design
Development
Implementation
Transition
Utility
Senility
Replacement

The Seven Ages of System

(Replacement might appear to make the number of ages sum to eight; however, Replacement is generally an event and, like death, cannot be considered as part of the living process). Managing Systems Creation also presents the view that a customer purchasing an IDA system is actually paying for five systems, and will need to buy, own and maintain three of them. The five systems are as follows:-

The IDA system which the customer wants (eg an air transport system)
An IDA crew / operator training system (eg flight simulators for the aircrew)
An IDA maintenance and servicing system (eg aircraft maintenance systems)
A test and integration system for use in the factory during development
An in-company support system to maintain the developing equipments

The first three systems in the list are needed by the customer; the last two are paid for as part of the contract; all are IDA systems in their own right. *Managing Systems Creation* presents a variety of methodologies for analysing systems, for trading between options, for connecting systems for integrating systems and for organizing systems to create systems.

Managing Systems Creation seeks also to present the systems engineer's view of creating an IDA system which will last and be valuable to its users and owners for its lifetime. In some respects, the paper presents the systems engineer's ethic; in others, it presents a simple blueprint for how to organize a project-management-based organization; and in still others, it presented a set of graphical methodologies for understanding design requirements.

Managing Systems Creation, as the title implied, is concerned with the creation process. It is intentionally more abstract than the intitial framework and, as such, more widely applicable. Some attention had been given to front-end issues, such as "the wider system" and system boundaries, without being too specific about the meaning of either term and certainly without showing ways to address these issues. The approach adopted in *Managing Systems Creation*, for good and sound engineering reasons, was based on successive, formal specifications to ensure a firm basis both for understanding between customer and supplier and and for progressive, well-regulated development and implementation activities.

5.3. SEAMS and Systems Factories

Developing a creative framework on paper is one thing; encouraging an organization to use it is another. Organizations are, after all, human activity systems and they are robust and resistant to change. They also lack corporate memory, that is the ability to remember approaches that worked and ,particularly, that did not work, and to learn from the lessons by transferring the memory between projects.

An organization, particularly an engineering organization, needs a corporate memory. In the past, such memory has resided in long-serving employees who have reached a position of respect and who remember successful ways to achieve objectives or pitfalls to avoid. Nothing is more galling than making the same expensive mistake twice simply because nobody remembered the lessons from the previous occasion. Attempts to store such learning in company standards usually fail because of the robustness of the human engineer-system which will often revert to redesigning the wheel from first principles, and which is generally averse to being governed by thick and dusty tomes full of procedures.

With projects becoming more complex and diverse, and with labour mobility becoming the norm rather than the exception, depending on written procedures and on the memory of individuals is inadequate. There is a need to build procedure, process and experience into a live, constantly-evolving framework which can incorporate the avoidance of pitfalls. The enclosed paper "*Getting SEAMS Straight*" introduces the concept of a computer-based, company-wide system for such an enterprise. SEAMS is an acronym from Systems Engineering, Analysis and Management Support (Environment)

SEAMS introduces a variety of elements, including:-

SEAMS as a meta-system
A core framework of engineering tasks
The system engineers toolset
SEAMS as a corporate memory
Standards for Procedures
Standards for Specifications
Standards for Project Control
Standards for Quality Assurance
Standards for Reviewing the Standards
SEAMS as a company-wide CAD/CAE "backplane" into which project-specific tools and techniques could be "plugged" according to need
A systems engineering Code of Practice

(The Code of Practice is being actively pursued by The IEE's M5 Executive Committee on Systems Engineering)

SEAMS is only one example of the general move toward what are sometimes called system factories. The Alvey Software Engineering Programmes had as its keystone the Information Systems Factory (ISF) which sought to provide a continuous on-line environment of computer-aided tools to permit the analysis of requirement, high-level-design, module design, code, link and test, of software, together with documentation production and configuration management.

Latest in the line of systems factories is the ESPRIT project, ATMOSPHERE, which will seek to provide a similar environment for the creation of systems as ISF sought to provide for software. The success of these very large enterprises is by no means assured - if it were possible to develop an effective metasystem for the creation of new systems, it is by no means evident that it would be effective and affordable; the degree of investment which would have to be made by an organization would require that the organization have a systems production throughput on a massive scale. Another concern revolves around the definition of "system" as it applies to ATMOSPHERE - is a system defined around a computer system, or could it encompass wider systems, such as aircraft or ships, or could the definition run to companies and organizations?.

Other Engineering Frameworks are also in evidence, although most are rudimentary. The most impressive is that of M'Pherson (1980) in which he describes graphically the process of developing a system design, par excellence

SEAMS stands at present as the only proposal before UK industry for an affordable solution to the need for better ways of creating systems. It is not a factory, for that is presently both too difficult and too expensive. It does incorporate a framework of tasks and it does include features which amount to a corporate memory. It contains no tools as such.but rather the ability to interface with tools. The present status of SEAMS is as a proposal within industry, being actively pursued. In course of that pursuit, I produced a network of the related set of tasks that SEAMS would support in the process of conceiving any IDA system. This network, or framework for conceiving design concepts, is unlike any of its predecessors and is the subject of following chapters.

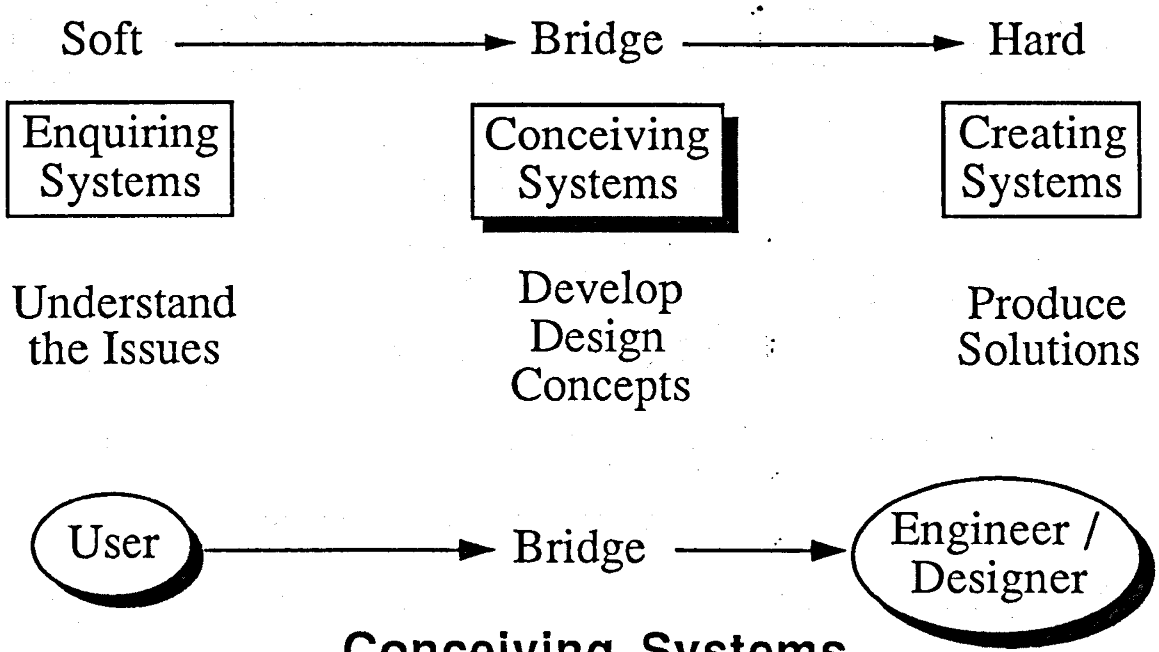
Summary

The chapter presents the topic of Creating Systems, an important topic in its own right, but one which is presented here particularly to show the development of engineering frameworks into which the output from any Conceiving System must feed.

6. A CONCEIVING SYSTEM

6.1. The Purpose and Role of a Conceiving System

Previous chapters have presented Enquiring Systems, to enquire into problem situations, and Creating Systems, to develop systems, ab initio, to meet a requirement. Between these two metasystems is a gap; the purpose of a Conceiving System is to fill that gap.

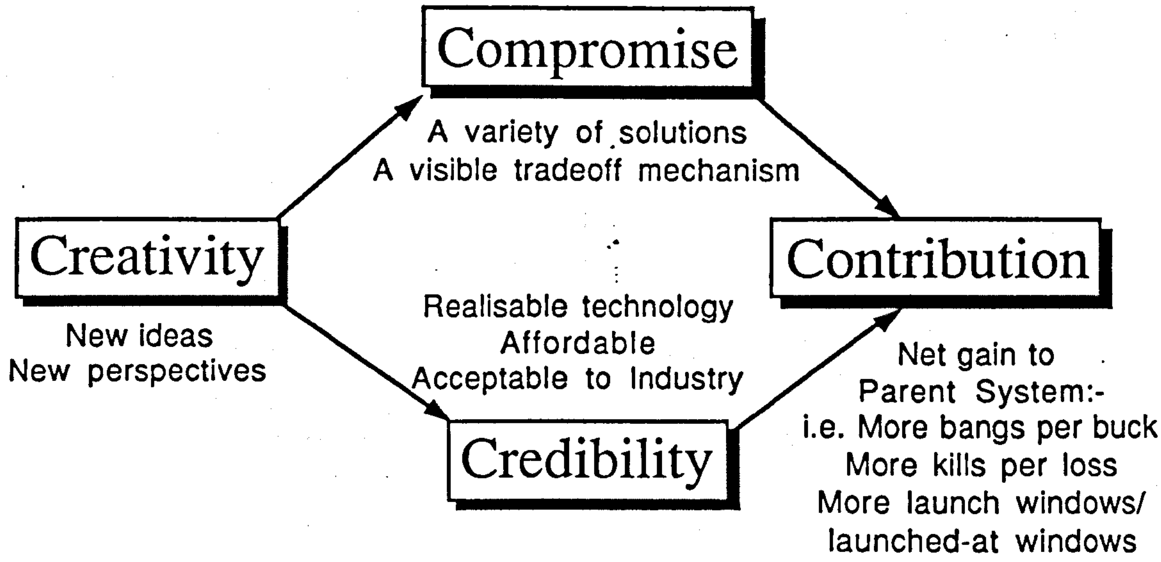


Conceiving Systems

A Conceiving System is intended to bridge the gap from soft issues where there is no established requirement, simply a problem situation suggesting a need, to the development of a substantial design concept for a solution to (part of) the problem. A Conceiving System should also, by virtue of its concern with operational domain issues, provide a bridge between user / operators, those who understand the domain, and designer / engineers, those who understand the technology. A Conceiving System is a meta-system, too. Its objective is to provide a coherent basis for structured innovation of, particularly but not exclusively, IDA systems in their broader definition.

Coherency requires that a design concept be realized comprehensively and traceably and that each step of the path be justifiable. In the real world, as experiences with Creating Systems testified, work is more often carried out by the inexperienced, rather than the experienced designer, particularly where innovation is the order. Coherency therefore must be judged relative to the inexperienced.

Innovation is a much bandied term. Innovation is concerned with more than creativity. Innovation involves the realization of creative ideas, and that is a much harder process than simply creating ideas.



Innovation - the 4 Cs

As the figure shows, innovation may be thought of under the headings of:-

Creativity, to generate the ideas and concepts

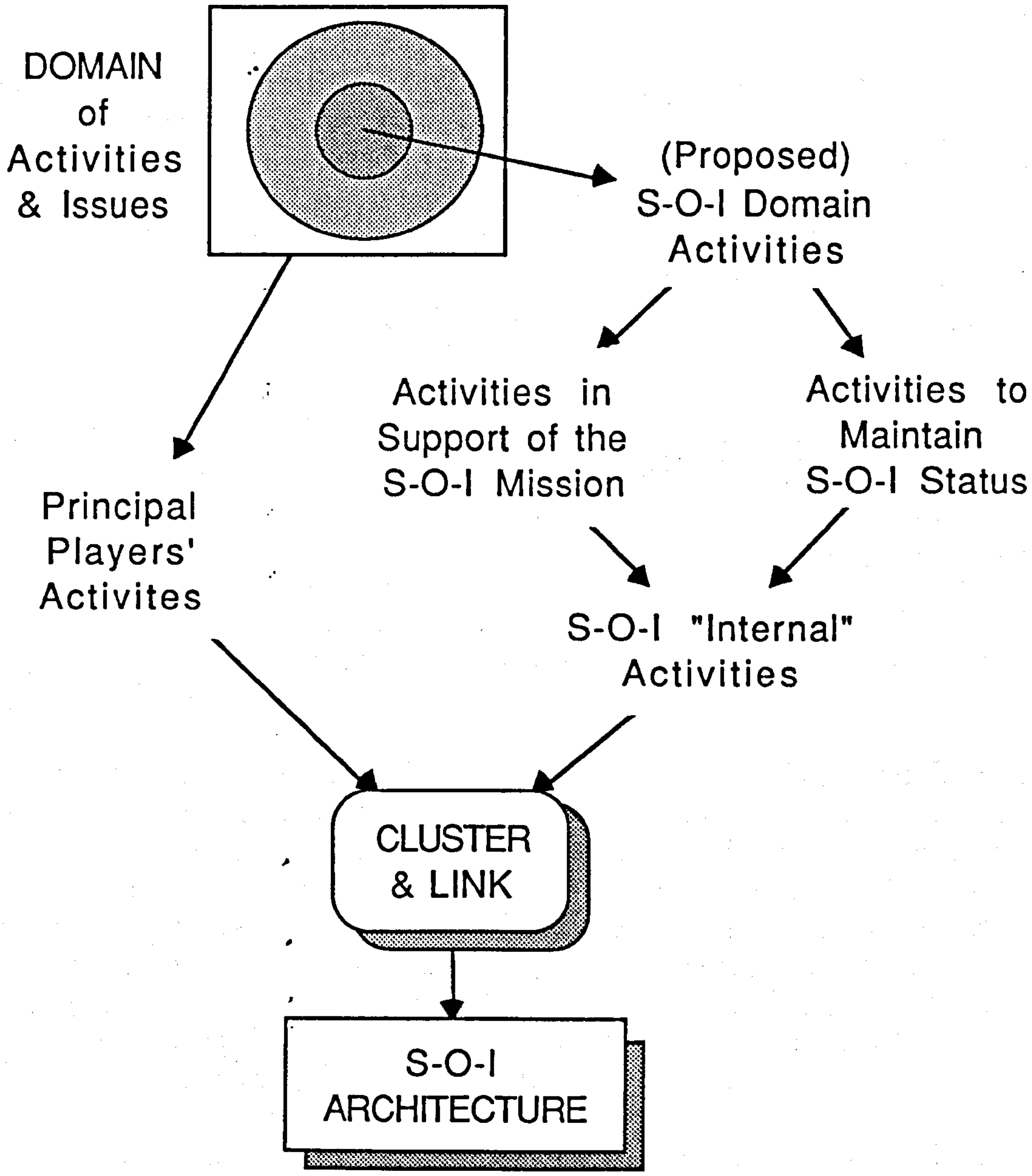
Credibility, to ensure that the creative ideas ar more than flights of fancy

Compromise, to recognize that any real-world solution has to be adapted to its environment

Contribution, to ensure that the creative concept is really useful

6.2. Issue-Based Approach

Problem situations exist in domains of activity. The problem situation concerned with command and control of a new frigate exists in the management of naval warfare domain. It is not possible to understand the problem issues without understanding the domain issues.



THE DEVELOPMENT OF ARCHITECTURE

The figure shows a domain at the top within which are many activities and issues. At the centre of the domain is shown a set of activities proposed to resolve one or more issues and to be undertaken by a proposed System of Interest (SOI) which may be virtually undefined and unbounded. Surrounding the central core is a set of activities related to those proposed for the SOI within the domain (outer ring) and beyond that are the remaining domain activities which have no meaningful relationship with those of the SOI

The Proposed SOI activities comprise two major groupings: those in support of the primary mission, to resolve the domain issue(s), and those designed to maintain the status of the SOI itself, since it will need to perform and survive in a hostile environment and it may absorb energy and resources from within the domain. Together, these activities may be regrouped and considered as the SOI Internal Activities; typical internal activity groupings might be entitled "Conduct Operations", "Collect Intelligence / Information", "Assess Situation", and so on

The Domain activities related to those of the SOI are undertaken by associated systems, referred to in the figure above as Principal Players. These activities, since they relate to SOI Internal Activities, influence the organization of those Internal Activities and a process of activity clustering and linking recognizes the influences and develops the basis of structure, or architecture. Note that SOI architecture is, in a sense, a reflection in miniature of the of domain of activities which generated it, consistent with ideas of Russian Dolls

Within the domain and influencing it are political, economic, technological and environmental factors, at the very minimum. In real instances, ethnic and religious factors can predominate, and clan, tribal and territorial imperatives are always present, often in the guise of "doctrine" or "practice". Domains of human activity also exist in a variety of states. In conflict, such states would include, peace, tension, transition to war, and many others. In less primitive domains, terms such as "standby", "shift", "silly season", "working-to-rule" and"normal service" suggest finite states. The imperatives and influences differ markedly between finite states in the same domain.

Against the backcloth set by these many factors, there may be a problem situation that requires examination. It is in the ethic of a conceiving system as here defined that it anticipates the possibility of solution; in this respect, it differs from enquiring systems.

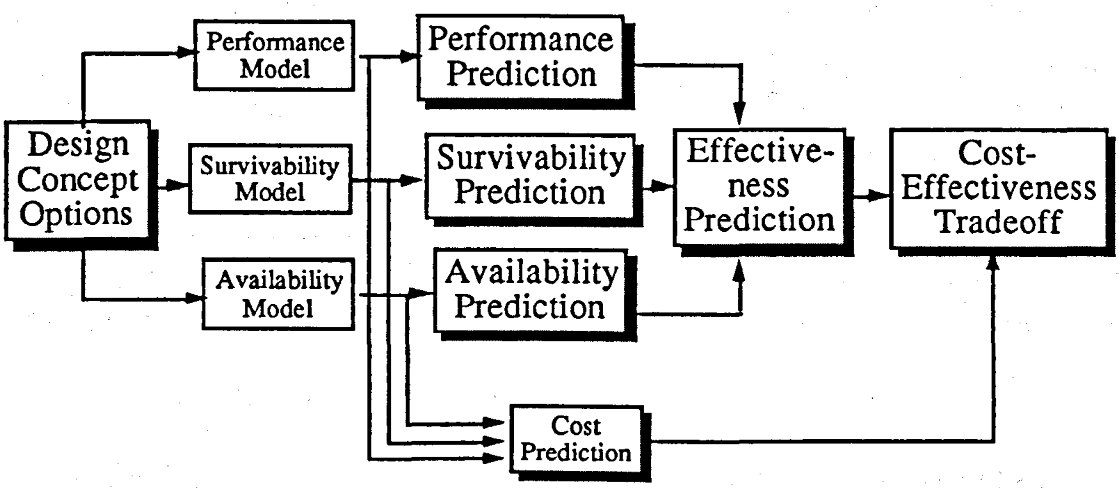
The domain will exhibit a variety of activities, undertaken by the domain occupants. Issues arise because these activities: are inadequate, superfluous, inappropriate, untimely, aggressive, expensive, inconsiderate, offensive, politically irritating, etc. In exploring these activity issues, within the domain, under wider system influences, strains become apparent. It is possible in these problem situations to postulate a variety of solutions. It is likely that SSM, for example, will provide a cogent medium for unraveling the issues and perhaps even for postulating potential solution systems

Systems have parents and siblings. Any potential system will, by definition, form part of a larger system which is its parent. The goals, objectives and "drivers" of the System of Interest (SOI) must be substantially compatible with those of its parent system. Systems have peer systems, (siblings) and systems serve other peer systems.The characteristics of a SOI must be compatible with the needs of its served system; in this respect, I am in agreement with Checkland(1981) and his Law of Conceptualization

Issues will arise as a result of any new system which is proposed to be introduced into the domain. The analyst has, therefore, to be concerned not only with the issues already extant in the domain, but the changed set of issues that his work might introduce. This requires that he or she produce a putative set of activities which a solution SOI would undertake.

6.3. Cost Effectiveness and Net Contribution

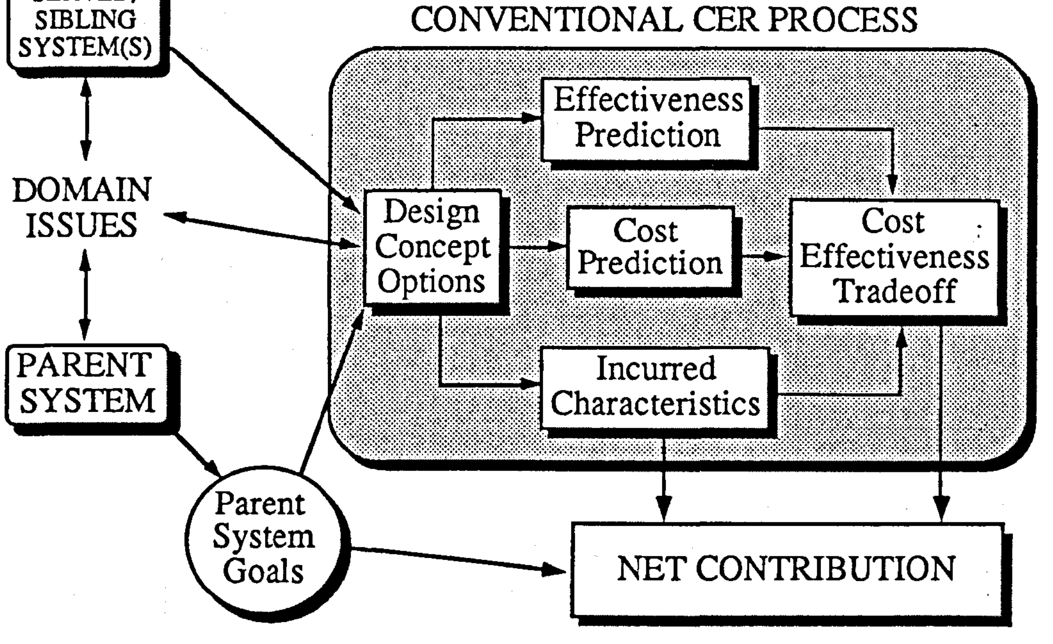
Systems to solve (or at least change) problems are often complex and expensive. It is practice to relate their effectiveness, however anticipated, to their cost, however estimated, and to rank potential SOI competitors according to cost-effectiveness rules.



Cost-Effectiveness

Such rules are valuable in the real world, but they do not tell the whole story. Each potential SOI, in addition to being effective for its particular, declared purpose, is also owned by its parent system. A proper measure of the worth of an SOI is its Net Contribution to that parent. Consider for example a competition between system concepts in which the winner, the most cost-effective solution in relative terms, also was (say) the heaviest in a case where the parent system had a limited payload. Analysis of that SOI would trade relative cost-effectiveness - a gain - against concomitant weight - a loss - and determine Net Contribution. The same SOI, now offered for a parent system with no payload concern, would exhibit a quite different Net Contribution. The concept of Net Contribution is, perhaps, new but important in the conceiving of systems

Net Contribution



The figure shows the cost-effectiveness paradigm from the previous figure in the shaded panel. It also shows that domain issues affect the parent systems and sibling systems in addition to the SOI. Net Contribution is seen, then, as the balance of contribution made by the SOI systems's goals. The figure also shows that the SOI influences the domain issues and indeed an SOI may create as many new issues as it seeks to solve - The US Strategic Defense Initiative is a case in point.

6.4. The Continuum Concept

Some methodologies, and SSM is typical, have very few steps; SSM has seven. The would-be analyst has to progress from step to step, iterating where need be, but essentially he is using his own intellect to find his way in a very loose framework. Advocates of such frameworks would say that it is the very "looseness" that provides the ability to address a wide variety of "soft" problems. I would agree, but it is also true that these frameworks are suitable structures only for practitioners of substantial, and unusual, intellect; what of the rest?

For the less experienced, or for those treading a particularly complex path where the wood may be obscured by the trees, a set of stepping stones is more appropriate. Purists might criticize the concept as reductionist, see Burrell and Morgan (1979) for categorization, but is that reasonable? A man approaches a river, to find no crossing. Using a stick to probe the water and some nearby boulders, he finds a route across the river and establishes a stepping-stone path over the water. Is his method reductionist? In a sense, yes, he has reduced the problem of crossing the river to a series of smaller problems. But in a sense, no, because what he did was exploratory - he had to find his way, there was no sense of decomposition, little of Descartes principles of 1637, so often referred to as the foundation of reductionism:-

> "Accept only that which is clear and distinct as true
>
> Divide each difficulty into as many parts as possible
>
> Start with the simplest elements and move by an orderly procedure to the more complex
>
> Make complete enumerations and reviews to make certain that nothing was omitted"

Whether or no the idea of stepping stones is reductionist, there is clear advantage to the typical analyst in having rather more than less guidance in a methodology, provided always that guidance does not become anti-creative.

Accepting, with that proviso, the idea of stepping stones, considerations of step-size come into focus. In real-world problems, the need for a structured conceiving system arises principally as a result of the need to manage complexity. Individual steps should therefore be such that they uniquely contain easily grasped concepts.on the one hand, and do not require unattainable leaps of the intellect on the other

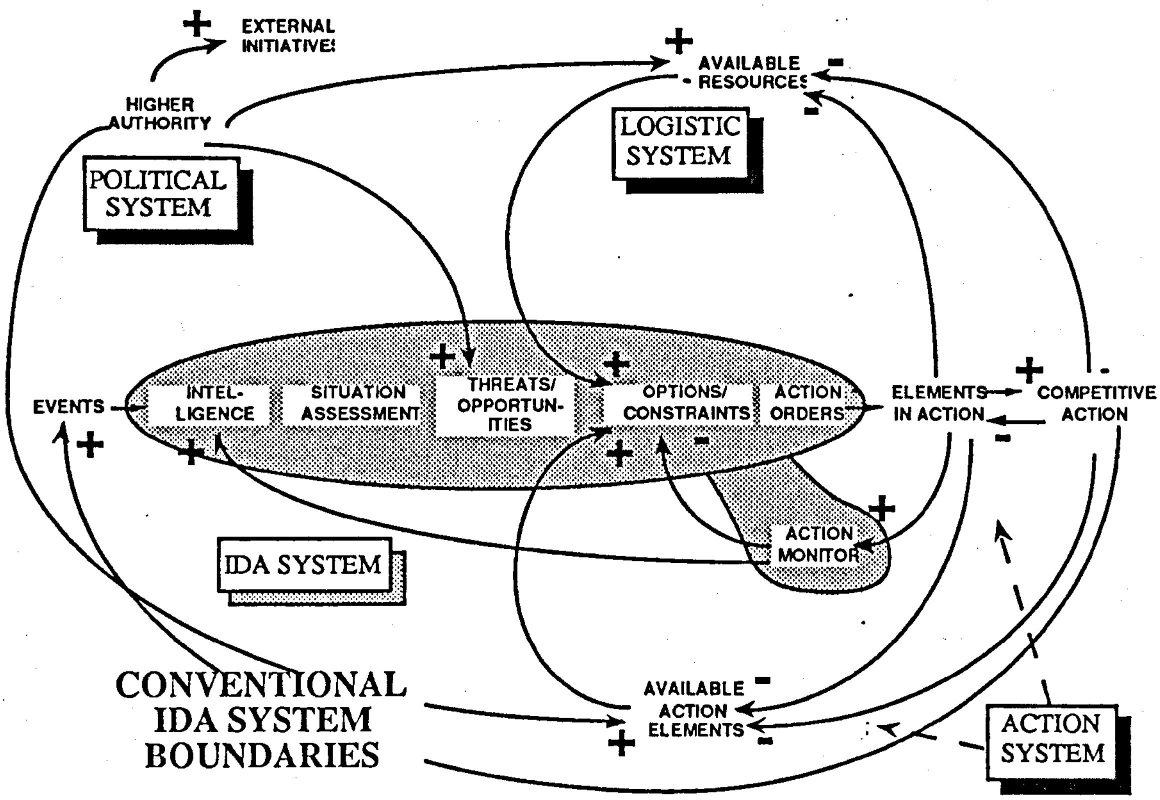
So emerges the idea of a continuum - a set of steps, so arranged as to provide a contiguous route from issues to well-formulated design concepts for solutions.

6.5. Bounding Concepts

Putting sensible boundaries around the SOI is a difficult task. There are few rules, and they are mostly negative; the most important of these is not to identify the SOI boundary with a convenient physical or organizational boundary. That rule is of limited value to the average analyst, serving only to caution, not to assist.

Sensible boundaries concern themselves with interfaces. The ideal boundary would not be crossed in either direction, but would uniquely contain the SOI. The only case of such a boundary is the Universe: again, not too helpful, but at least indicative. A reasoned, and reasonable, SOI boundary is one where the number of relationships with external entities is in some sense minimized and, where relationships do exist, they are well understood.

IDA systems have boundaries that are encouraged by convention. The diagram below shows the convention used by many western military systems. In this approach, the IDA system is separate from the Action System which it controls, remembering of course that the Action System contains one or more miniature IDA systems of its own. Other systems in the figure also contain their own IDA systems. Political and Logistic Systems are in evidence on the figure, and clearly they have an impact on the central IDA system. These influences are, however, determinable to a degree, and the usual practice is to contain within the IDA system, replicas in miniature of these other systems. An IDA system might have a logistics cell, for example, which "looks after" the interface with the complex external logistic system; this is a good example of the proper use of clan loyalties to accommodate tribal objectives, since the logistics cell would be manned by a logistics specialist - and similarly for political advisers, engineering cells, and so on.



An approach to bounding, presented by example later, arises from clustering domain activities and by then identifying the activity clusters with the resulting physical clusters. The process of clustering draws together those activities which have many, or strong, interrelationships. If the clustered activities are retained in physical realization, then relationships will not be impeded. Bounding the SOI then becomes a process of identifying the domain activities and clustering them according to some suitable strategy or rules Some of those rules have already been presented under the discussion about N^2 charts.

Other approaches to bounding are available. The systems dynamics figure above is a useful device for representing connected systems at a sufficiently abstract level for sensible groupings and inter-group influences to be apparent. Similarly, Data State Design offers a usefully abstract view for the same purpose - see the enclosed papers, *A General Theory of Command and Control* and *Managing Systems Creation* for details, the one at macro-level and the other more at the internal boundary level within an IDA system

Since boundaries form around clusters which are the stuff of architecture, the concept of bounding is basic to architecture too. Clusters, and hence, their boundaries, come in varieties. An entity may become a member of a cluster because it has a strong relationship with one of the existing cluster members. Alternatively, it may have many relationships with many of the cluster members. Different clustering rules result in tight, compact clusters, or loose, strung-out clusters. Some strategies leave entities without a cluster to which they should belong; others provide "homes" for all entities but by loose attachments in some instances. Examples of different strategies and their impact on clusters will be presented in Chapter 8.

Visual Q Analysis is developing as a powerful means of exploring clustering and hierarchical concepts; see Macgill (1983). Q Analysis examines clusters and hierarchies topographically and geometrically, enabling the analyst inter alia to, as it were, slice horizontally through a hierarchy and view the cross-section from above.

6.6. Middle-Out Design

Top Down design is a fundamental feature of the systems engineer's philosophy; it implies a high level of abstraction at the start so as to free the mind of prejudice and to encompass all aspects of the subject. Top Down, improperly applied, can be a disaster of overkill and misdirection. It is essential to have some knowledge of "the bottom" if Top Down is to be given sensible direction; understanding the bottom equates to having good domain and subject knowledge.

Most people, if they are honest, do not work Top Down. While it is often sensible to present the results of work top down, analysts faced with a new problem generally seem in practice to start somewhere in the middle of a problem, choosing a topic with which they are reasonably comfortable, so easing their way into the overall problem, as indeed proposed by Descartes. It is the mark of a good analyst that he or she can retract from this early effort and move up to the top level once a greater degree of understanding has been achieved without retaining undue allegiance to their initial work.

The real world in which design solutions have to be conceived operates under budget and time pressures. A methodology that does not make the best use of time will prove impracticable. It is a characteristic of some soft methodologies that they produce at the beginning a vast array of problem situation data much of which is subsequently discarded. SSM's rich picture can fall into this category (when being developed by the inexperienced), although it can also be argued that unless the full spread of situation factors is uncovered, it will be difficult to discern the issues. A compromise is needed which directs attention of analysts towards factors which could be relevant and away from factors that cannot be relevant.

For similar reasons of time and budget, and with the added incentives of traceability and completeness, it is important to employ formality in the concept evolution process. The software industry is presently in the throes of introducing formal methods for mathematically provable software design and test. The move is inevitable where a soft engineering practice meets stark reality in business. The process of conceiving designs is not at a stage of development where such formal methods could be applied, but a simpler form of such formal methods is a minimum feature.

A suitable approach to the kind of relaxed formality appropriate to Conceiving Systems is that of the template. The template is simply an empty set of pigeon holes, suitably annotated with headings indicating the contents with which each hole is to be filled.

A Requirement Template, available from the start of a concept formulation exercise, will enable analysts to concentrate their early work towards the intermediate goal of filling the template pigeonholes. Further templates, available from the start of the exercise, will guide the effort after filling the Requirement Template towards filling the solution templates. Completing templates also provides milestones of achievement and progress, which are necessary even in this essentially creative process.

A Requirement Template, at high level, might appear as follows:-

S-O I PRIME DIRECTIVE
Semantic Analysis
of
Prime Directive

STRATEGY FOR ACHIEVING PRIME DIRECTIVE

Behaviour		Management Set
Aggression Discretion Co-operation	Unifying Concepts Innovative Approaches Strategy Options	Mission Management Vehicle Management Resource Management

Measures of S-O-I Effectiveness	Negative Contribution Factors	Threat to Achieving PD
Performance Availability of Performance Survivability of Performance	Cost Parent System Degradation - Consumption / Dissipation - Opportunity Cost - Adverse Mission Effectiveness	External Internal Environmental Political / Economic / Technology

SOI Requirements Template

(Following paragraphs will explain the meaning and purpose of the various Requirement Template panels.)

The following template is an example of one used later in the process of conceiving a design concept, and it presupposes that several design concepts are vying for success. The Tradeoff Template, one form of the general solution template, is used for recording the characteristics of several design concept options and for trading between their respective merits and demerits:-

		OPTIONS 1	2	3	4
EFFECTIVENESS	PERFORMANCE - Mission Management - Vehicle Management - Resource Management AVAILABILITY - Reliability - Maintainability - Re-Configurability SURVIVABILITY - Avoidance of Detection - Self Defence - Damage Tolerance INTEROPERABILITY USABILITY SECURITY, ETC				
NEGATIVE CONTRIBUTION	COST ADVERSE PARENT MISSION FACTORS FAILURE PATTERNS CONSUMPTION DISSIPATION MASS / WEIGHT VOLUME / SHAPE, ETC				
	NET CONTRIBUTION				

SOI Tradeoff Template

It is immediately apparent that the two templates are related, since some of the titles are shared between them.That commonality follows the thread identified earlier in this chapter concerned with issues, effectiveness and net contribution

6.7. Creative Entropy

Templates are intended to direct the effort, not to inhibit creativity. Stimulating creativity within the template framework is an essential objective of a conceiving system. There is a number of approaches employed in this context; together I have dubbed them as promoting Creative Entropy - the term is intended to evoke an intellectual image rather than a mathematical formula, but it requires explanation

Entropy is a measure of the degree of disorder in a system. Guilbaud (1959) states: "To say that entropy increases spontaneously, in specified circumstances, is simply to say that the physical system in question tends naturally towards states that are more probable, being realizable in a larger number of distinct ways"

Entropy is concerned, then, with variety and possible states. Creative Entropy is the purposeful development of variety and the exploration of possible states, the purpose being to create an information base relevant to some concept-of-interest

It has long been noted that there is qualitative correspondence between entropy and information, supported by a notable mathematical similarity which de Broglie (1951) suggests is the "Most pleasing and most important of the ideas suggested by cybernetics...". Certainly, the generally accepted view of the relationship between the two measures is that an increase in entropy is analogous to a diminution of information, given suitable statistical models for both.

Negentropy, (negative entropy) is a term used by communication and information specialists to describe information content. Checkland (1981) points out that information and entropy are mathematically very similar except for the introduction of a negative sign in the measurement of information content in messages. "Now, since entropy measures degree of disorder, and information may be plausibly regarded as that which reduces uncertainty and hence increases order, it is extremely tempting to equate information with negative entropy".

I am less happy with Checkland's view than with Guillbaud's approach, since the negative sign in negentropy results from taking the logarithm of a probability that a message will occur, by convention less than or equal to unity.

Creative Entropy, as I define it, is the conscious development of concepts, data and ideas, using a variety of techniques to provide frameworks and environments to stimulate and direct this essentially creative process of design concept formulation.

Mind Sets

One way to generate concept entropy is to induce intentional mindsets in the analyst. The template is a high level mindset. Contained within it is a variety of mindsets. These mindsets are self-contained topic areas requiring concentration in which the analyst can roam, generating ideas by brainstorming, ideawriting or cerebral energy. By moving from mindset to mindset, a wealth of creative concepts can be stimulated. It is, of course, important to ensure that the sum of such mindsets is sufficient to cover the necessary range of design issues - that is the role of the template. A particularly useful application of mindset is the necessary and sufficient set.

Necessary and Sufficient Sets

Survivability may be thought of as comprising three sub-headings:-

Avoidance of Detection
Self Defence
Damage Tolerance

Avoidance of Detection concerns itself with camouflage, mimicry, covert communications, Stealth and the like. Self Defence addresses the ability of the system in question to fight off an attack. Damage Tolerance presumes damage and addresses the ability of the system to continue operation..

Together, they form a set which has some of the formalism of a Necessary and Sufficient set. Why necessary and sufficient?. Avoidance of Detection seeks to survive by not being seen. Self Defence presumes Avoidance to have failed. Damage Tolerance assumes Self Defence to have failed. Together, all three cover the range of feasible situations. Provided the individual sets can be filled, then the composite set is both necessary and sufficient.

Necessary and Sufficient Sets can be applied more widely than survivability and are of particular value in Conceiving Systems; they provide an unobtrusive formality to the process of generating and developing design concepts.

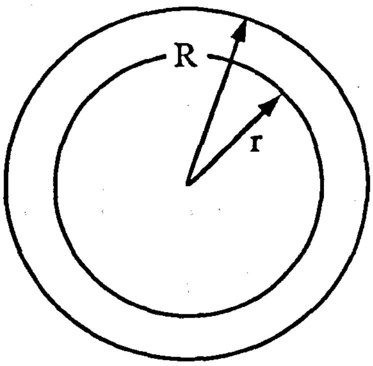
Creative Tension

Creative Tension is realized by identifying objectives in the design, formulating strategies for achieving those objectives and at the same time elaborating the threats to the achievement of those strategies, see Popper (1972). A typical objective might be the improvement of system availability. Creative tension requires attention to the *threat* to achieving that improvement, which might be political, skill shortages, cost, inaccessibility, or a host of other threats. Creative tension prevents narrow focusing on particular aspects of the problem situation which, for improved availability, might have been simply to increase spares in anticipation of failures.

The Spectrum of Options

The degree of Creative Entropy is clearly related to the diversity of optional design concepts that are generated. Choosing both the degree of diversity, and the number, of options is non-trivial.

It is possible to approach an understanding of this issue geometrically, which method I developed when working on the ACCS programme in Brussels to provide a coherent and traceable means of reducing the ideal number of computer simulation runs from 2000 to an affordable number of less than 100. Consider a circle, or disk. It has area and two dimensions. We can show simply that the area is proportional to the square of the radius.



Anulus area as % of whole = 1 - (r / R)^2

Take anulus to occupy, say, 10% of area

Then r = √0.9 * R

r = 0.949 * R

Information Shells

In the figure, only 5% of the radius (1 - 0.949) accounts for 10% of the area. If the axes of the circle represented dimensions of interest, then we could say that 10% of the information content is contained in the outer 5% of the radius. This simple concept applies to any number of dimensions. Consider a problem situation with, say, twenty independent parameters of interest, including speed, capacity, power dissipation, weather and so on. These parameters are independent, or dimensionaly orthogonal. They can thus be represented in N-dimensional space, where N = 10

The mathematics as applied to the circle can be applied equally to the N-dimensional information sphere, but this time taking the Nth, or tenth, root. Keeping with 10% as the figure of choice, we calculate that 10% of the sphere's volume is contained in an outer skin only 1.05% of the radius. I refer to this as the orange peel effect - as the number of orthogonal parameters increases, the burden of the information content in the resulting N-dimensional information sphere migrates rapidly to the surface of that sphere. Moreover, as the number of dimensions increases, the opportunity to pursue design concepts that examine each dimension independently along its length, diminishes due to time and budget constraints. In practice, time allows us to pursue only a few options - the greater the number of dimensions, or degrees of freedom, the fewer the number of options due to their increasing individual complexity.

The choice of options is difficult but, continuing along the multi-dimensional orange analogy, the choice should be made near the orange surface and equidistant around the surface from the ends of axes. For example, consider three dimensions - speed, weather and height - in an N-dimensional information set. Given limited choice, it is better to choose combinations such as: {maximum speed, minimum height, bad weather} than to select {cruising speed, average height, good weather} for the following reasons. The first set has extremes, evidently; the second set contains inter alia, average height. Average height lies between minimum and maximum height

- it is the origin point of the height axis, and so we are exercising the other parameters along the height axis. We will explore the limits of the design envelope less by so doing. Better by far to combine an extreme of height - high or low - with an extreme of speed - high or low, and so on. In practice, many extreme combinations prove uninteresting, as in this example might {low height, low speed, good weather} which would often be mundane. A process of elimination is thus possible in which the less interesting combinations of extremes are progressively eliminated until an acceptable number of options consistent with time and budget is achieved

6.8. The Concept of a Prime Directive

Central to the idea of Conceiving Systems is the Prime Directive (PD). The PD is the highest level of abstract, objective statement of SOI purpose. The expression, Prime Directive, is borrowed from life sciences and is exemplified by homo sapien's prime directive, "propagation of the species"

Mention was made earlier of the difficulty faced by analysts in trying to functionally decompose high level requirements, a truly reductionist process. If functional decomposition starts too low in the hierarchy or if there is no real understanding of "function" then difficulties and shortcomings ensue. The difficulty, previously mentioned, facing radar analysts who start decomposition at the "sensor, communications, processing" point (actually not functional at all, but physical decomposition) is that they can never justify "what *kind* of radar" The homo sapiens PD above is ideal; starting at that point allows the following approach, by comparison with the radar decomposition.

> Humans evolved physically and socially. Social evolution, behavioural evolution during one lifetime, required minimal birth-imprinting. Human children were therefore born helpless and were protected and educated socially within the family circle. Their helplessness necessitated shelter by night, and, homo sapiens operated principally by day. Optical sensors consequently evolved to be optimal in the green part of the visual spectrum, where greatest solar light energy falls by day

Returning to the radar case, were the radar analysts to,work from an equivalent prime directive for their radar,stating what the ultimate objective of their radar was to be, rather than from a set of pseudo-functions, then they would learn what kind of radar sensor was needed rather than simply that one was to exist.

The prime directive is the point to which all design concepts should be traceable. Clearly its formulation is of considerable importance. The PD is not synonymous with SSM's root definition which seeks to describe a system. The PD presumes a general need and describes the need in the highest possible terms. A PD must be abstract. It should, therefore,comprise a phrase containing only one verb, in the infinitive.In that one phrase, only the highest principles should be included.

Perhaps the best way to understand PDs is to observe some. We have seen the archetypal PD for homo sapiens The Royal Air Force have a Prime Directive (my term, not theirs) for the Air Defence of the United Kingdom: "To neutralize enemy air incursions into the UK Air Defence Region". This is an excellent PD. It expresses succinctly and precisely the raison d'etre, the limits of action and the sphere of activity of the Air Defence forces.It does not over-specify; "neutralize" is vague yet entirely sufficient for purpose and there is no hint of solution in the PD's wording. These, then, are the features that characterize a good PD:-

Highest Level of Abstraction Ultimate Purpose Sphere of Endeavour Solution Transparency

Semantic Analysis

Semantic Analysis was learnt in Brussels working with French designers on the ACCS Programme - they appear to use it incessantly. The process is straightforward: each word in a statement (the PD in this instance) is examined and expanded as far as it can be to extract all meaning, stated and implied, that it might contain.

On first association, semantic analysis may seem boring and pedantic. Experience suggests considerable value, however. A group engaged in semantic analysis come to form a comprehensive, consensus view concerning the statement. Where the process is applied to a document, progress - at first apparently slow, rapidly accelerates as prior ideas that have been understood contribute to fuller understanding of later ideas.

I have developed my own style of semantic analysis for the application under consideration; an example follows, using the RAF Air Defence PD as start-point. The PD is , again, "To neutralize enemy air incursions into the UK Air Defence Region"

"To neutralize...	To eliminate the threat from....
....enemy.. hostile	.......those declared by government to be who....
...air incursions..	.... enter by air without permission......
..into the UK ADR". defined sovereign UK	......into the designated airspace legally and internationally promulgated as air space
Implied Means:-	Use of U.K. Air Defence assets

The semantic analysis thus comprises three parts: the PD, the analysis expanded as a continuous sentence, and statement of implied meaning - where such exists. The semantic analysis expands the understanding but without impairing the characteristics presented in the box above.

6.9. Strategy and Threat

The formulation of a sound PD is necessary, essential even, but not sufficient. Consider the case of homo sapiens; he shares his prime directive with every other biological entity on the planet. The PD alone lacks discriminatory power. That power is developed by identifying both a strategy for achieving the PD and a threat to its achievement (creative tension)

Using homo sapiens as an example again, the strategy adopted by our forebears in meeting the prime directive of propagation of the species could be described at great length and it certainly had many variations, but it probably included the following:-

> The formation of family groups and groups of such groups into clans or villages. The nuclear family related to a wider family, often matriarchal owing both to the greater longevity of the hardier females and to their continuous presence within the family unit. Around the family groups, the males hunted in cooperation, while the females and young children gathered food. Families and clans banded together for strength, sheltering in caves and constructions by night and in inclement weather to protect the young, the infirm and the old.

Threats could similarly be elaborated, but might be addressed very briefly as follows:-

> Threats arose from a variety of sources: competing groups of homo sapiens; disease; weather effects on food supplies; carnivores; internal competition amongst aggressive, ambitious younger males, and so.

The homo sapiens sapiens pattern shows how to tackle the IDA problem. Having established a threat to the achievement of the PD in some depth, various strategies can be conceived for an SOI and tried out against it, until a robust, high-level strategy, or set of strategy options is developed. These will give purpose and direction to the design concepts.

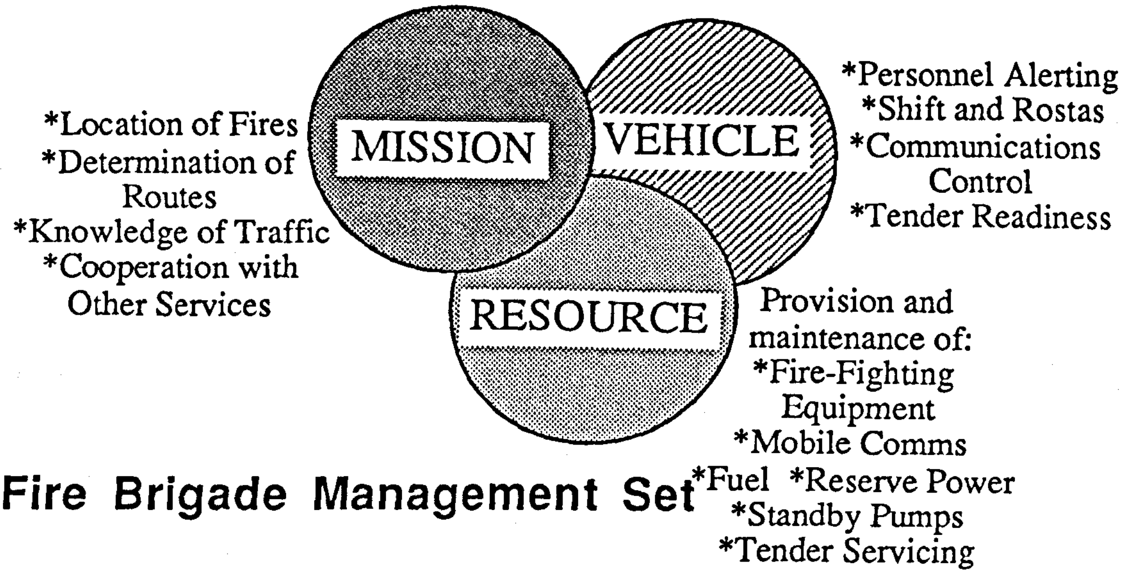
Behaviour

An aspect of IDA system seldom discussed is their 'behaviour". The Soviet Air Defence system that shot down the Korean airliner in recent years certainly exhibited behaviour - fierce territorial imperative. IDA systems generally exhibit character; it is developed within the design concept by the management structure, the speed of response, by the information presented to the operators, and of course by the behaviour of the operators themselves. If the IDA system is viewed as a single entity, humans and machines together, then behaviour is attributable.to the whole. If we look at the two elements separately, then each element enables behaviour on the part of the other.

6.10. Management Sets

Management Sets are necessary and sufficient set which describe the sum of activities undertaken *within* the SOI. Generally, the Prime Directive implies a system mission. The first of the three Management Sets is the sum of the activities needed to effect the mission. The IDA system may be viewed as a vehicle for executing the mission. That vehicle has to be run, organized and protected within and from its immediate environment. The second of the management sets is the sum of all the activities needed to sustain the vehicle as it pursues the mission. Both the vehicle and the pursuit of the mission absorb resources. The third management set is the sum of all the activities necessary to resource both the vehicle and the overall mission.

As an example, consider the following notional example for a fire brigade



Fire Brigade Management Set

Note that mission management is concerned with the overall objective, while vehicle management refers, not to the fire tenders - which are the action elements and which contain their own IDA system, the fire-crews - but to the means of managing local operations within the fire HQ control centre, such as calling-in off-duty staff, alerting senior staff, preparing all the fire-station facilities, etc. Finally, resource management ensures that both the mission management and vehicle management activities are resourced. This means, for example, that resource management will provide mission management with backup tenders as called for to support the mission, while mission management will deploy those backup tenders operationally.

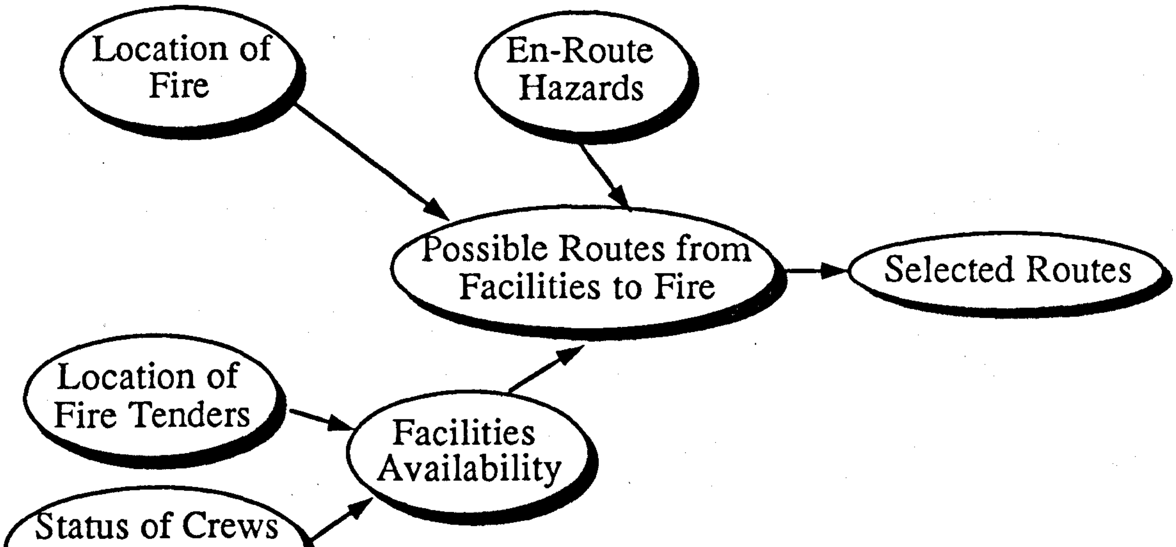
The management set is another Necessary and Sufficient set. Closer examination shows that it enables provision of Performance, Availability of Performance and Survivability of Performance The three elements are mixed, but clearly discernible: Performance is vested in the skill and effectiveness with which control staff respond to fires by deploying and managing fire-fighting facilities and Availability and Survivability are provided by vehicle and resource management, which together ensure men and machines are provided, either from within owned resources or by borrowing from other forces.

The Management Set, as elaborated during the process of design concept formulation, is the precursor to software, since it embodies purpose and meaning into the solution.

6.11. Architecture and Process

The eventual design concept will be characterized by a combination of architecture and process. The Management Set described above leads first to process and then to software. Each of the activities in the fire brigade diagram will result in a set of processes which, taken together, will achieve the activity.

For example, consider the following:-



Determination of Routes - Fire Brigade

The activity,"Determining Routes" is comprised of a number of processes, contained within the arrows of this data state diagram, but none-the-less highly visible.Note the level of abstraction that is being maintained by the choice of representation - the activity is clearly explained as a combination of overt data and presumed process. The diagrammatic language is simple and allows designers, domain experts and engineers to discuss on a common basis, meaningful to all.

The various activities developed in the management set, together with those in the domain, have inter-relationships. The foundations of architecture are established by clustering those activities into organizational sets. This process will be elaborated in following chapters

Summary

The chapter has presented the important, fundamental concepts of Conceiving Systems which act as a bridge between those who simply enquire into problems without a drive to find solutions, and those who drive towards solutions without perhaps enquiring as much as they should.

The components of a Conceiving System are presented as follows:-

- Middle-out Design, using a set of templates such that the exploration of concepts and the generation of data and information is structured and directed, rather than haphazard and capricious

- Creative Entropy, a device for generating a diverse information base from which to develop solution concepts

- Necessary & Sufficient Sets, a means of introducing formality into the information generation, to promote complete coverage of subject matter

- Prime Directive, the concept that any conceived system should be directed towards one overriding objective

- The Triad of Prime Directive, Strategy for its achievement, and Threat to its achievement as a means of generating useful information

- Management Sets, standard Necessary & Sufficient Sets which apply to any organism, and which comprise only three headings:-

    - Mission, or Objectives, Management, concerned with the direct achievement of the Prime Directive

    - Viability, or Vehicle, Management, concerned with maintaining and enhancing the capability of the system to pursue its objectives

    - Resources Management, concerned with supporting the first two above.

7. THE SEVEN STEP CONTINUUM

7.1. A Bridge from Soft to Hard?

The previous chapter described features of a Conceiving System; this chapter presents a conceiving system, and subsequent chapters show it in use. To reiterate, a Conceiving System is intended to fill the gap between:-

> Enquiring Systems, sometimes called Learning Systems, which seek to understand complex problem situations, and
>
> Creating Systems, sometimes called Engineering Frameworks or Design & Development Systems.

A Conceiving System seeks to use methods, from whatever school of endeavour, to understand domain issues, to identify potential solutions and to develop robust design concepts. It is an aim of a Conceiving System that operators and users, familiar with the domain and its issues, could / should form a part of the Conceiving System together with system analysts and designers.

The Seven-Step Continuum introduced and described below is a Conceiving System.

7.2. Addressed Systems

The Seven-Step Continuum (SSC) is designed primarily with IDA systems of the classic variety in mind, i.e. those which relate to, but essentially do not contain, either sensor systems or action systems. However, the framework is sufficiently relaxed to encompass a wider variety of concepts, and chapter 9 will apply the SSC to an avionics architecture which is at the limit of the definition.

7.3. Developing the Continuum

The SSC has been developed gradually over a number of years, emerging slowly from the Creating System presented in Chapter 5. It has been tried in use in industry, evolving in the process and so its progress may be described as action research

User populations were originally design engineers, but more latterly the SSC has been used with non-engineer students in academia. The students were a mixed-sex, mixed ethnic-origin, mixed interest group concerned with the study of systems and management. Their surprisingly-ready grasp of the methods and approach inspired the writing of this thesis.

It is concerning to note that industry, comfortable enough with Creating Systems, presently has little opportunity to apply Conceiving Systems in the real world at present. Generally, the conception of a system is the responsibility of a customer organization while its design and development falls to a contractor. Thus in UK defence, for example, the Ministry of Defence Operational Requirements branches, populated principally by operator / users, detail the requirements for new systems, while MOD Procurement Executive branches let contracts for the study and design & development of those systems. This division might be counter-productive, but in the absence of an accepted Conceiving System, it is the present, wide-spread practice. It is, in my view, one of the major, if not the principal, cause of dissatisfaction with procured systems.

7.4. The Seven Steps at Level Zero

The Seven Step Continuum suggests, perhaps, that the seven steps themselves form a continuum. That is not quite accurate. The seven steps form a progression from soft to hard, from issue to solution (if there is one), but each step is quite large - too large to be termed a continuum. Each step could be viewed as a set of lesser steps, and each of those decomposed further. The SSC does not operate in quite that formal, reductionist manner, either. Instead the steps are "magnified", revealing more and more detail with increasing magnification. This differs from reductionism in that "smaller" ideas become visible at higher magnification, and are not necessarily simple decompositions from the higher level.

First, the Seven-Step Continuum at the highest level, level zero:-

STEP 1 - Understand the Issues

STEP 2 - Establish the SOI Requirement Template

STEP 3 - Develop Process and Structure — Softer

STEP 4 - Assign Man-Machine Roles

STEP 5 - Develop SOI Performance

STEP 6 - Develop SOI Effectiveness — Harder

STEP 7 - Develop the Preferred Design Concept

Seven-Step Continuum from Soft to Hard

Much of the content of the seven steps has been presented in the previous chapter and in the enclosed papers, particularly *Managing System Creation.*

However, some explanation of the step structure and sequence is warranted

The steps clearly move from issue to design concept - the intention is to find a solution, the goal is to attain a result, even where the result is that no sensible solution exists even at concept level.

Steps are taken in sequence, starting a Step 1.

There being no established consensus on what constitutes a design concept, Steps 5, 6 and 7 develop successive degrees of robustness in the concept. It is, for some purposes, feasible to stop at Step 4, having become satisfied with the level of understanding for a particular design concept. For others, stopping at Step 5 will satisfy. The full seven steps will provide a well justified, traceable design concept where one exists.

The steps do not, essentially, presuppose an SOI with objectives; in the real world, it is common for systems to be loosely prescribed by customers without apparent, or at least declared, objectives or with the declared objectives quite different from the real objectives which have to be inferred.

The SOI, or System of Interest, is the complete IDA system and *at all times* includes the people and their machines; even when it is necessary to distinguish between the contained (human) decision forming system and the decision support system which may (but need not) be technologically based, measures of effectiveness will continue to regard the system as one.

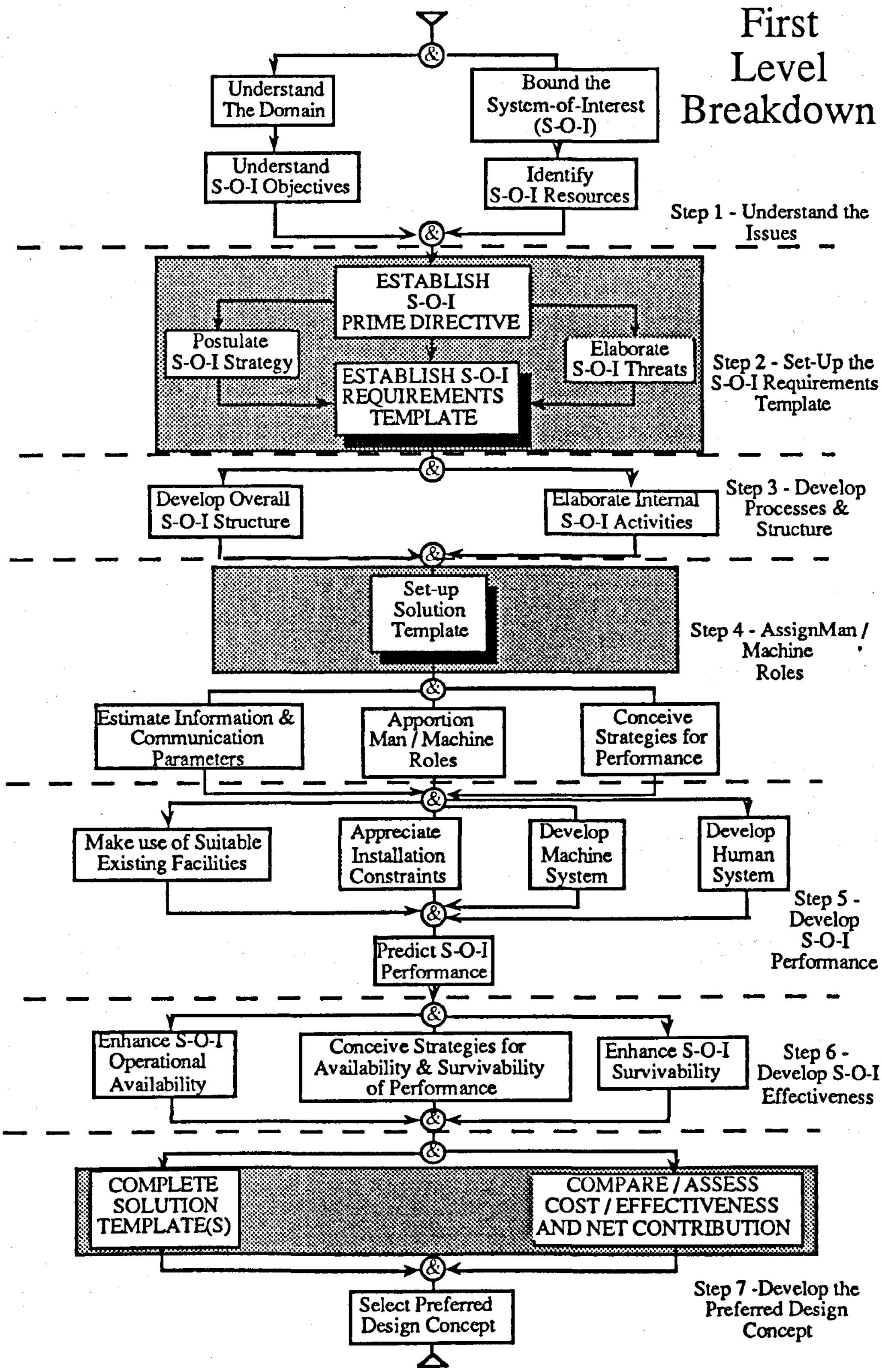
The SOI is influenced by parent and sibling systems, but it is not those systems. There is a relatively clear demarcation around an IDA system; the influences on the SOI are experience through its many interfaces

Some of the steps require understanding, technique and method which may not presently be available either in an acceptable form, or in a form for which there is consensus approval. None-the-less, it is possible to delineate and describe the characteristics of such needs; indeed, it is useful and important so to do, in order to direct attention to the shortfall in tools and techniques.

The development of Performance is seen to precede that of Effectiveness. This sequence is pragmatically based; there is no value in developing effectiveness in a system concept that lacks performance.

7.5. The Seven Steps at Level One

The first level of resolution is as follows:-



The individual steps will elaborated in the following paragraphs. Note, at this first level of magnification, the use of a restricted set of action words, at the beginning of each task; typical action words and their meanings follow:-

Understand	Gain an in-depth knowledge of all facets
Bound	Describe the limits of the system
Identify	Generate or discover, recognize, describe, categorize
Postulate	Propose, put forward
Elaborate	Amplify, expand, decompose
Develop	Create, expand, detail
Estimate	Calculate numerically, approximately
Conceive	Create, generate the idea
Appreciate	Survey and gain an understanding
Enhance	Add features, capability
Formulate	Develop justifiable, traceable rules
Cluster	Group according to formulated rules
Predict	Forecast on the basis of model or parametric analysis
Map	Transfer viewpoints
Compare	Model options
Etc	

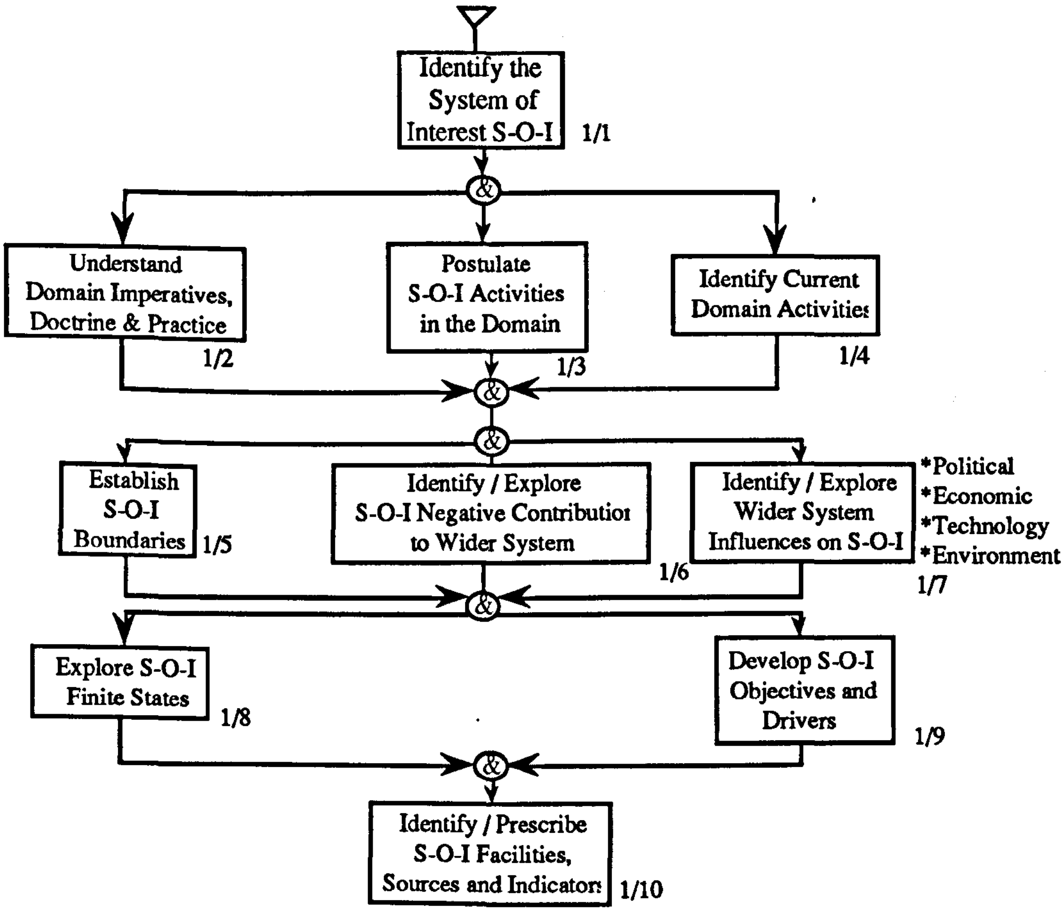
Some restriction to the range and meaning of action words is a useful adjunct to the formality of the metasystem process; it militates against misunderstanding without at the same time introducing a set of jargon terms that would deter the non-specialist or newcomer.

Note in the figure above that the Requirement Template is set up as the output of Step 2 - in fact, Step 2 is devoted to setting up the Requirement Template on the basis of work undertaken in Steps 1, and 2. The Requirement Template is a hinge-pin of the SSC, directing earlier work and setting the stage for the the Solution Template of Step 4, which provides a similar hinge-pin for the remainder of the design concept formulation process going as far as is needed

7.6. The Seven Steps at Level Two

This paragraph will briefly present each of the seven steps in turn at the second level of resolution or magnification

Step 1



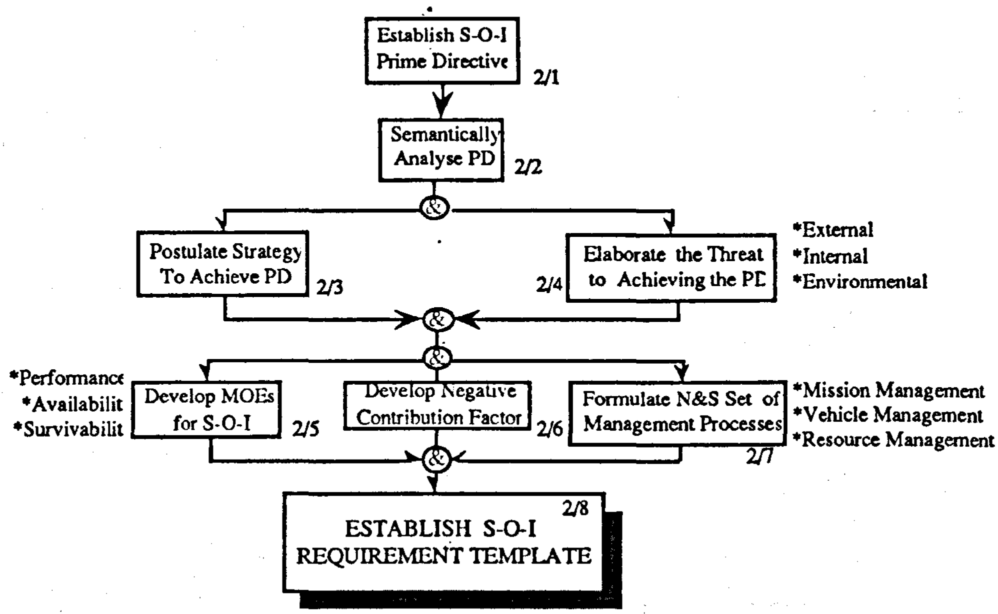
STEP 1 - UNDERSTAND THE ISSUES

The first step explores and examines the full range of influences in the domain, and in the wider political, economic, technological and environmental arenas. The SOI either exists as a concept (1/1) or as part of the set of domain activities, (1/3 & 1/4).

It is essential to fully comprehend, in the fullest sense, the domain imperatives, doctrine and practice (1/2) and it is here that the fallibility of knowledge elicitation comes into focus. Can it be possible for someone not immersed in the domain to capture not only the spoken and written, but the understood word of a particular domain in which the expert may have spent a lifetime and yet still be learning. Knowledge elicitation is a vicarious activity at best and - well-intentioned though the practitioners may be - it can surely be rated no better than second to the expert's deployment of his personal understanding.

Note also the deliberate opening up of the issues of Negative Contribution to the parent system and the exploration of differing finite states in which the SOI might be required to exist, co-exist, operate, etc; both are examples of Creative Entropy being introduced to expand out the issues at an early stage

Step 2



STEP 2 - ESTABLISH THE S-O-I REQUIREMENT TEMPLATE

The second step establishes the essential Prime Directive, from and to which all subsequent results should be traceable. The PD is semantically analysed to expand understanding and viewpoint within the (supposed) team of concept developers. Creative Tension can be seen in (2/3 &2/4) in Postulate a Strategy and Elaborate the Threat to Achieving the PD

Step 2 adds Creative Entropy in the form of Necessary and Sufficient sets, in three discrete areas:-

Threat - {External, Internal, Environmental}

Measures of Effectiveness - {Performance, Availability, Survivability}

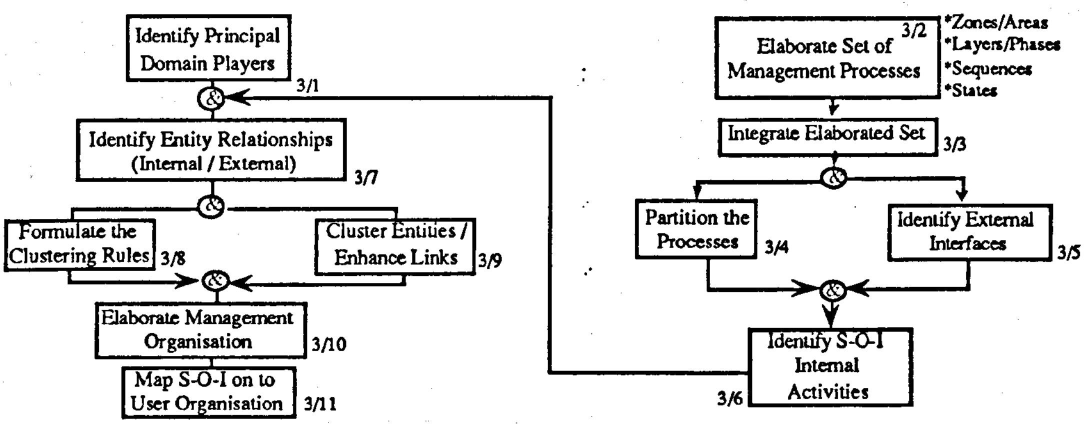
Management Processes - {Mission Management, Vehicle Management, Resource Management}

There is also development of the negative contribution factors identified in the first step.

It would be quite usual for a design concept to fail at step 2. To make the process more robust, several strategies and threats (2/3 & 2/4) would be preferable in step 2, allowing not only for a better prospect of uncovering the unexpectedly good concept, but also providing insight to those factors which discriminate between options. *System Creativity*, enclosed, presents an example of variety in strategy options resulting in an unexpected bonus in the choice of company divisional structures

The Requirements Template has been shown in the prior chapter; it is the output from step 2.

Step 3



STEP 3 - DEVELOP
PROCESSES & STRUCTURE

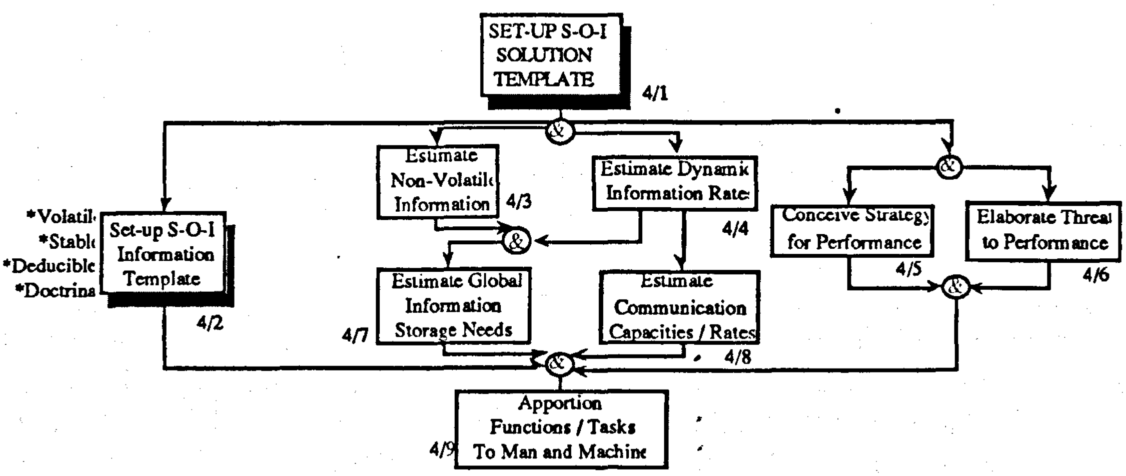
It is in Step 3 that structure, balance and flow make their first appearance.The Set of Management Processes is elaborated into a full description of all the activities taking place within the IDA system, without regard to internal interface or boundary. Flow diagrams of various sorts, thread diagrams and - as shown in the previous chapter - data state diagrams are but a few of the methods. This process of elaboration, which is essentially creative, requires excellent domain understanding if the domain practices and imperatives are not to be inadvertently contravened.

The elaborated processes for each of the many activities which might be represented in each of the Set of Management Processes can all be joined together (3/3); this is essentially so, due to their origin as related activities within the Necessary and Sufficient Set bounds. Joining them may produce a large diagram, but it is generally a worthwhile procedure since the result represents all the functionality of the relevant design concept. It is also worth integrating the processes in order to partition the SOI at its boundaries

The elaborated processes of necessity incorporate entities which are external to the SOI. The boundary of the SOI is superimposed, (3/4), and interfaces between the SOI and its parent and sibling systems can be established at that point, at least in terms of category and priority of information. At this point we may perform a degree of aggregation by grouping the internal processes into generic Internal Activities which are generic because they are repeated in different processes. Typical internal activities would usually include: information collection, operations management, contingency planning, resource management, etc.

Relationships between internal and external entities are identified, (3/1 & 3/7), as the beginnings of architecture analysis. Architecture is formed initially within the SSC by grouping entities, (3/8 & 3/9), which are principally activities, both within and external to the SOI. Examples will be shown in Chapter 9. The term "Management Organization" is used instead of architecture, (3/10), as more usual, but the organization of the system operators / users / managers is undoubtedly architecture within my definition.

Step 4



STEP 4 - ASSIGN MAN-MACHINE ROLES

Step 4 commences by setting up the solution template, based on previous work, to guide all further activity until the end of Step 7; Step 4 then addresses information, communications and performance. Creative Tension is invoked at 4/5 and 4/6 above, and 4/2 introduces another N&S Set, this time for the information which will be the life blood of the eventual IDA system. - {volatile, stable, deducible, doctrinal}. An information template from *A General Theory of Command and Control,* enclosed, is reproduced here:-

	Stable	Dynamic	Deducible	Doctrinal
Performance	Terrain Action Element Capabilities Interoperability Orbats Plans	Positions Strengths Going Weather Plans Engage/Disengage/ Move	Positions Balance of Forces Threats Rates of Closure Weaknesses Enemy Intentions Penetrateion and Egress	Cost-Exchange Ratios SOPs Enemy Tactics Tactics & Strategy
Survivability of Performance	Availability of Alternates Self Defence Damage Tolerance Signature	Alternate Status Self defence Status Emcon Status	Attrition Overrun Forecast Action Element Strengths	Mobility vs Armour Attack/Defend Ratios Stand/Withdraw Ethic SUCOC/COLOC
Availability of Performance	Reserves Replenishmnet Cycles	Warcon, POL Spares Transport Readiness	Combat Days Remaining	Predicted Conflict Duration

Information Template

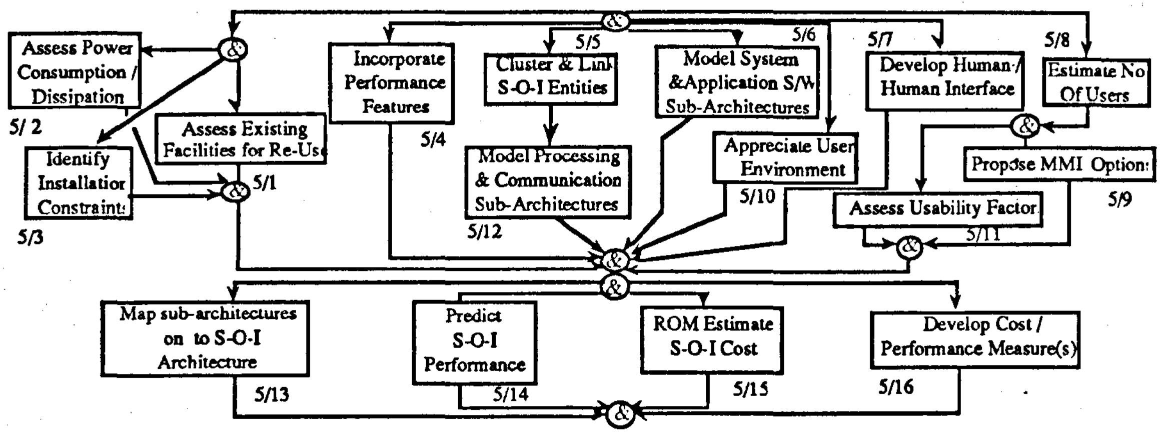
The Information Template, as the figure shows, is the meshing of two N&S Sets since the row headings also form the, by-now-familiar, set of Performance, Survivability and Availability.

Step 4 also contains, at (4/9), the task "Apportion Functions / Tasks to Man and Machine". Much study has been undertaken in the area of Man-Machine Interface, Human Computer Interaction and the like, but this task still presents problems. For IDA systems which operate "near-real-time", i.e. they are on-line, transaction-based rather than in a real-time loop where failure to process in the allocated time results in system failure, the present practical approach is based on treating the information handling system as "very dumb". Thus the machine is used to handle high volume, highly-repetitive tasks requiring not only no intelligence, but no substantial calculation, concentration being on simple processes, sorting, presentation, aggregation, etc.

For real-time systems, the situation is more severe; the IDA system has a set period, usually part of a cycle, in which to complete all activities. It is best to concede at this stage that task 4/9 is beyond the aims of this thesis, and is a specialist subject, the study of much research in its own right.

Step 5

Step 5 is the most complex in the SSC, but owing to its interwoven nature, with both human and machine elements of the SOI being considered in parallel and together, the complexity is inevitable. For many applications, this step will decide the validity or otherwise of system design concepts

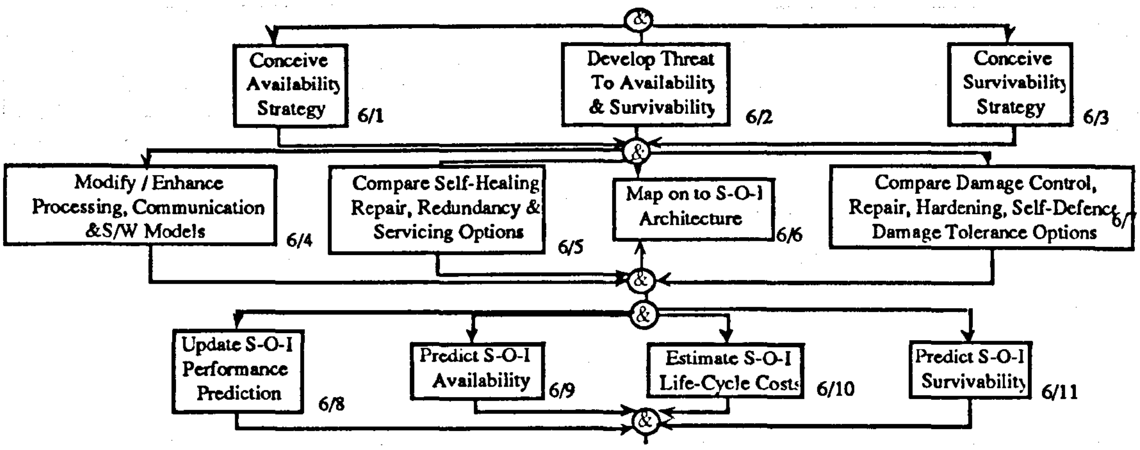


STEP 5 - DEVELOP
S-O-I PERFORMANCE

At the left are the power, installation and existing facility constraints endemic to any viable IDA concept.In the centre is the group of activities associated with processing and communications sub-architectures. At the right the human system is developed. Along the bottom, the elements are brought together again, (5/13), to predict performance, estimate cost and develop cost/performance measures which may be used to compare optional design concepts.

Step 6

Level 6 is concerned with well-understood concepts of availability and less-well understood concepts of survivability, see *System Survivability Science*, enclosed



STEP 6 - DEVELOP S-O-I EFFECTIVENESS

Step 6 starts by Creating Tension, between strategies and threats for both availability and survivability. By definition, of course, their threats are quite different, the one being inherent in the system. its components and their environments, while the second is due to hostile acts.

The content of individual tasks within Step 6 will not be pursued further here, since they are largely standard practice except, perhaps for self-healing system design, which is at the cutting edge of design concepts at present.

Step 7

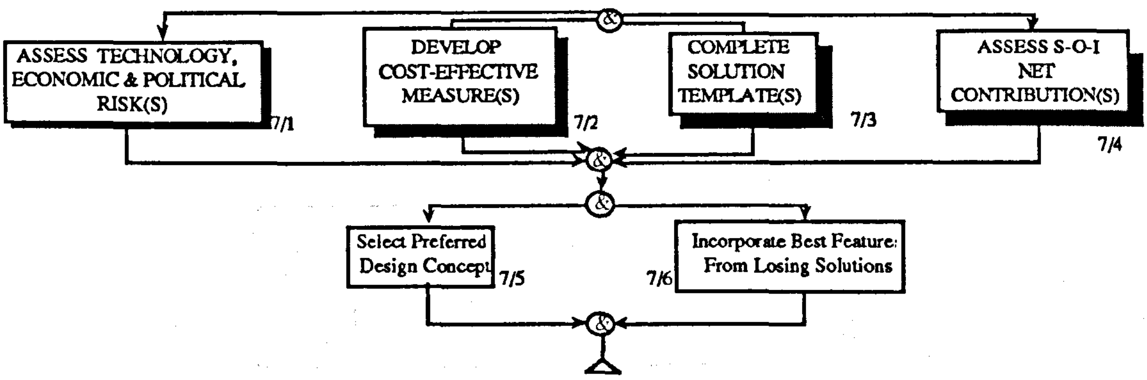
The final step is concerned with trading between design concept options to select the best, particularly in respect of the following:-

Risk, economic, political and technological, which has been carried through from task 1/6 in the first step

Cost-Effectiveness as a means of ranking options

Net Contribution, generally to the parent system, to establish whether *any* of the putative solutions is worthy.

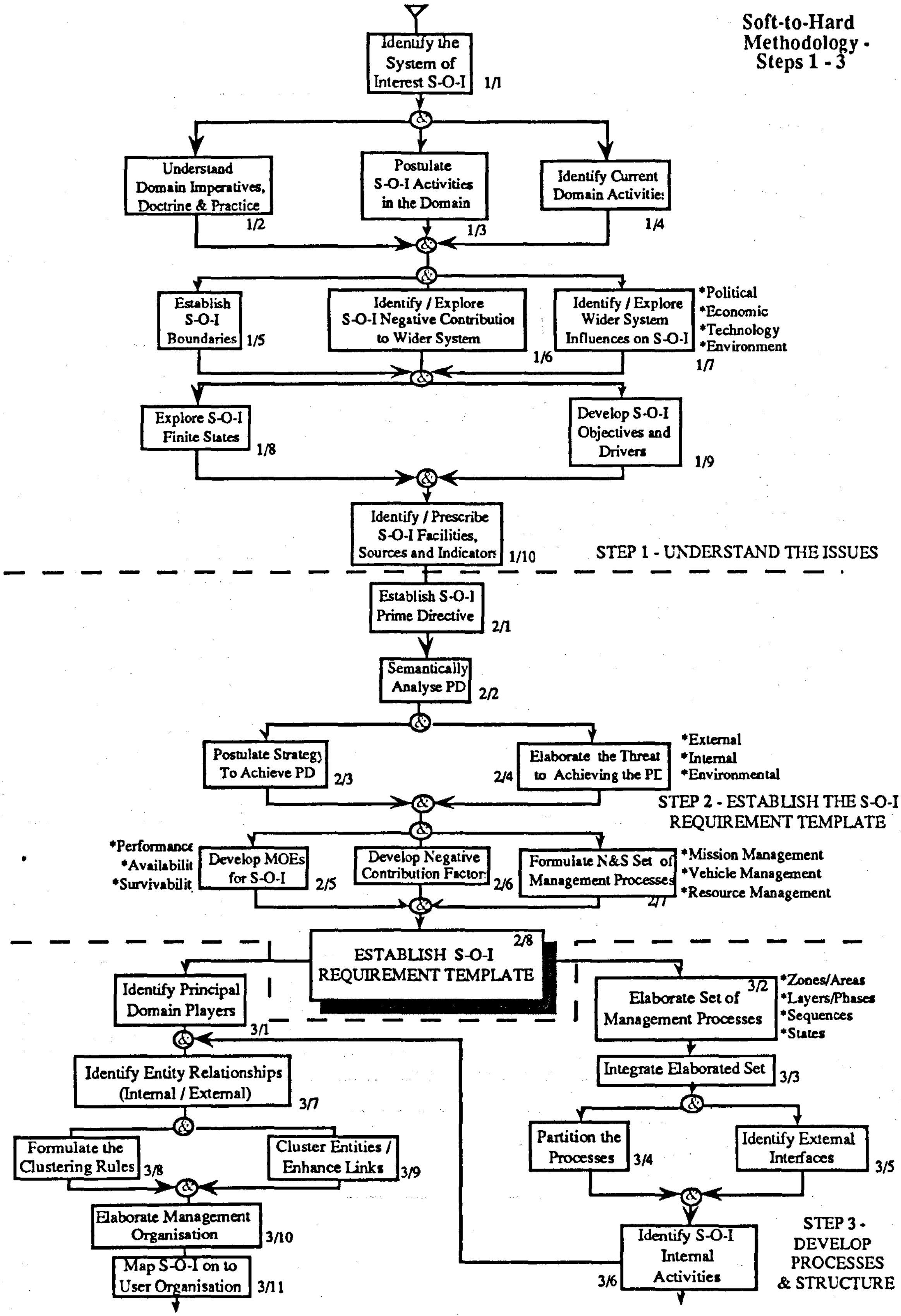
Incorporating those features from losing concepts into the selected preferred option - which may, of course, require its design to be substantially re-balanced.



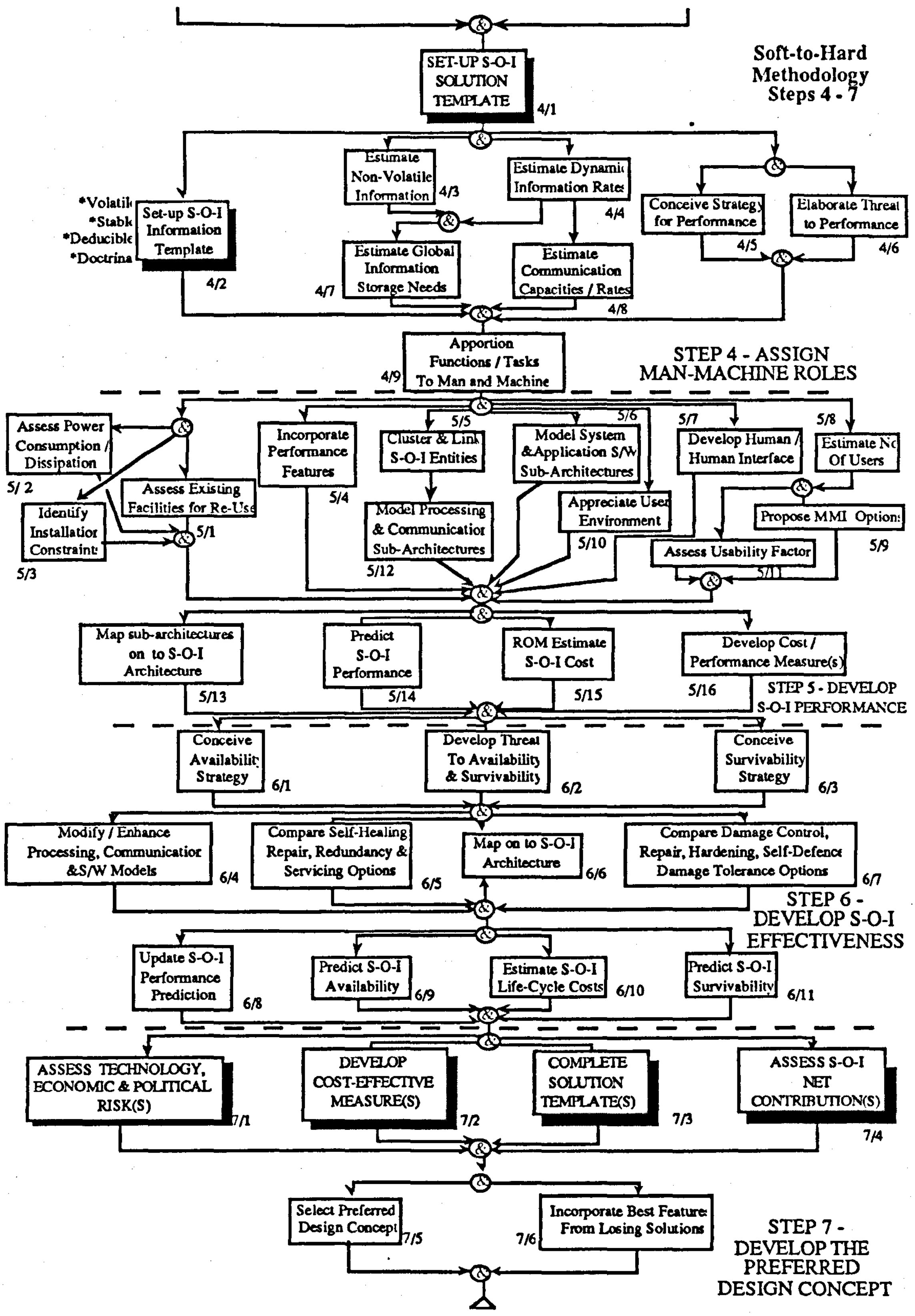
STEP 7 - DEVELOP THE PREFERRED DESIGN CONCEPT

The Seven-Step Continuum Integrated

It is, of course, possible to join the seven steps at level 2 into one larger network, enclosed at the end of the text.



Soft-to-Hard Methodology - Steps 1 - 3
Identify the System of Interest S-O-I 1/1
Understand Domain Imperatives, Doctrine & Practice 1/2
Postulate S-O-I Activities in the Domain 1/3
Identify Current Domain Activities 1/4
Establish S-O-I Boundaries 1/5
Identify / Explore S-O-I Negative Contribution to Wider System 1/6
Identify / Explore Wider System Influences on S-O-I 1/7
*Political
*Economic
*Technology
*Environment
Explore S-O-I Finite States 1/8
Develop S-O-I Objectives and Drivers 1/9
Identify / Prescribe S-O-I Facilities, Sources and Indicators 1/10
STEP 1 - UNDERSTAND THE ISSUES
Establish S-O-I Prime Directive 2/1
Semantically Analyse PD 2/2
Postulate Strategy To Achieve PD 2/3
Elaborate the Threat to Achieving the PD 2/4
*External
*Internal
*Environmental
STEP 2 - ESTABLISH THE S-O-I REQUIREMENT TEMPLATE
*Performance
*Availability
*Survivability
Develop MOEs for S-O-I 2/5
Develop Negative Contribution Factors 2/6
Formulate N&S Set of Management Processes 2/7
*Mission Management
*Vehicle Management
*Resource Management
ESTABLISH S-O-I REQUIREMENT TEMPLATE 2/8
Identify Principal Domain Players 3/1
Elaborate Set of Management Processes 3/2
*Zones/Areas
*Layers/Phases
*Sequences
*States
Integrate Elaborated Set 3/3
Partition the Processes 3/4
Identify External Interfaces 3/5
Identify S-O-I Internal Activities 3/6
Identify Entity Relationships (Internal / External) 3/7
Formulate the Clustering Rules 3/8
Cluster Entities / Enhance Links 3/9
Elaborate Management Organisation 3/10
Map S-O-I on to User Organisation 3/11
STEP 3 - DEVELOP PROCESSES & STRUCTURE


Soft-to-Hard Methodology Steps 4 - 7
SET-UP S-O-I SOLUTION TEMPLATE 4/1
*Volatile
*Stable
*Deducible
*Doctrinal
Set-up S-O-I Information Template 4/2
Estimate Non-Volatile Information 4/3
Estimate Dynamic Information Rates 4/4
Conceive Strategy for Performance 4/5
Elaborate Threat to Performance 4/6
Estimate Global Information Storage Needs 4/7
Estimate Communication Capacities / Rates 4/8
Apportion Functions / Tasks To Man and Machine 4/9
STEP 4 - ASSIGN MAN-MACHINE ROLES
Assess Power Consumption / Dissipation 5/2
Identify Installation Constraints 5/3
Assess Existing Facilities for Re-Use 5/1
Incorporate Performance Features 5/4
Cluster & Link S-O-I Entities 5/5
Model System &Application S/W Sub-Architectures 5/6
Develop Human / Human Interface 5/7
Estimate No Of Users 5/8
Propose MMI Options 5/9
Appreciate User Environment 5/10
Assess Usability Factor 5/11
Model Processing & Communication Sub-Architectures 5/12
Map sub-architectures on to S-O-I Architecture 5/13
Predict S-O-I Performance 5/14
ROM Estimate S-O-I Cost 5/15
Develop Cost / Performance Measure(s) 5/16
STEP 5 - DEVELOP S-O-I PERFORMANCE
Conceive Availability Strategy 6/1
Develop Threat To Availability & Survivability 6/2
Conceive Survivability Strategy 6/3
Modify / Enhance Processing, Communication &S/W Models 6/4
Compare Self-Healing Repair, Redundancy & Servicing Options 6/5
Map on to S-O-I Architecture 6/6
Compare Damage Control, Repair, Hardening, Self-Defence Damage Tolerance Options 6/7
STEP 6 - DEVELOP S-O-I EFFECTIVENESS
Update S-O-I Performance Prediction 6/8
Predict S-O-I Availability 6/9
Estimate S-O-I Life-Cycle Costs 6/10
Predict S-O-I Survivability 6/11
ASSESS TECHNOLOGY, ECONOMIC & POLITICAL RISK(S) 7/1
DEVELOP COST-EFFECTIVE MEASURE(S) 7/2
COMPLETE SOLUTION TEMPLATE(S) 7/3
ASSESS S-O-I NET CONTRIBUTION(S) 7/4
Select Preferred Design Concept 7/5
Incorporate Best Features From Losing Solutions 7/6
STEP 7 - DEVELOP THE PREFERRED DESIGN CONCEPT

Summary

The chapter presents and describes the Seven Step Continuum at three levels of decomposition, defines the terms used, and ends with a complete, third level network of tasks which identifies the sequence and relationship between activities in the development of a robust concept for a system, should such a concept exist. Tools and techniques for undertaking the tasks were not presented, those being topics for the following chapter.

8. TOOLS AND TECHNIQUES FOR CONCEIVING SYSTEMS

8.1. The GEC Survey

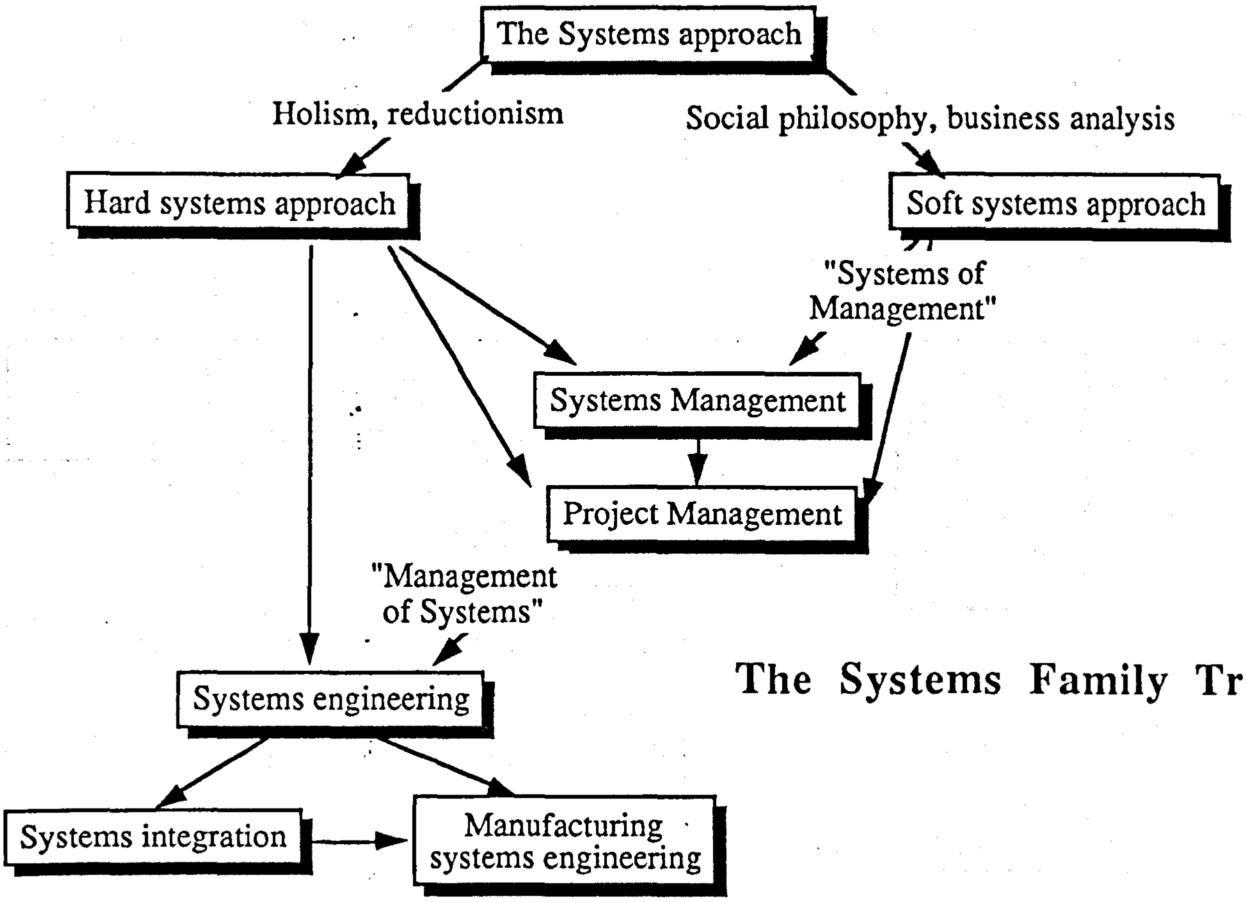
In 1987 I chaired a General Electric Company (GEC) Working Party, tasked as follows:-

> "To assess the models and modelling techniques available within the GEC product units and on the market for systems engineering
>
> "To assess the models and techniques not yet available"

The three-month working party covered the wide range of GEC company, and external, systems engineering and model potential in some three months. Much of the report is confidential, but some aspects are not.

First, the unwritten task was to explore the potential of existing models and tools - the terms being virtually synonymous in this context - for cross-use between GEC companies. The working party found pockets, as might be expected in such a vast and varied conglomerate, where no systems engineering was undertaken consciously at all, and some of the companies were either unaware of, or positively opposed to, systems engineering. Others had varying degrees of interest in systems engineering, many wishing greater understanding but fearful of cost. Not too surprisingly, the systems companies were the main custodians of expertise and models, but even here there was a wide variation in practice and understanding.

We, the working party, felt that categorization would help to clear the mind, so we developed the following figure:-

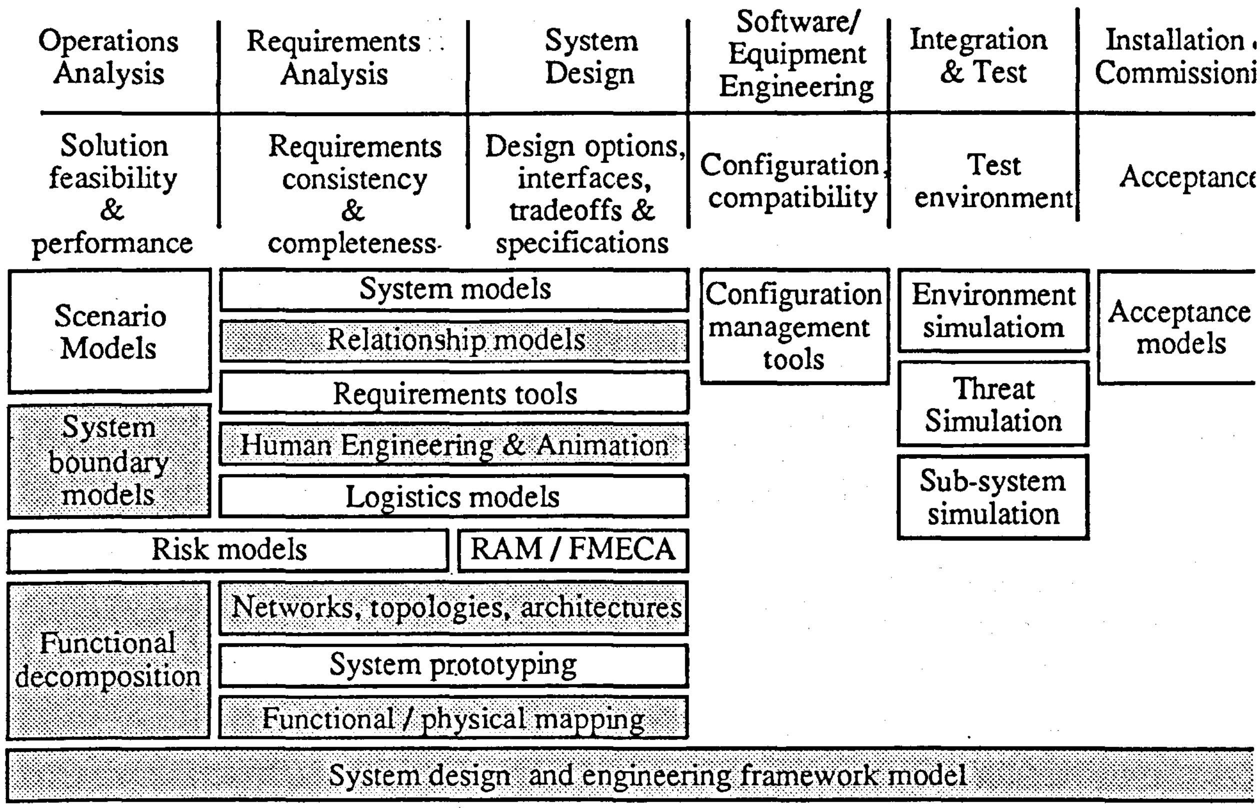


The Systems Family Tr

Although the categorization might be argued on purist grounds, it none-the-less represents a valid industry view. Noteworthy is the subordinate position allocated to systems engineering vis-a-vis project management, for example, and the expression of holism as relevant to hard, but not soft, systems approaches.

The working party developed a taxonomy of tools. It was clear to us that we needed to establish a spectrum of tool "needs" before proceeding to see if the needs were being, or indeed could be, satisfied. The taxonomy is presented below diagrammatically, in the form of a tool shadowboard.

Systems Engineering Tool Taxonomy



The figure requires some explanation. The top row establishes the phases of any project as seen from industry's viewpoint and as developed in *Managing Systems Creation.* The second row identifies the principal objectives of tools and models in the respective phase, so "solution feasibility and performance" is the objective that tools and models might assist in achieving during the Operations Analysis phase. Beneath each of the column headings are the tools / models that the working party, drawn from all quarters of GEC, felt would be of greatest value - regardless of whether they were available, or even realizable.

8.2. The "Missing Tools"

Subsequent investigation showed that the "shadowboard" of tools was not evenly filled. Scenario models, software requirements tools, environment simulation, logistic models and many more were in abundance. (Requirements tools generally support functional decomposition, but do not execute the decomposition themselves - that is in the mind of the operator.) RAM / FMECA - Reliability, Availability, Maintainability / Failure Modes, Effects and Criticality Analysis - was undertaken quite comprehensively. On the other hand, the following were not in evidence:-

System Boundary Models
Functional Decomposition
Functional to Physical Mapping
Relationship Models
Architecture Design tools
System Design and Engineering Framework Models
Risk Models

The shaded boxes on the figure show the research areas represented in this paper; there is a high degree of correspondence, as follows:-

System Boundary Models	System Dynamics
Functional Decomposition	Creative Entropy, N&S Sets
Functional to Physical Mapping	N^2 Chart and cluster analysis
Relationship Models	N^2 Chart
Architecture Design Tools	Various; see later sections and chapter 9
Frameworks	Creating and Conceiving Systems
Risk Models	Not addressed directly

Although the methods, on the right, were found to be available in principle to meet the needs of design, on the left, the methods were not in a suitable form. N^2 charts, for example, can be extremely cumbersome in use, having to be re-drawn manually whenever any entity is moved on the chart. Unlike a rectangular matrix, the diagonal nature of the N^2 chart makes re-drawing a non-trivial task, and effectively renders its use as a means of forming and recognizing patterns impracticable for charts where N exceeds about ten. Unfortunately, it is the larger, complex, charts which are more interesting, and the more likely to contain "hidden" structure.

The use of processing power is an obvious way to address the N^2 charting problem; even here, however, limitations appear. A 25 x 25 N2 chart can be arranged in 0.5 x 25! different ways; it is possible to exceed sensible processing limits with quite modest N^2 charts.

In researching several of these areas, I realized that current methods used in industry were producing rigid, inflexible results. I therefore became interested in how Nature manages to "develop" such flexible, adaptable systems

8.3. Nature's Approach to Complex "Design"

Darwin proposed that Natural Selection is the mechanism by which living things evolve. Natural Selection favours the survival and breeding of offspring which are best-adapted to the changing and predatory environment. The cornerstone of this process is variability between offspring.

Natural Selection is a continuing process - natural systems are evolving now from generation to generation. But Natural Selection has also produced species such as Man that can evolve socially during their lifetimes, regardless of offspring. While lower orders are imprinted at birth with fixed behaviour patterns, higher order animals have less imprinting of rote and more ability to learn and adapt to the contemporary situation. This "lack of fixed rules"phenomenon is itself, undoubtedly evolving too, but at a very slow rate commensurate with the rate of evolution of other characteristics in higher order animals. So, if we have the wit to comprehend, we may be able to emulate Nature in some small degree and develop flexible systems which can adapt to their changing environments during their active lifetimes - a worthy goal.

Cumulative Selection

The key to understanding Natural Selection is cumulative selection. Consider a combination lock; suppose it has, say, three tumblers. If we set the first two tumblers correctly, but miss the third, then we have to start again at the beginning. Nature does not work like that Instead, nature produces a variety in offspring; although all offspring are similar, they all have some dissimilarities, too. Some of these variations favour survival-to-breed better than others, but the survivors do not "start again"; they still retain, and pass on, the vast majority of features which they shared with their siblings. Thus the basic model stays, but features accumulate beneficial changes, and lose disadvantageous characteristics. This process is called cumulative selection and is the basic mechanism underpining the operation of the three tools to be described below.

The school of reasoning to which these tools belong is sometimes called genetic computing. My work in this area is at an early stage, but has already drawn comment from the technical and non-technical press, see The Engineer (1989) and the Financial Times (1989), and from companies wishing to co-operate in the development of the techniques. SYSMORPH, see below, has been the subject of patent action.

*SYSMORPH - morphological design of aircraft, ships, rail carriages,etc

*COMMORPH - physical layout of networks, communication systems, oil pipe-lines, routes of penetration and egress, corridors and pathways, circuit board layout, etc

*TREEMORPH - shortest non-redundant path connecting a set of points

*CADRAT - functional-to-physical mapping for organisational and management structures, systems engineering of advanced avionics architectures,etc

CAD TOOLS & THEIR APPLICATIONS

The tools and their applications are explained in the enclosed papers:-

Creative Design by Cumulative Selection looks at the broad perspective of design tasks which cumulative selection could undertake from the design of rail carriages and airport lounges to communications networks, printed circuit board designs and railway carriages

BM / C^3 Design by Cumulative Selection looks at the application of cumulative selection to the ab initio design of complete, complex platforms, such as an aircraft or a ship and to communications networks; the paper presents the idea of systolic cumulative selection, see below

Avionic System Architecture examines the fundamental characteristics of avionics systems and generates sets of clustering and repelling influences that together promote different types of avionics architecture

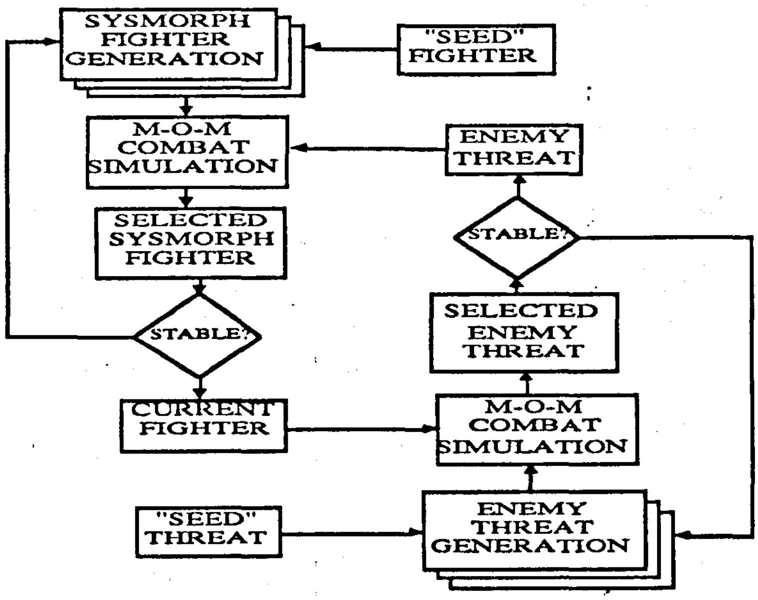
8.4. SYSMORPH

SYSMORPH attempts to emulate in a very small way how evolution might work. A model of a system is drawn by a standard drawing routine from a set of "genes". Each gene codes for a variety of physical and behavioural characteristics in the model entity, which might represent a ship, an aircraft, a train, a car, a yacht, etc.The standard drawing routine then draws a set of offspring, each of which is the same as the original - the parent - except that one of the genes has been incremented or decremented.

The offspring are passed through a simulated hostile environment and their relative performance, survivability and availability are rated, scored and aggregated. There is also a cost model in SYSMORPH which aggregates the cost for each offspring. The offspring producing the greatest cost-effectiveness ratio is selected to breed in its turn.

It might be thought that interest would centre on the final configuration produced by a SYSMORPH run. However, the route by which the program reaches its destination can be just as interesting. The program can be used to identify the first, and by definition, most cost-beneficial, change to be made to an existing design.

The SYSMORPH idea is exciting. Suppose, for example, it were possible to evolve the hostile environment? This approach, executed in a limited fashion only at present, I call systolic cumulative selection. Consider the following figure, which concerns itself with developing the initial design concept for a new fighter aircraft:-



SYSTOLIC CUMULATIVE SELECTION

The process of producing offspring and subjecting them to a hostile environment is shown at the left and proceeds until a stable fighter design is reached. The hostile environment contained a model of an enemy fighter, in a Many-on-Many (M-O-M) combat model. Suppose now that the positions are reversed. The evolved fighter design is frozen and the enemy fighter is allowed to evolve in the same M-O-M combat model. Once its design plateau has been reached, the process can revert to the original state, to evolve the original fighter even further.

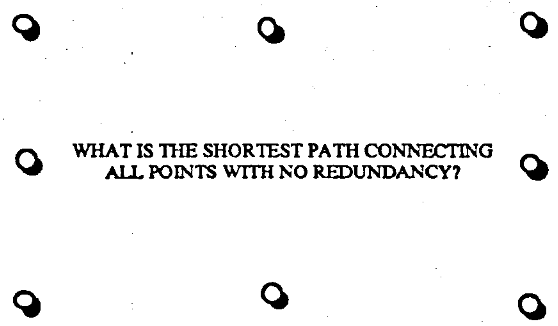
Here we have the potential to undertake some advanced design concept development. In the fighter world, each new design on one side of the cold conflict results in a counter-design some ten to fifteen years later by the opposing side. Using SYSMORPH systolic cumulative selection it may be possible to anticipate the counter. There is, of course, the potential for runaway, with assumption being piled on assumption.

Such risks are evident, and safeguards can be taken. One safeguard is to start from a well-known design point and allow only small, traceable increments in design evolution to occur at each systolic "heartbeat". A second safeguard is in the use of trusted models, developed and validated over many years perhaps, as the simulated hostile environment; these are areas for research.

8.5. COMMORPH

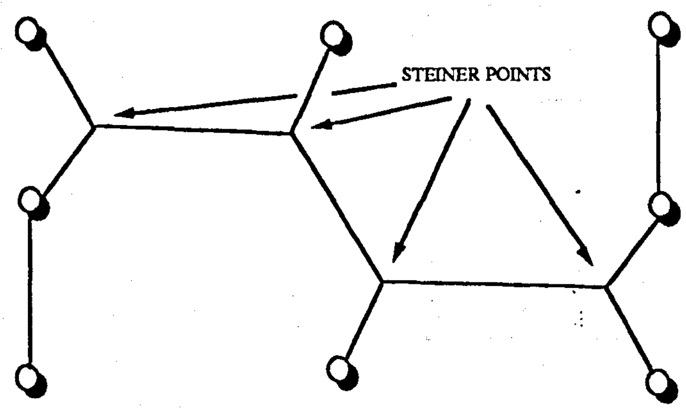
COMMORPH was presented at the Third International Conference on Command, Control, Communications and Management Information Systems in May 1989, and immediately caught the attention of the British Army as a means of laying out mobile and transportable communications systems in real time, in the field.

COMMORPH concerns itself with communication system trees, loops and networks. It too employs cumulative selection, but not in the same way as SYSMORPH. Together with its offspring, TREEMORPH, COMMORPH lays out communication networks according to some optimum objective. The process is best understood by example. Consider the following arrangement of points, which it is desired to connect by the shortest non-redundant route, i.e. no alternative routings permitted.



THE CLASSIC STEINER TREE PROBLEM

As the title suggests, this is a classic problem similar to the well-known travelling salesman problem, and equally difficult to solve by conventional means. Cumulative selection approaches the problem by scattering a number of potential connection junctions into the space between the points, joining junctions and points according to tree rules. The total length of all the connections is measured. Then each junction in turn is taken through a small figure-of-eight "dance", with overall length being measured at eight point during the dance. Junctions are repositioned after their dance at the position on the dance offering shortest overall length. The tree gradually evolves to this:-



STEINER TREE SOLUTION

As with most optimization techniques, COMMORPH and TREEMORPH may provide solutions which are locally minimal but not globally minimal. I have devised a series of "shake routines" which seek to nudge the trees or networks to see if a global minimum exists that is shorter than the current offering. Such an approach carries no guarantees, of course, but practical experience using the tools strongly suggests that local-only optimization is apparent to the user, who may seek the global optimum - if needed - by manual intervention.

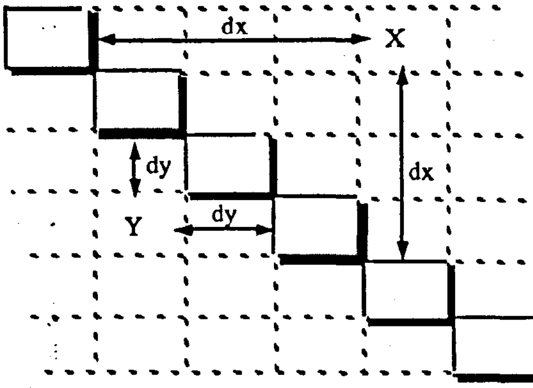
The advantage of COMMORPH comes, not from addressing clean problems such as the Steiner Tree, but from more practical applications The process of optimizing may account for more than simply overall length - see *Creative Design by Cumulative Selection,* where the concept of optimizing to an array of mutually incompatible influences is addressed

8.6. CADRAT

The N^2 chart has already been presented in the section entitled Binding and Coupling and will be seen again in the next chapter. CADRAT, the Computer Aided Design - Relationship Analysis Tool automates the processes involved in drawing and rearranging N^2 charts, which can be extremely tedious for large matrices. But CADRAT does rather more than simply rearrange.

Scoring N^2 Charts

It is possible to introduce the concept of scoring an N^2 chart. To score a chart, one simply determines the distance of each interface from its associated entities in terms of the number of squares, and multiplies the number in the interface by some function of distance. The sum of all such scores for all interfaces makes the N^2 chart score



Score = X * f(dx) + Y * f(dy)

Scoring N2 Charts

The process is illustrated in the figure. The number, X or Y in the figure, represents some characteristic of the interface as viewed from system level, such as strength of association, priority, etc., determined in a systematic marking scheme. Typical functions, f, of distance would be direct proportionality, square or cube. The purpose of these choices will be explained below.

If two interfacing entities are separated on the chart by a number of intervening entities, then the interface joining them will score high due to the distance functions, dx and dy above.If they were adjacent, their interface score would be low. It is therefore possible to arrange the entities on the chart such that the chart has an overall lowest score. In this configuration,all interfaces connecting entities will be grouped such that their scores are low; competition between entities, each trying to occupy the same position relative to a third entity will be resolved by the strength of their respective interface.

The reason for the function of distance can now be explained. If the distance function is made a square instead of linear, then tighter clusters result. The cube function produces even tighter clusters. Examples of the effects will be shown in the next chapter. CADRAT refers to these distance functions as "First Moment", "Second Moment", etc. The reason for the title can now be explained, too. If the N^2 chart is considered like a lattice, then the leading diagonal containing the entities is a central septum. Taking a linear function of distance about this septum equates to taking moments, and is a torque, or first moment function. Squaring the distance equates to an inertia-type, or second moment function. And so on.

Automatic Clustering of N2 charts with CADRAT

The value of being able to score an N^2 chart is insufficient of itself to be really useful, since for a matrix, say 25 x 25, the number of ways in which the matrix can be arranged is tens of millions. CADRAT uses Cumulative Selection and the scoring system to rearrange the matrix until it adopts the minimum score configuration The resulting configuration clusters associated entities according to the strength of that association.

The ability to cluster automatically, and according to a selection of simple rules, presents great architectural richness. Entities present emergent properties not visible in any other practical way, as clusters, often unexpected, appear. The clustering has, however, come about on the basis of relationships and strengths identified individually by the would-be architect, and the CADRAT tool simply reveals the structural implications of his or her supplied information

CADRAT and Hierarchy

Experiments with the CADRAT prototype showed two unexpected features:-

> During the testing phase, I entered known optimal solutions, but the machine occasionally returned seemingly non-optimal solutions

> Replacing the decimal numbers used as strengths with a simpler binary choice made very little difference to the resulting clusters

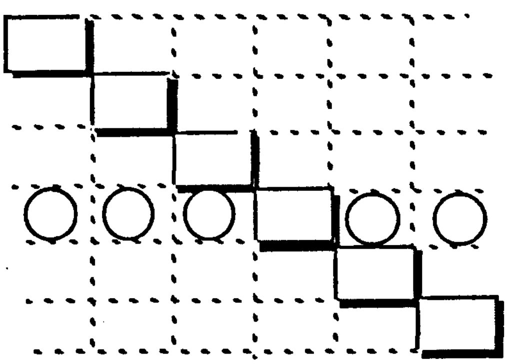
Research into related ideas, Macgill (1983), suggested that the result using binary numbers was to be expected, since the existence of an interface was the strongest influence, and adding a strength number constituted a second order effect. The situation with the seemingly non-optimal solutions was more complex. First, the program behaved correctly; it produced a lower score configuration than the entered, supposedly ideal solution. But second, the program achieved this by placing some entities right in the middle of tightly-bound functionally grouped clusters.

I eventually realized that the program was drawing in the two dimensions of an N^2 chart, but attempting to resolve an issue in three dimensions; the entity being interposed between members of a functional cluster was better viewed as being *above* that cluster.

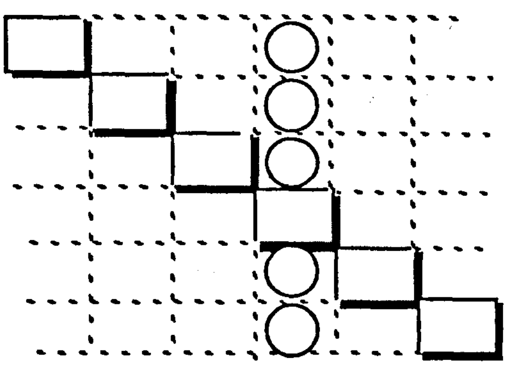
It is therefore possible and useful to use the numbering scheme, not only for its original purpose, but also as a means of examining the vertical, hierarchical structure of an organization or, more generally, of an entity-set The CADRAT program now offers the opportunity to show and cluster a section of the decimal interface numbers. By choosing the numbering scheme to represent, say, rank, seniority, authority, type of communication medium, etc, it is possible to view the N^2 chart either as a whole, or as a set of component charts each with its own purpose, and to investigate clustering at different levels within the same hierarchy.The effect is like being able to slice horizontally through a pyramid and to operate within the slice

Using CADRAT to Develop Architecture

The clustering power of CADRAT enables the user to use his or her pattern recognition abilities in conjunction with the machine. Previous sections on Binding and Coupling introduced tightly-bound functional blocks, system nodes and the waterfall. Other features may be learned and are valuable in instiling structure and balance into the architecture. Consider the following:-

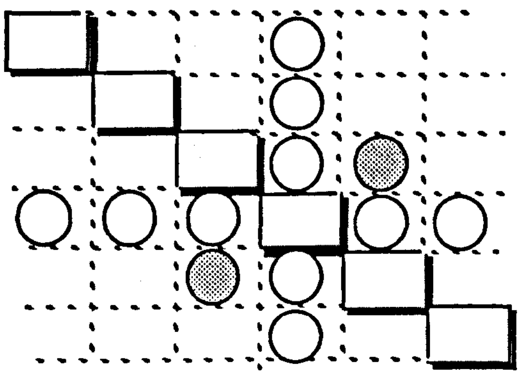


Source Node

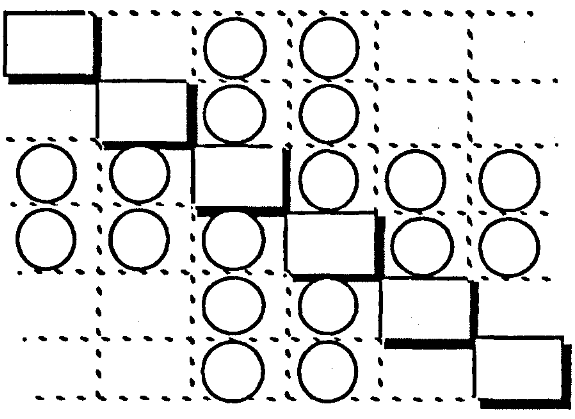


Sink Node

The entity on the row of marked interfaces supplies information to all other entities; it is a source node. The entity at the centre of the column receives information from all other entities, but provides no output. It is a system liability, as drawn.



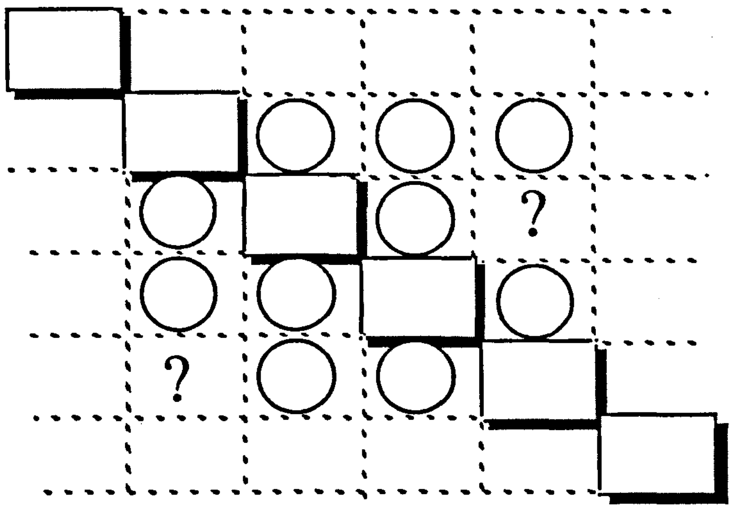
Supported Node



Duplicated Node

The left-hand figure shows a full system node, i.e. the formation of a cross from a source and a sink node. A full system node is of concern because its failure disconnects the system into two separate parts. In this case, however, the shaded interfaces provide some alternative routes for information to flow should the central node become prejudiced

The right-hand figure presents the example of a fully replicated node, a situation sometimes not detected in a large matrix prior to clustering. There are two nodes, side by side; if the system is subject to attack, then replicated nodes may be essential - military Alternate War HQs represent just such a case. Generally, however, the expense of such duplication cannot be justified, and the nodes can either be totally merged or, better, reduced to a supported node



Missing Interfaces

Perhaps the greatest opportunity that CADRAT's clustering offers is the detection of missing interfaces. I have mentioned architectural "balance" before; the figure shows architectural imbalance, since interfaces almost invariably occur in pairs, a "send" interface from entity A to entity B and a "reply" interface from entity B to entity A. These interface pairs appear also as corresponding links in control loops. A concept-developer seeing the above figure would have to question deeply the validity of the missing interfaces. While such missing interfaces can occur occasionally, for example, in the case of remote sensors which are totally open loop i.e. not controlled at all, they are rare and generally non-robust. Where the missing interface refers to human interchange, the absence of an interface in one direction suggests that A is passing information to B, quite unaware of whether it is being received or, if received, whether it is satisfactory. Such lack of feedback causes human links to fade.

The usual practice, then, on seeing missing interfaces either in a node or in a functional block, is to fill them in. Without a tool such as CADRAT there would not be such visibility of clusters, and the human skill at pattern recognition would not be brought into play.

Summary

The chapter first presented a "Tool Taxonomy" for systems engineering, a notional shadowboard showing the tools and techniques which were needed, although not necessarily available or satisfactory.

Nature's approach to complex system design was examined, in terms of natural selection and cumulative selection. Four prototype tools were introduced which sought to emulate Nature's methods in a very simple way, and some of their potential was highlighted.

9. THE SEVEN-STEP CONTINUUM IN ACTION

This chapter is effectively a paper within the thesis, presenting the results of applying the Seven-Step Continuum and the CADRAT tool in particular. The chapter is indicative of the processes, activities and results which would obtain in an integrated SEAMS environment although, in the examples below, each of the activities was carried out separately.

9.1. The three examples

The Seven-Step Continuum (SSC), so called because the sub-steps form an unbroken chain from soft issues to solid design concepts, can be validated as a useful approach by using it against real-world problems. Three different types of design concept development have been undertaken using the SSC:-

Air Traffic Control. The SSC was used with non-engineer university undergraduate students, to explore a well-understood system, that for the control of air traffic at small airfields. This process was undertaken to establish the comprehensibility of the SSC to non-specialists

The Channel Tunnel. The SSC was used with the same group of non-engineers to explore the concept of a Crisis Management system for the Channel Tunnel (Chunnel). No such Crisis Management system exists at present, although I have worked in industry on the design of a putative solution. Confirmation of the validity of the evolved concept was supplied by the Chunnel Head of Security.The use of the SSC in this case was designed to explore its role as a concept generator

A detailed design concept, conducted in industry with engineers, analysts and designers at various levels of experience for an advanced avionic architecture. Validation of solution is not meaningful in this instance, but the efficacy of the SSC as a route-map and concept generator can be assessed.

Extracts from the work undertaken in the three projects will be presented in following paragraphs, to illustrate the SSC in action against the varied problem-set.

9.2. The Seven-Step Progress

Understanding the Issues

The Channel Tunnel stirs different emotions in many people; it is noteworthy that benefits to be anticipated by its introduction are seldom, if ever, alluded to in public debate. An examination of the wider system issues following the SSC framework developed the following set of issues:-

Optional Transport Systems

Reduced Channel shipping
Reduced Ferry income
Increased rail freight > 300 miles
Reduced road freight?
Increased car / passenger traffic

Emergency Services

Terminal road traffic congestion
Ambulance/hospitals/staff
Police / fire facilities
New terrorist target

Infrastructure

Energy Generation
Power Distribution
Road / rail feeder links
Food / waste / sewerage
Fuel supplies
Housing

Ecosystems

Altered traffic patterns
Petrol fume pollution
Urban tunnels
Assembly area conurbations
Countryside noise

Economic Systems

Private Investment
Export / import ratios
Foreign exchange
Wealth generation centres
Exchange dynamics
- speed of money movement

The various issues influenced both the view of the Chunnel's objectives as a purposeful system and the activities that would need to be undertaken. Note that, although the SOI is a notional crisis management system, it is the issues facing the parent system, the Channel Tunnel itself, which are being developed and explored.

Chunnel Objectives and Drivers

Increased cross-Channel Capacity
Weather Independence
Faster long-haul road-rail freight
Flexible / expandable
Upgradeable
Safe
Damage Tolerant
Continuous Operation
Affordable
Profitable Business

The activities which the Chunnel system will undertake include those which might, or might not, fall due to the, as yet unbounded, crisis management system.

Domain Activities

Carrying Passengers	Carrying private cars
Carrying freight	Carrying lorries
Inspecting - freight, passengers, cars, lorries, luggage	
Assembling, ticketing passport checks	Entraining / detraining
Customs	Immigration
Shops and restaurant	General repairs
Maintenance of rolling stock, rail system, tunnel fabric	
Signalling	Patrolling
Guarding	Intelligence gathering
Issue of alerts and warnings	Crisis Management
Emergency repairs	Passenger evacuation, etc

Note how the subject of crisis management arose as a natural progression of thoughts about looking after the passengers and freight and maintaining system momentum

At this stage in the concept evolution imperatives and doctrine should be composed - in extant domains, of course, they already exist.

Channel Tunnel Imperatives and Doctrine

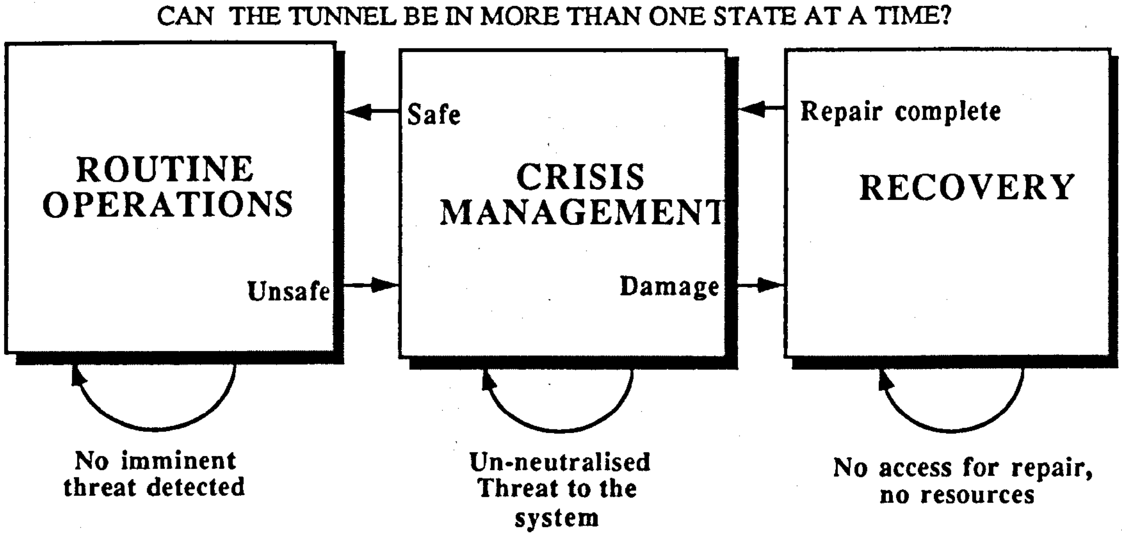
Safety and Confidence first

Performance	Day, night, all-weather, all times
Availability	Reliable - non-interruptive maintenance to tunnel, rolling-stock, etc, dedicated maintenance / repair crews reserve stocks, tools, power, lighting
Survivability	Protected Control facilities Perimeter fencing Dedicated police, fire service? Independent tunnels and controls Sanitized NO-GO zones Damage limitation and repair facilities

The first Necessary & Sufficient set, Performance, Availability and Survivability emerges in the development of the parent system imperatives. It is against parent system objectives,imperatives and doctrine that the net contribution of any design concept will be measured.

Lastly, before turning to the other projects issues, it is worth examining the potential of a crisis management system to occupy various finite states

CAN THE TUNNEL BE IN MORE THAN ONE STATE AT A TIME?



Chunnel Finite States

The conclusion drawn from the deceptively-simple finite state/transition diagram is that the Channel Tunnel could be in several different states at one and the same time, since some threats could be neutralized by localizing them such that they presented no immediate hazard to operations

For Air Traffic Control (ATC) sound domain knowledge is essential; it is possible, however, for students to establish some measure of understanding simply by observing typical small airfield activities, where the complexity and apparent confusion of larger airfields and ATC centres is missing. A typical set of activities might include:

General Flying	Approaching
Landing	Taking off
Marshaling	Controlling Ground Crews
Controlling ground movements	Controlling approaches, take-offs
Controlling airspace	Replenishing
Engine running	Servicing and repairing
Clearing runways (e.g. snow)	Bird scaring
Extinguishing fires	Recovering crashed aircraft
Rescuing endangered aircrew, passengers	Diverting
Forecasting weather	Assisting in airborne emergencies

These and the many other activities routinely carried out at small airfields, lead naturally to the idea of resources (facilities, sources and indicators in the SSC, Step 1). Typical air traffic control resources - remembering that not all of the above activities fall to ATC since we looked at the wider airfield system - include:-

Air space
Taxi-ways
Runways
Approach Aids
Local Aids
Crash facilities
Fire-fighting facilities
Weather forecasting facilities
Area radars
Communications and radio direction finding
Traffic barriers
People barriers
De-icers
Crash barriers
Fuel Bowsers and dumps
Immigration control..............etc

The need for each of these resources can be predicted by considering the activities above - the process is an example of Mindset, introduced earlier, which concentrates the attention on a series of pertinent topic areas in turn. Sensors and indicators necessary to support the activities and implicit in the resources follow naturally in the same Mindset.

Developing the Prime Directive

ATC presents an interesting example of selecting an SOI Prime Directive (PD). There is a clear choice between ATC being an executive authority and being an advisory service. Corresponding PDs might appear as follows:-

"To manage air operations safely and effectively"

OR

"To provide a safe, effective environment for aircraft operations

The present policy, world-wide - i.e. a doctrine - is to vest the ultimate authority for the safety of passengers in the pilot and captain of an aircraft. This approach, sometimes difficult to reconcile with the close control necessary to prevent mid-air accidents, favours the second PD. The PD is analysed semantically as follows:-

"To provide..."	A supporting, resource-providing, non-authoritative role
"...a safe..."	Threats and risks are known, excluded, neutralized or avoided, such as weather, collision, crash, fuel shortage, fire, airborne emergency
"...effective..."	Performs well, is capable, available, survivable/durable/robust
"...environment for..."	All the salient features and facilities relevant to air activities
"...aircraft operations"	Approaching, leaving, local flying, taxiing, parking, loading / unloading, replenishing, repairing, diverting

Aircraft operations may, or may not, include parking and often do not include repairs, so the semantic analysis above reveals issues which must be resolved. The purpose of the semantic analysis is to find and define the boundaries to the SOI's responsibilities - these bounds to responsibility are often major issues. An example for the Channel Tunnel Crisis Management system illustrates the point: in this case, the semantic analysis forms a complete, if long, sentence - a ruse to make it more understandable:-

"To neutralize...	To detect, locate, identify, render safe
..any..	whatever constitutes a hazard
..imminent..	in the near-term
..threats....	to the status quo of
..to the passengers,.........	those legitimately carried by the company ....in exchange for fares,
..staff,....	those in the paid employ of the company,
..fabric,...	the tunnel itself,
..facilities,....	rolling stock, rails, signals, shops, ...offices, franchises, barriers, etc,
..and continued operation......	and the uninterrupted transportation ....of passengers and freight
..of the Tunnel".	within the bounds of legal responsibility

(Notes:- *"Imminent" is unclear
*Sequence of "passengers, staff, fabric, etc" is significant
*"Continue Operations" implies continued revenue and confidence
*No allowance has been made for the franchise workers)

So, the semantic analysis can put rather formal specification boundaries to the SOI. This is, of course both proper and important if the subsequent design concepts are to be useful. The notes, in the above example are particularly interesting - the people working in the franchises appear not to be covered by the Crisis Management PD; the emergence of such issues is one of the major contributions of semantic analysis.

Finally, here is the PD for an advance avionics architecture which shows both PD and semantic analysis as before:-

Semantic Analysis

To provide..............................A service to support...

...an effective..............................performance, availability of performance and survivability of performance, giving......

...means of..................................a vehicle for conveying, processing and handling information in support of.............

...concurrent management.. simultaneous awareness, understanding and husbanding of capability for...

...mission.......................................the end-purpose of the particular set of activities to be effected by the aircraft........

...platform.......................................the aircraft itself in respect of its ability to survive and support the mission and the ....

...resources..................................sensors, emitters, weapons, power, energy, processing, communications, reserves, reversions allocated to the platform and its mission

IMPLIED

The contained architecture should show a net positive contribution to the effectiveness of its containing platform - i.e. any negative aspects, such as carried weight, loss of payload volume, etc should be more than compensated for by the S-O-I

Here the emergence of Net Contribution is apparent in the Implied Section at the bottom, since the architecture could occupy a significant proportion of the aircraft volume and payload - depending on the aircraft type.

The significant difference between a root definition in SSM and a Prime Directive in the SSC should now be evident. The latter is the highest level of abstract description of the system's purpose or raison d'etre, while the former concerns itself with the definition of systems in the "systems-thinking" mode of analysis

Threat Elaboration

Elaboration of the threat to achieving the PD invokes the N&S set of External, Internal and Environmental Threats. A typical elaboration, in this case for the Channel Tunnel, would be as follows:-

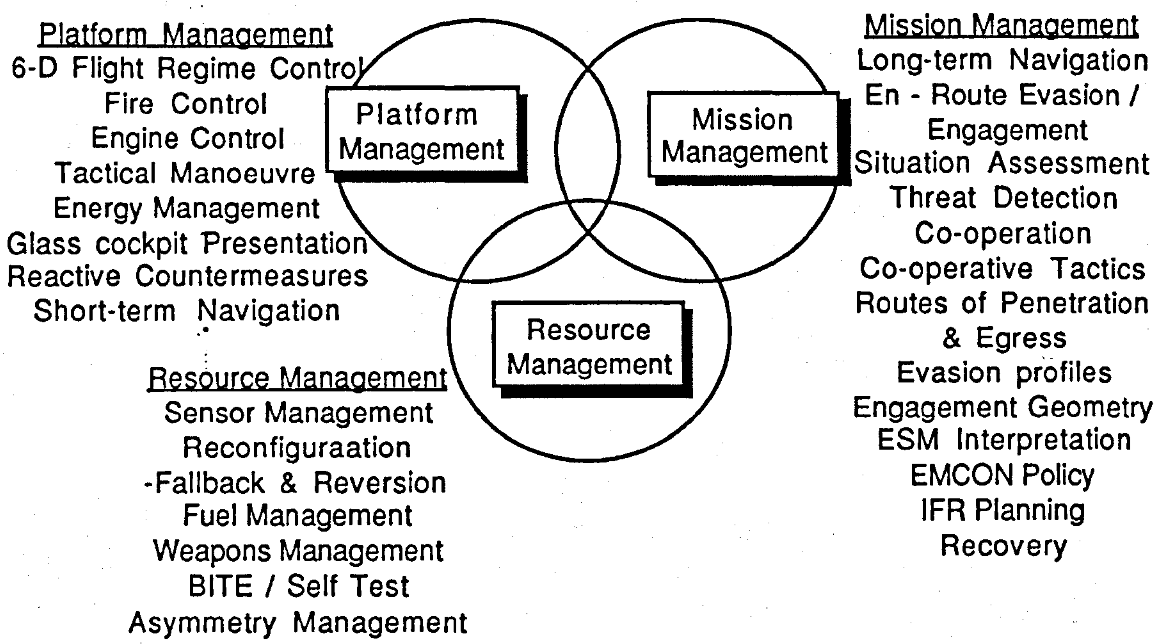
CATEGORY	SPECIFICS	SENSOR
Natural Hazard	Water Ingress	Moisture sensors
	Earth movement	Strain gauges
Accident	Rail	Rail currents, CCTV, Guards
	Aircraft Crash	?
Freight Hazard	Noxious substances	Auto-sniffers
	Explosives	Sniffers, Detectors
System failure	Watertight Bulkheads	Built-in Test (BIT)
	Power distribution	BIT
	Signals / Points	BIT / staff
	Communications	On-line test
	Automatic braking	Rail currents
	Ventilation	BIT & sniffers
Terrorist	Posing as passenger	?
	Carrying weapons	X-ray, detectors, sniffers
	Explosives in luggage	"
	Explosives in freight	"
Disaffected staff	Sabotage	Sensors, patrols, inspections, passwords, initial staff selection

In this example, threat elaboration has been coupled with the ideas of decomposing the broad threat and identifying both the sensors where they can be envisaged and - more importantly - the threats with no sensors

The Set of Management Processes

An example of a set of management processes was given earlier for a notional fire control centre, Below is a set for the advanced avionic architecture project.

Functional Categories



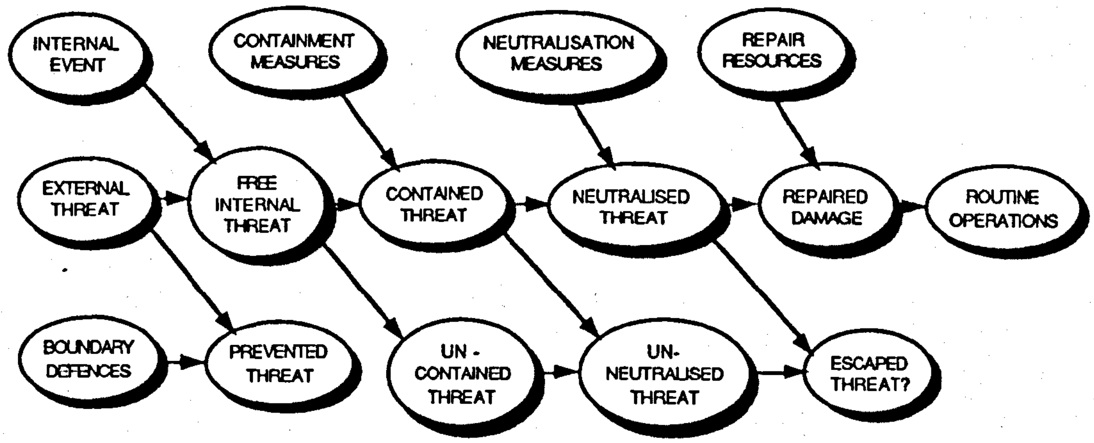
In the figure it can be seen that Mission Management refers to the goals of the sortie, Vehicle Management refers to the aircraft itself, and Resource Management refers to the husbanding of resources in the aircraft in support of both mission and vehicle. In this case, the Management Set can be seen to support a sequence of activities associated with the mission which is evidently directed towards reaching some remote target or rendez-vous. Other projects develop the management set around different schema as follows.

Elaborating the Management Processes

Once the activities of the Set of Management Processes are identified, they may be elaborated by any of a variety of standard techniques, including flow charting and data state design. Since the latter is less common in the UK, the following Channel Tunnel examples use that approach. It is, of course, necessary to have some form of strategy for the elaboration process; generally strategies for managing involve one or more of:-

Zones or Areas of Segregation
Layers of Phases of Operations
Sequences
Finite States

In the case of a Crisis Management system for the Channel Tunnel, a layered defence against the intruder seems most appropriate - if only by trial and error. In such a scheme, a series of defensive screens is envisaged, as follows:-

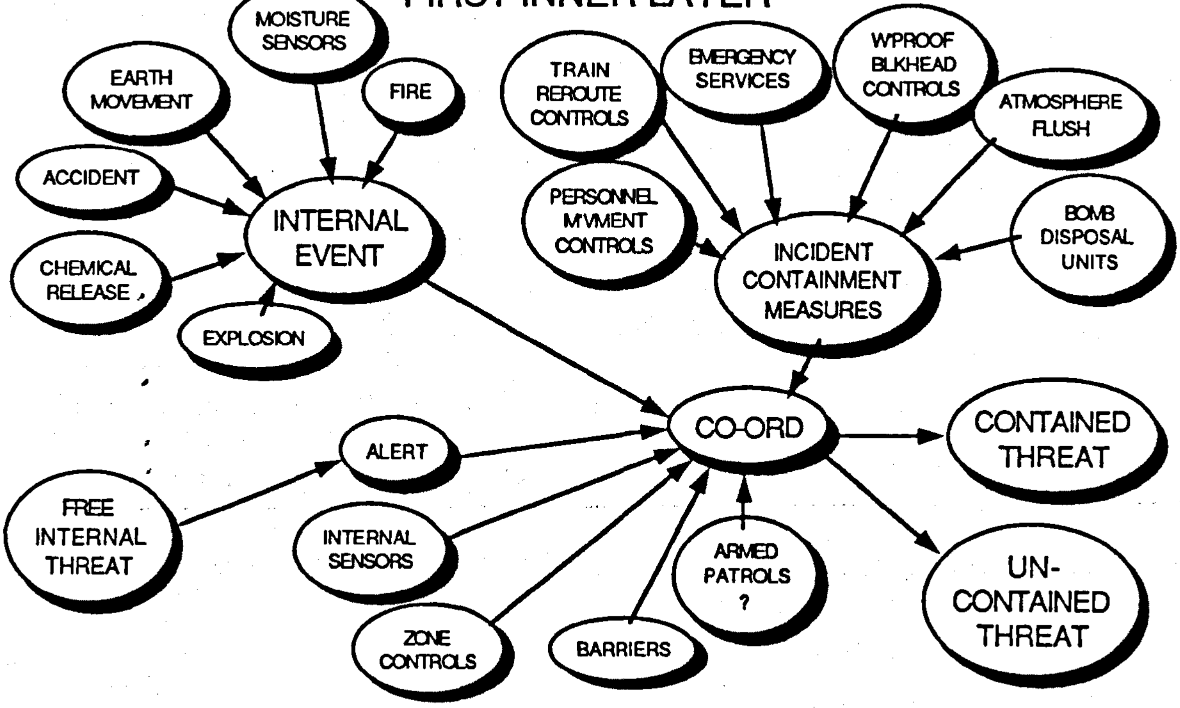


LAYERED DEFENCE

Efforts are first made to exclude the threat. It must be assumed that these will not always be successful. Efforts are then made to contain the free threat, now internal to the SOI. If contained, the next step is neutralization, and so on.

Each of the layers can then be decomposed in turn; The following example shows the First Inner Layer:-

FIRST INNER LAYER



Similar elaboration of the other steps in this case lead to similar diagrams which, since they all emanate from the first diagram in this section, can be joined together into one large diagram describing the complete set of processes. This picture represents the functional description of the SOI to a second level of resolution and provides the basis for subsequent software design. But first, it is necessary to develop an architectural, or structural base on which to develop the functionality.

Developing Organization and Architecture

In any IDA system organization there will be internal and external organization. The major internal activities will be known from the process elaboration; these need to be connected with the external world, the parent and sibling systems. I refer to these external features as the Principal Players. For the Channel Tunnel Crisis Management System they might be:_

Baggage Inspection	Customs
Immigration	Local Police
International Police	Intelligence (part of CM)
Operations (hub of CM)	Security Control and Forces
Emergency Services	Damage Repair & Control
Damage Repair Resources	Logistics (part of CM)
Activity Sensors (CCTV, etc)	Environment Sensors
Safety Controls (Bulkheads, train re-routing, etc)	
Routine Train Operations	

The Principal Players are then grouped as follows:-

Establish an N^2 chart, using intuitive groupings to begin

Create quantitative cell values for each interface, either on a basis of urgency, importance of information, quantity of information, hierarchy or similar according to purpose

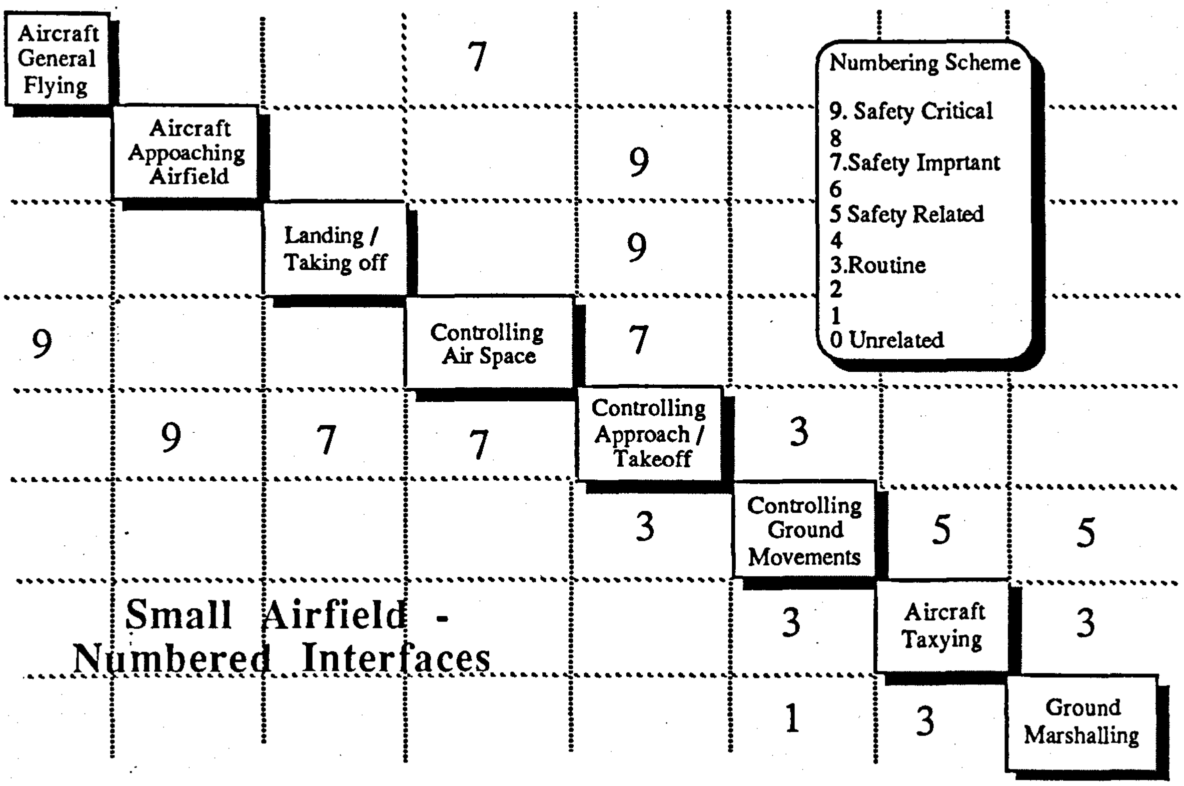
Cluster using weighted cells

Cluster using sets of the un-weighted cells to explore in three dimensions

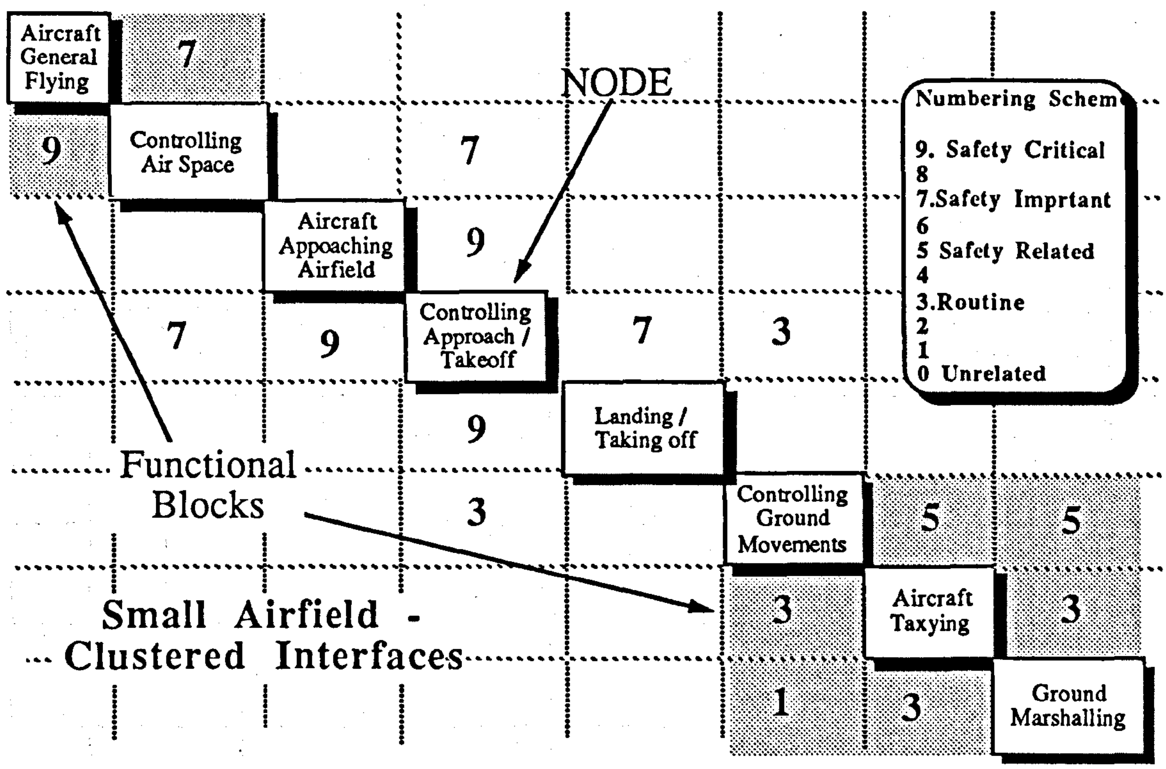
Find design weaknesses - incomplete clusters, system nodes, etc

Strengthen architecture - add links, replicate or protect nodes, etc.

Some examples of the procedure first for a simplified Airfield Project, and then for the Channel Tunnel follow; first, the Airfield Project:-



The N² Chart, or incidence matrix, is shown, in simplified form above, as it might appear after simply observing activities. The domain activities are represented in the leading diagonals, and the relationships between them gave been entered in the appropriate interface rectangles using the numbering scheme in the inset box. At this point, unclustered, the chart has been used simply to record, but already we can see a functionally bound block of activities at the bottom right hand corner, while "Controlling approach and Takeoff" has some of the features of a partial system node. We may rearrange the activities along the leading diagonal in detect further interface patterns, as follows:-



Re-arranging reveals a cleaner node and another bound pair at the top left-hand corner. The node also couples the two functional blocks. This example is simple but highly illustrative; the activities have grouped in domain patterns with "free" flying, controlled flying near the airfield and the movement of aircraft on the ground as the three domains

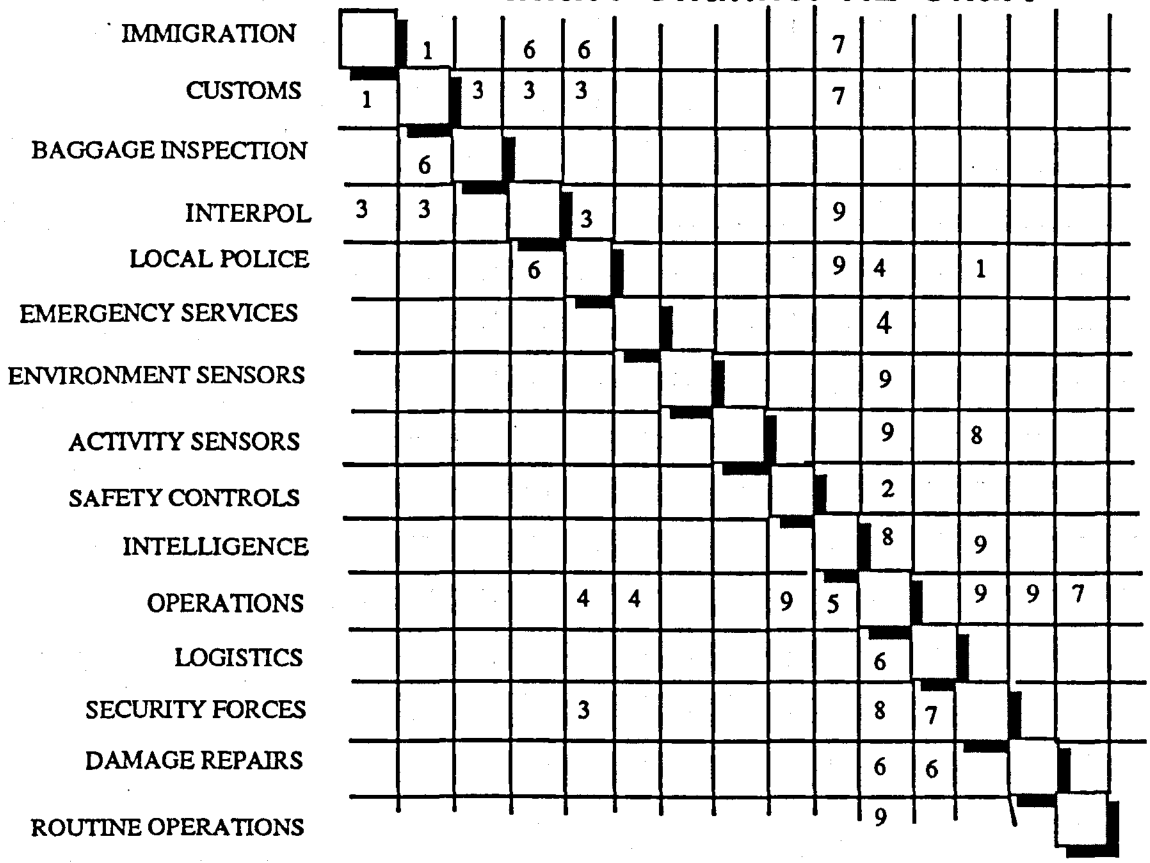
The number of activities above was so small that the use of computer assistance was unnecessary; following examples all made use of processing power, CADRAT, to untangle the patterns. First, the Channel Tunnel Crisis Management System:-

INTERFACES AND INFORMATION

IMMIG	LIAISON		FINDS	FINDS					FINDS					
LIAISON	CUST	ALERTS	FINDS	FINDS					FINDS					
		BAG INSP												
ALERTS	ALERTS		INTER POL	ALERTS					THREATS ALERTS					
			FINDS	LOC POL					THREATS ALERTS	STATUS RESRCS		LIAISON		
					EMGC SVCS					STATUS RESRCS				
						ENVT SENS				STATUS EVENTS				STATUS EVENTS
							ACTV SENS		ACTIVI-TIES	ACTIVI-TIES STATUS		ACTIVI-TIES		ACTIVI-TIES
								SFTY CNTR		STATUS				
									INTEL	ASSMNTS CNSTRS		WARNGS		
				ASSISTNCE PROGRESS	PROGRESS ASSISTNCE			ACTIVATE	INT GAPS PLANS	OPS		ACTIVATE DIRECT	ALERT DIRECT	ALERT STATUS
										RESRCS TRNSPRT	LOGS			
				LIAISON						STATUS FINDS RESULTS	RESRCS	SCTY FRCS		
										STATUS PROGRESS	RESRCS		DAMG RPRS	
								ACTIVATE		STATUS ALERT				RTNE OPS

The process is started by collecting and collating all the information flowing, or likely to flow, between the various entities concerned with the Management of the Channel Tunnel. The above N^2 chart is a typical example of such a chart at the early stages of design concept formulation for a simple system. Matrices with over 100 entities are quite common for more complex systems. (The abbreviations are, hopefully, self-evident, but in any event it is unimportant to read the detail; it is the overall pattern formed by the interfaces that is more important at this stage.)

Initial Chunnel N2 Chart



The second chart, above, introduces a marking scheme which rates each of the interfaces on a scale of zero (no interface) to 9 (strong relationship) as seen from the viewpoint of the proposed SOI. In the example, the viewpoint being expressed is that of the putative Crisis Management System, as can be seen by the relatively low numbers being allocated to interfaces not directed into the SOI, represented here by:

Intelligence
Operations
Logistics
Damage Repairs (the SOI-associated action element)

Some pattern is already visible in the Chart above; Operations has a number of interfaces and appears to be something of a local node. Perhaps CADRAT can reveal more pattern.

Clustered Chunnel N2 Chart

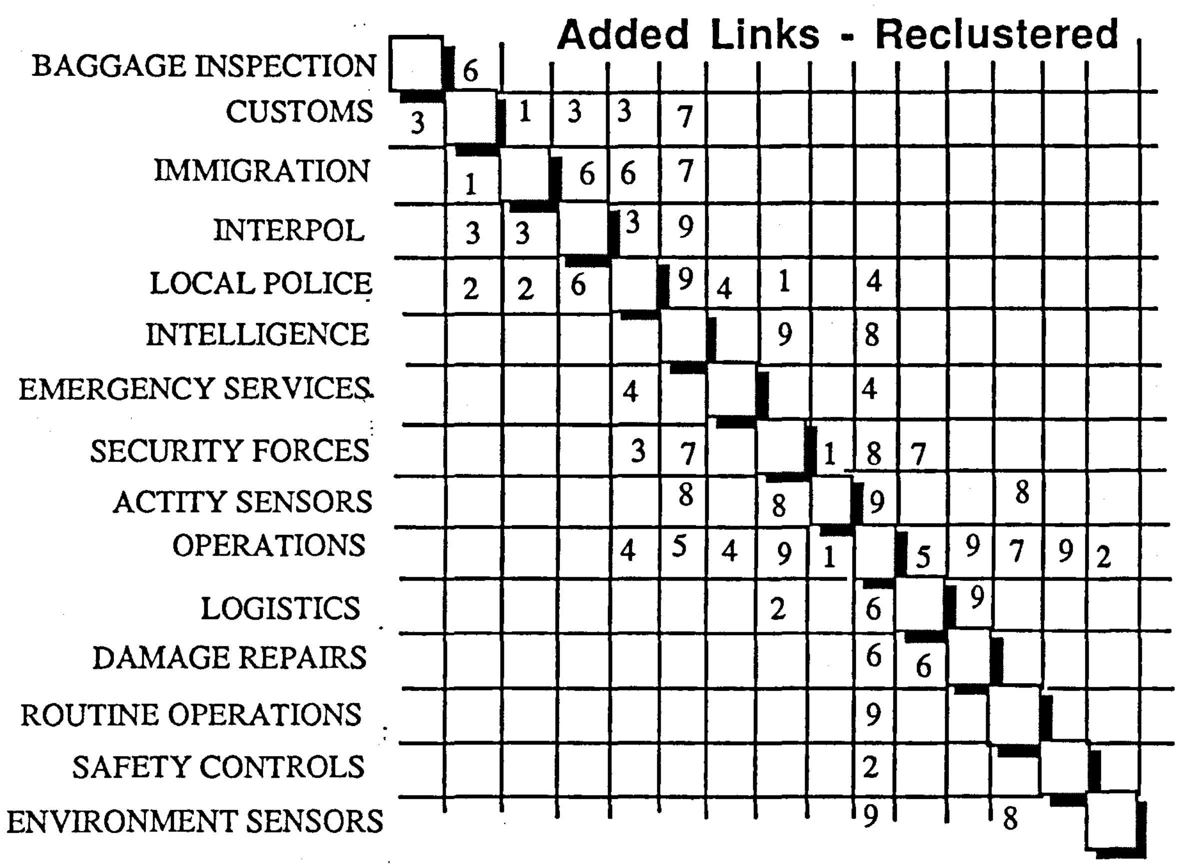
BAGGAGE INSPECTION		6													
CUSTOMS	3		1	3	3	7									
IMMIGRATION		1		6	6	7									
INTERPOL		3	3		3	9									
LOCAL POLICE				6		9			1	4					
INTELLIGENCE									9	8					
EMERGENCY SERVICES										4					
ACTIVITY SENSORS						8				9					
SECURITY FORCES					3	7				8					7
OPERATIONS					4	5	4		9		7	9		9	
ROUTINE OPERATIONS										9		9			
SAFETY CONTROLS										2					
ENVIRONMENT SENSORS										9	8				
DAMAGE REPAIRS										6					6
LOGISTICS										6					

The clustered chart reveals some pattern at the upper left of the screen brought about by clustering Customs, Immigration, Interpol and the Local Police, but the expected square associated with a functionally-bound block is clearly incomplete. It is also clear from the grouping that the emergency services, although connected into Operations, has no connection with the Local Police - a clear oversight in the original chart, revealed by the clustering process. The node centred on Operations is incomplete in that it has interfaces missing in the limbs of the cross and it is asymmetric, with the left limb longer than the right Lastly, there is a partially bound functional block joining Action Sensors, Security Forces and Operations.

All of these deficiencies are not only capable of repair, but are architectural weaknesses and should be repaired to instil necessary balance and structure in the architecture. Balance can now be seen, literally, as symmetry of pattern within the N2 Chart. The following chart repairs the omissions:-

BAGGAGE INSPECTION		6													
CUSTOMS	3		1	3	3	7									
IMMIGRATION		1		6	6	7									
INTERPOL		3	3		3	9									
LOCAL POLICE		*2*	*2*	6		9	*4*		1	4					
INTELLIGENCE									9	8					
EMERGENCY SERVICES					*4*					4					
ACTIVITY SENSORS						8			8	9	8				
SECURITY FORCES					3	7		*1*		8					7
OPERATIONS					4	5	4	*1*	9		7	9	*2*	9	*5*
ROUTINE OPERATIONS										9		9			
SAFETY CONTROLS										2					
ENVIRONMENT SENSORS										9	8				
DAMAGE REPAIRS										6					6
LOGISTICS									*2*	6				*2*	

The added interfaces have been italicized to highlight their positions. Of course, now the clustering needs to be re-executed, since the new interfaces result in a non-minimal matrix score, or simply, the entities can be clustered rather better. The following chart re-clusters the new set of interfaces:-



Enhancing the architecture by adding links has significantly changed the optimum cluster pattern. In particular, a new functional block has appeared uniting Operations, Logistics and Damage Repairs. Common sense might have suggested that these three would be closely associated; happily, proper application of the CADRAT tool makes up for any deficiencies in that area.

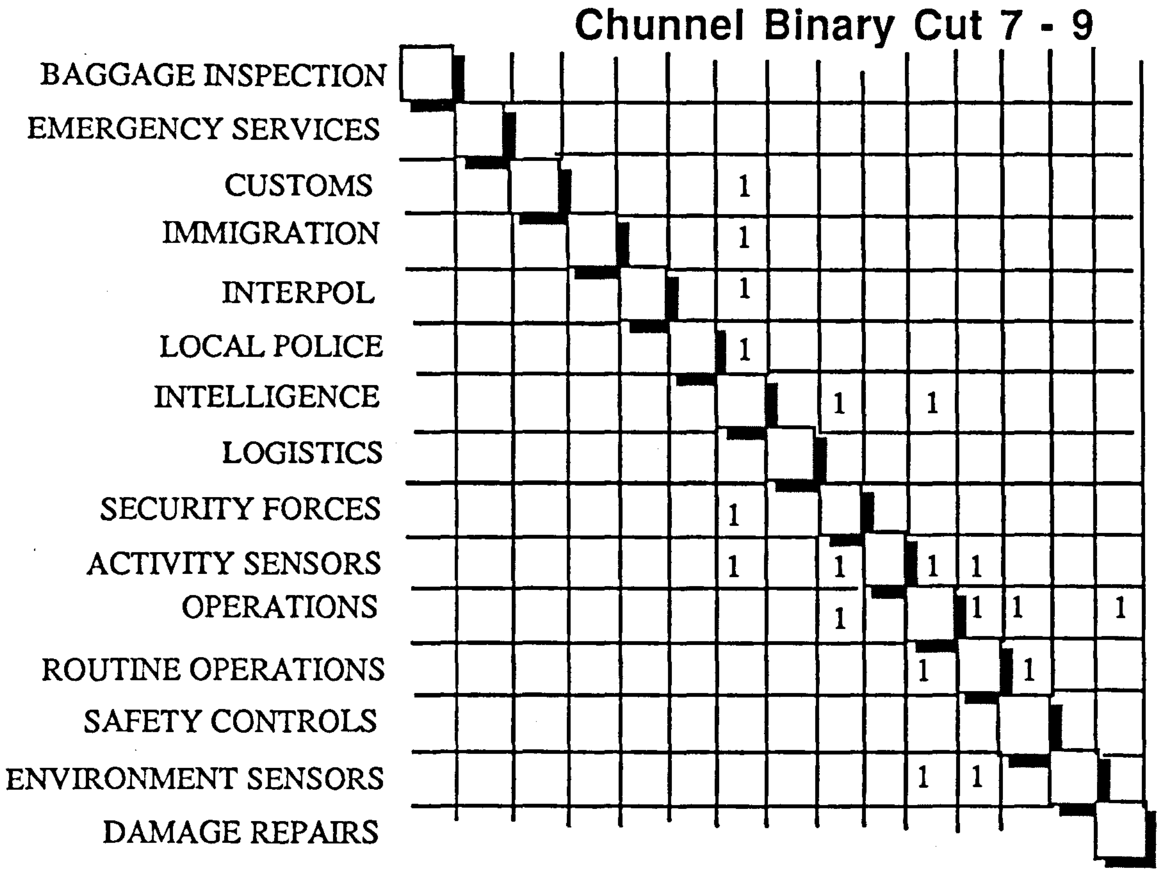
It was mentioned earlier that CADRAT could section the decimal numbers in order to examine the architectural pyramid layer by layer. This process can be very revealing, and can reveal weaknesses and enable remedial action to be proposed and tried. The tool simply takes any band of decimal numbers in the interfaces and reduces numbers in the selected band to unity, putting all others to zero. The chart can be clustered in this state to reveal the structure connecting peer activities. The following two Chunnel charts select two bands, 1 to 3 which, because of the numbering scheme chosen, looks at connections designated low priority as see from the Crisis Managers viewpoint, and 7 to 9 which conversely looks at the high priority links. The viewpoint caution is important; priority is very much in the eye of the beholder, and in this case the beholder's prejudice is revealed by the CADRAT clusters

CHUNNEL BINARY CUT, 1 - 3

BAGGAGE INSPECTION															
IMMIGRATION			1												
CUSTOMS	1	1		1	1										
INTERPOL		1	1		1										
LOCAL POLICE		1	1			1									
SECURITY FORCES					1				1						
LOGISTICS						1		1							
DAMAGE REPAIRS															
ACTIVITY SENSORS															
OPERATIONS									1			1			
SAFETY CONTROLS										1					
ENVIRONMENT SENSORS															
ROUTINE OPERATIONS															
EMERGENCY SERVICES															
INTELLIGENCE															

The chart reveals that, while Immigration, Customs, Interpol, Local Police and the Security Forces are all well connected at lower level, their connections with Crisis Management Intelligence is nil and with Operations is sparse, and via the Security Forces only. This is architecturally unsound because the Crisis Management Team must have low-level communication links with their essential allies in incident-containment. high-level links are well enough, but as shown at present, the Police would be hard pushed even to alert the crisis management team, and alerting is a priority communication.

The next figure shows the supposedly high priority links:-



This chart for high priority links shows Intelligence as rather more of a sink node than a source node, as evidenced by the asymmetric nodal cross about the activity. For example, Intelligence returns no information to the police, Interpol or any of the other agencies. This architecture may have clustered well under the full decimal regime, but the 3-D analysis reveals serious shortcomings which could be remedied quite simply. It is the ability of the clustering tool to reveal distinctive patterns with direct structural correspondence that gives such strength.

After repairing architectural weaknesses the process of architectural development proceeds along a series of simple steps:-

The functionally bound activity blocks are candidates for physical juxta-position in the design concept so as to shorten and facilitate links and interfaces for the overall system.

Such functional blocks can be collapsed to single functions in subsequent versions of the chart, which consequently becomes smaller as the architectural elements coalesce or aggregate, so gradually reducing the complexity. Collapsed functional blocks can, of course, be reclustered to continue the architectural build-up.

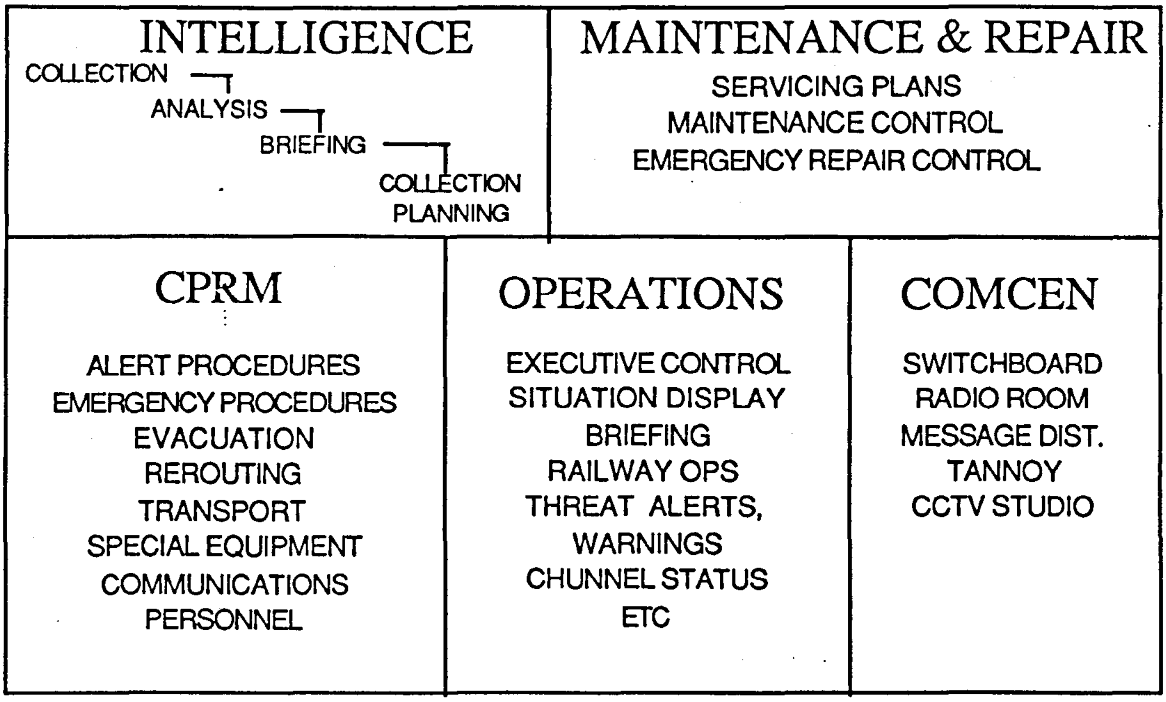
Where a member of a functional block cannot be physically located with the other members, as in the example where it alone is mobile, then the communication links with other block members should receive special attention, since they become extended and hence crucial to good operation.

Nodes also demand special attention. In a hostile environment, nodes must be either hardened or replicated or both. In a benign environment, such measures are unaffordable and unnecessary, but it is as well to support the node (see section on CADRAT) with alternative interfaces around its centre-point. The node in the final decimal cluster, three figures above, is partially-supported.

Some examples of the successive activities in developing design concepts for the Channel Tunnel follow:-

CRISIS MANAGEMENT FUNCTIONS

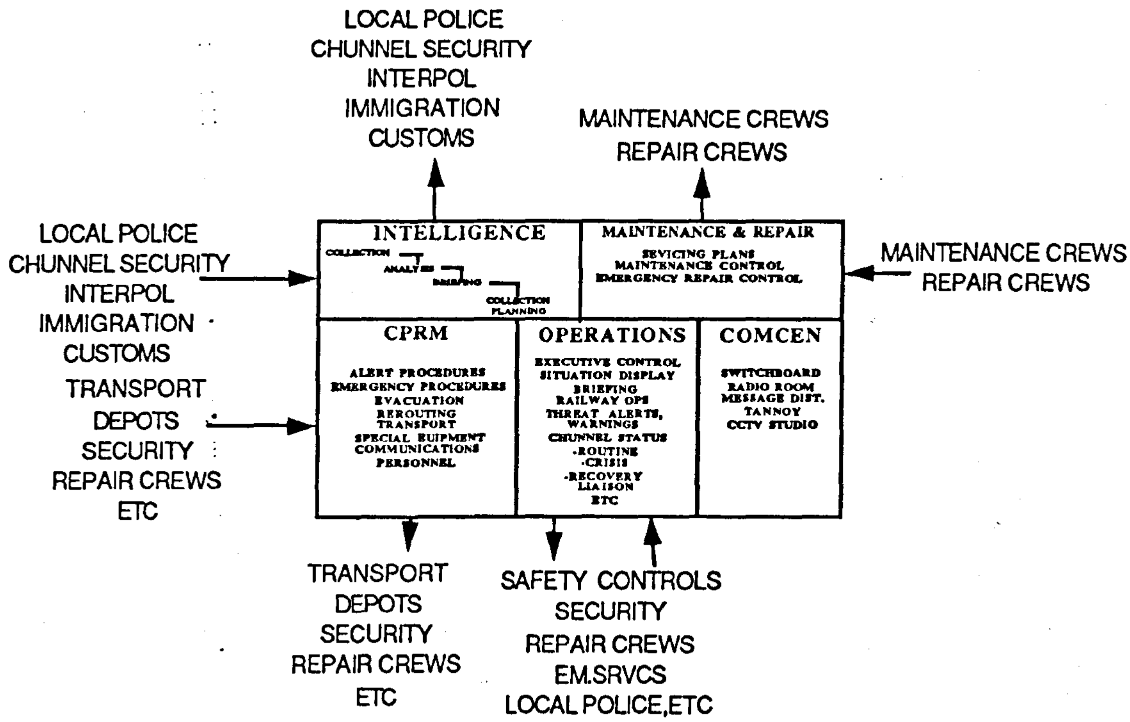
First the decision circle. Set against the circle are the principal factors of interest in the formulation of decisions. The various set of factors fall naturally to different skill-groups of people; Review Constraints for example, is often the realm of engineers and logisticians, while Assess Situation might be more appropriate to operations staff



CELL ORGANISATION

The decision circle forms the basis for physical organization in larger organizations where cooperative decision analysis is necessary. The figure above shows a notional layout for the Chunnel Crisis Management Control Centre layout. CPRM stands for Contingency Planning and Resource Management, two functions which are often combined

Having postulated a control centre and cell structure, it must be connected to the other system agencies in the manner developed in the N^2 charts above:-

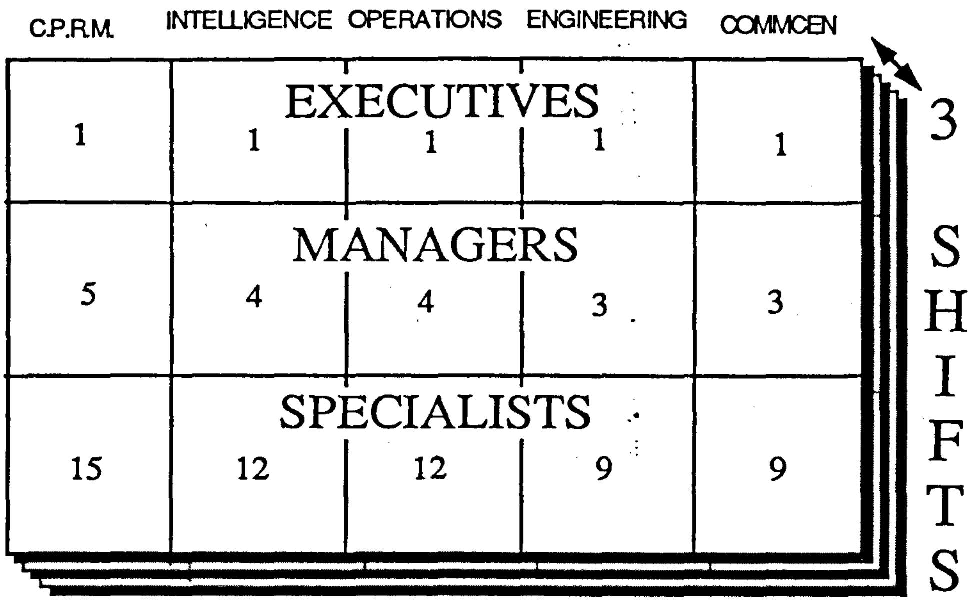


CRISIS MANAGEMENT CONNECTIVITY

The figure repeats the previous diagram in miniature and shows its in / out connections

Having looked at logical external connectivity, the concept can now develop in terms of communications capacities. This soon becomes an engineering task, but to begin it is necessary to establish the number of people operating within the Crisis Management Control Centre. Previous activities have developed the processes to be executed by some undefined combination of man and machine. It is now necessary to estimate the amount of information to be handled and processed, and to invoke the use of machines where useful or necessary, particularly to relieve operators of routine, but important, tasks and to develop a cost-effective design concept.

It is also necessary to consider the human activity structure within the Control Centre. Such centres can be viewed as three-tiered - see *Systems in Command*, enclosed - and there is also likely to be a need for shift working to provide continuous cover. It is useful to develop a human architecture viewpoint, as follows:-



MANAGEMENT ARCHITECTURES

The numbers in the squares show the estimated staff numbers each shift, before allowances are made for leave, sickness, etc, and considering only one Control Centre; survivability considerations might suggest a standby control centre with its own operatives.

Clearly, Crisis Management is highly labour intensive during a crisis. The concept of having such large numbers of people on duty twenty four hours a day, 365 days per year to anticipate a crisis is, however, unrealistic

Summary - Channel Tunnel Design Concept

It is interesting to note that the full analysis for Channel Tunnel Crisis Management does not favour a separate crisis management control centre on grounds both of cost and of inability to maintain the drive of the people involved over protracted periods when, hopefully, no crises occur. Instead, as the last N^2 chart above amply illustrates, there is such a close liaison needed between crisis management and routine operations management that the better approach is to combine the two. By providing dedicated crisis management facilities in the operations room, and maintaining a cadre of highly trained individuals within the Chunnel work-force and within the emergency services, the same objectives can be met more effectively and economically. There is a risk that such a cadre will not be maintained, and so its formation and training would best be mandatory, externally inspected and comprehensively undertaken, with exercises and command team training.

Advanced Architecture Design Concepts

This section touches on design concepts which are expanded and explained much more fully in *Advanced Avionic Architectures*, enclosed, which was presented at the Royal Aeronautical Society in March 1989. The concepts expressed in the paper have contributed to a major UK initiative in the realm of future aircraft design, presently the subject of much international activity, particularly in the US.

Since the concept of clustering and linking is fundamental to architecture, it should be applicable to the most advanced of projects. In such cases, the challenge and demand might be much greater, but the concept should be similar. Following examples are taken from an advanced avionics architecture project concerned with physical relationships and positioning of avionics devices - boxes - inside the fuselage. There are many influences on positioning:-

Antennae must be at the fuselage skin

Receivers must be near the antennae to avoid unnecessary signal loss in the connecting cable

Displays and controls must be near the crew

Critical devices must be replicated and isolated to avoid a single hit from damaging performance

Centre of Gravity must be maintained within limits, etc.

As can be seen, while some factors tend to cluster, others tend to repel. The CADRAT tool can accommodate this by a simple change of numbering scheme, as follows:

9.Proximity Essential
Proximity v. Important
7.Proximity Important
Proximity Useful
5.Proximity Convenient
4.Don't Care
Distance Convenient
2.Distance Useful
1. Distance Important

Effectively all that is happened is a shift of origin to the middle of the range;

this does, of course, exclude zero as an option and in consequence every interface is filled, making clustering by eye and simple matrix manipulation virtually impossible. To the CADRAT tool, nothing has changed, however.

With the more complex situation presented by avionics architectures, the usefulness of the First and Second Moment Options in the CADRAT program are realized. It also becomes swiftly evident that we are not in the realm of "right answers". There is no correct answer to architectural problems; some are simply better than others in some respects.

Consider the following situation. An avionics system is comprised of a number of discrete modules which can be put together to build operating functions. There are 16 such modules in this case; they are as follows:-

A power amplifier (PA)
Two antennae
A special transmitter
An special receiver
A radar transmitter (Tx)
A Communications, Navigation and Identification (CNI) transmitter
A CNI receiver
A radar receiver (Rx)
Two TX/RX Controllers
A frequency synthesizer
Crew Displays and Controls (D & C)
Weapons

Adding the aircraft fuselage and the crew themselves completes the picture. The first of the resulting clustered N^2 charts is as follows:-

Power Amplifier		9	9	9	3	9	5	9	2	2	1	2	2	5	2	2
Special Tx	9		6	6	7	5	5	5	4	4	5	4	7	5	5	5
Radar Tx	9	6		6	7	5	5	5	4	4	7	4	7	5	5	5
CNI Tx	9	6	6		7	5	5	5	4	4	7	4	7	5	5	5
Controller 1	3	7	7	7		5	5	5	7	7	7	7	1	5	5	5
Antenna 1	9	5	5	5	5		9	1	9	9	5	9	5	5	1	1
Fuselage	5	5	5	5	5	9			5	5	5	5	5	5	5	9
Antenna 2	9	5	5	5	5	1	9		9	9	5	9	5	5	1	1
Special Rx	2	4	4	4	7	9	5	9		6	8	6	7	5	5	5
Radar Rx	2	4	4	4	7	9	5	9	6		8	6	7	5	5	5
Freqency Synth.	1	5	7	7	5	5	5	5	8	8		8	5	5	5	5
CNI Rx	2	4	4	4	7	9	5	9	6	6	8		7	5	5	5
Controller 2	3	7	7	7	1	5	5	5	7	7	7	7		5	5	5
D & C	5	5	5	5	5	5	5	5	5	5	5	5	5		9	5
Crew	2	5	5	5	5	1	5	1	5	5	5	5	5	5		2
Weapons	2	5	5	5	5	1	9	1	5	5	5	5	5	5	2	

ALL TRANSMITTERS PLUS CONTROL

PIVOT

ALL RECEIVERS PLUS CONTROL AND FREQUENCY SYNTHESISER

ISOLATED FROM RADIATION

Advanced Architecture - First Moment

The first moment cluster shows several interesting architectural features:-

> The Transmitters have been grouped with one antenna and one controller "above" the Fuselage
>
> The Receivers have also been grouped with the other antenna and the other controller"below" the Fuselage
>
> The Fuselage is a pivot, around which the design clusters
>
> The Displays and Controls, the Crew and the Weapons have been positioned remotely from the Transmitters with their unwelcome and potentially hazardous radiation
>
> The Controllers have been mutually separated for survivability.

The first moment clustering is interesting, but that may not be the end of the story, since first-moment clustering produces only a few, fairly loose clusters. The second moment cluster produces tighter clusters as follows:-

Power Amplifier		9	9	9	9	9	5	3	2	2	2	3	1	5	2	2
Antenna 1	9		1	5	5	5	9	5	9	9	9	5	5	5	1	1
Antenna 2	9	1		5	5	5	9	5	9	9	9	5	5	5	1	1
Special Tx	9	5	5		6	6	5	7	4	4	4	7	5	5	5	5
Radar Tx	9	5	5	6		6	5	7	4	4	4	7	7	5	5	5
CNI Tx	9	5	5	6	6		5	7	4	4	4	7	7	5	5	5
Fuselage	5	9	9	5	5	5		5	5	5	5	5	5	5	9	5
Controller 1	3	5	7	7	7	5			7	7	7	1	7	5	5	5
Special Rx	2	9	9	4	4	4	5	7		6	6	7	8	5	5	5
Radar Rx	2	9	9	4	4	4	5	7	6		6	7	8	5	5	5
CNI Rx	2	9	9	4	4	4	5	7	6	6		7	8	5	5	5
Controller	3	5	5	7	7	7	5	1	7	7	7		7	5	5	5
Frequency Synth.	1	5	5	5	7	7	5	5	8	8	8	5		5	5	5
D & C	5	5	5	5	5	5	5	5	5	5	5	5	5		5	9
Weapons	2	1	1	5	5	5	5	5	5	5	5	5	5	5		2
Crew	2	1	1	5	5	5	5	5	5	5	5	5	5	9	2	

ALL RADIATORS AND ANTENNAE AT FUSELAGE SURFACE

PIVOT

ALL RX

CONTROLLERS SEPARATED FOR SURVIVAL

COMMON TX / RX

ISOLATED FROM RADIATON

Advanced Architecture - Second Moment

The tighter clustering produces the following features:-

> The Fuselage forms a design pivot point as before
>
> All the Transmitters and their Antennae are grouped together "above" the Fuselage
>
> All the Receivers are grouped immediately "below" the Fuselage, some distance from their respective Antennae.
>
> The Frequency Synthesizer and one of the Controllers have grouped as common Tx / Rx facilities and indeed they serve both Transmitters and Receivers equally.
>
> As before, the Displays and Controls, the Weapons and the Crew have been isolated from the radiation, but rather more effectively in this case, by virtue of the way in which the radiators have been forced to the top. Note also that the Crew and Weapons have exchanged positions, the Crew now being furthest from the radiation.

The two diagrams show differences that are more than subtle architectural variations, and, since there is no "right" answer, they present alternative clustering concepts. Thus the CADRAT tool is, at this level, an idea generator, a concept proposer, with the operator interpreting, choosing, modifying and perhaps rejecting as he or she proceeds.

9.3. Summary

This chapter has illustrated some of the more unusual aspects of design concept evolution. Several linked techniques have been employed which are valuable, but the Seven Step Continuum is not based on any particular techniques or tools. It is a framework for developing IDA system design concepts from issues, and any methods which achieve tasks within the framework are valuable.

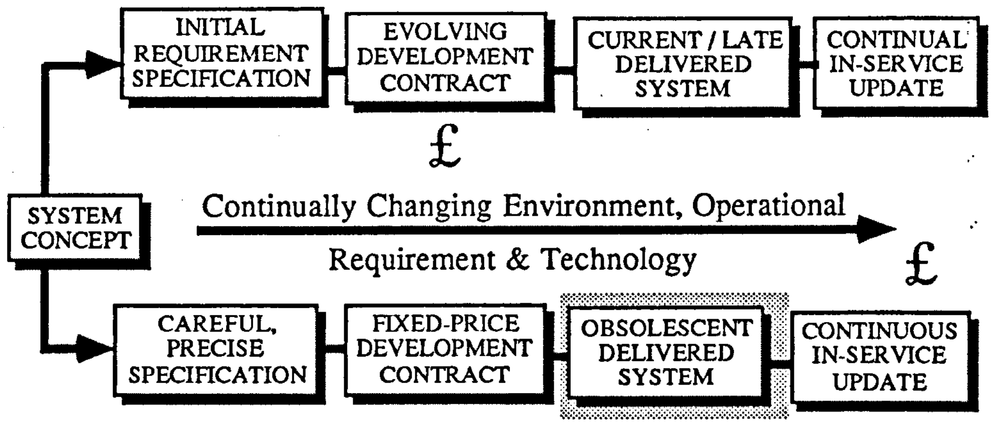
The SSC has been formed by action research, and it is a living framework in the sense that its design is not finalized and fixed. It has also been tested on only three subjects, admittedly all different. To form a judgement on its efficacy would be premature. It is, however, reasonable to suggest that it shows promise.

10. BUILDING FUTURE SYSTEMS

The enclosed paper, *Building Future Systems,* is an invited paper at the 1989 MILCOMP Exhibition and Symposium, and is concerned with the present tendency to design and implement large-scale IDA systems which are obsolescent at delivery.

10.1. Cost Effective Systems

The process of procuring large-scale IDA systems is presently going through a series of changes, brought about principally by past overruns on time and budget. The current mode is for fixed-price contracts, instead of the previous "cost-plus" regime. As the following figure, extracted from the enclosed paper *Building Future Systems,* shows, this change may not bring quite the expected benefits



CAREFULLY - SPECIFIED OBSOLESCENCE

The moving "£" indicates that the cost must be borne, by one customer or another. The dilemma facing customer and contractor is severe. If costs are to be held in check, the argument goes, then contracts must be based on a very careful specification which will result in precisely that which is required. Trying to carefully specify something as complex as an IDA system is, however, not practical for three reasons:-

> The Heisenberg syndrome - the more detail is required in the description of any human activity, the more difficult that detail is to convey. An example is the inability to specify in detail the dynamics of a real-time situation display. Even if the customer has a well-formed understanding of the future requirement, the words do not exist to meet the need for detailed description

Supposing perfect definition were feasible, the resulting implementation would produce a time capsule; the external environment within which the system has to operate evolves during the period of development, making the newly developed system obsolescent at delivery.

The more detailed the specification, the longer it takes and the more expensive it becomes, so making itself more likely to be outdated and unaffordable. This process is visible at present in the use of knowledge elicitation at the start of some major projects, in an attempt to tease out the softer issues at the start of the specification process.

Chasing the illusion of the perfect specification clearly presents diminishing returns. Recognition of the problem has resulted in the introduction of prototyping as a means of resolution. Prototypes of IDA systems represent some aspects of the eventual system in reasonable detail while other parts are modeled in coarse grain. The display example given above is typical; a display representing the final display is produced, with the remainder of the system which contributes information and dynamics to the display being simulated relatively crudely. A group of potential operators can sit in front of the display, can judge its potential and can modify the presentation to their consensus view.

The prototype concept is a step forward; it involves the operators and users at an early stage and enables them to apply their domain understanding directly, without the vicarious interpolation of a knowledge elicitor. But even here there are limitations: consensus means compromise, and some operator needs may be sacrificed; the domain environment will still evolve during development; and of course, the operators were not facing the real situation in the real environment, so such very real emotions as fear were not brought into play. Fear can make the simplest of actions difficult to effect.

So what is the answer?

10.2. Evolving Systems

Nature does not approach the problem in the same way: Instead, higher-order animals are born rather helpless, with limited imprint, but with the ability to evolve socially. Could IDA systems be designed similarly?

It seems likely that they can.There are two approaches which hold promise, as follows:-

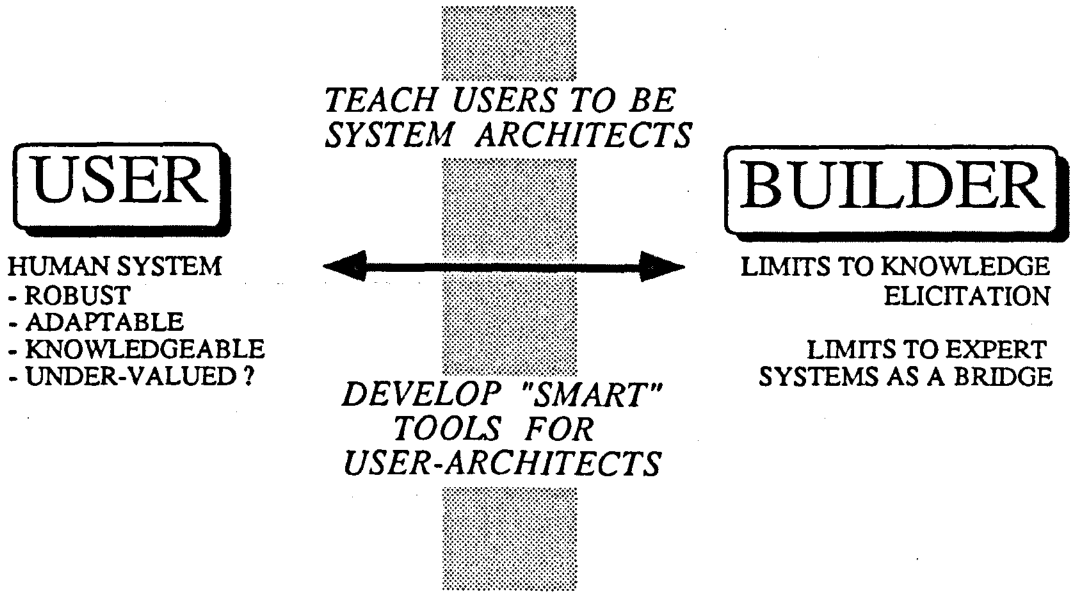
Deliver systems into service that contain their own building

capability. For example, instead of designing systems with a set of fixed display formats, design systems with a facility for the users to format their own displays. On delivery, the desired set of formats is already available, but as situations develop, as users find new applications, or as environments evolve, the users can themselves evolve their displays in step. This approach, designing user-adaptable systems, is technologically feasible now, and could be much more widely applied to advantage

Design Systems which can self evolve. Using the approaches outlined in the papers and the section concerned with *Cumulative Selection*, it is within technological capability at present to design systems which, for example, layout their own communications networks in response to the contemporary environment. If the environment changes, either because the IDA system is mobile or due to some hostile action, the IDA system can be designed *in total* can be designed to evolve towards the optimum dictated by the new situation; at present, only the human part responds effectively, rendering the man-made parts ineffective for the new purpose and environment.

10.3. User-Architects

Reference has already been made to the limitations implicit in knowledge elicitation. The alternative is to engage the users in system-building. They alone possess the domain knowledge. They lack, however, the training and skills which engineers and systems designers acquire. There is a need to bridge the gap between these two groups, and to develop a new breed - the user-architect - as illustrated below:-



BRIDGING THE USER - BUILDER DIVIDE

It is an objective of the Seven Step Continuum, as an instantiation of a conceiving system, to bridge that gap. The tools, SYSMORPH, COMMORPH and CADRAT described above and in the enclosed papers are also intended as supports for the bridge, since they require principally domain knowledge and understanding in their use and very little skill in operation. The information which the programs access may, however, contain much detailed engineering knowledge and understanding, and the outputs are certainly meaningful to designers and engineers.

Here, then, are the ways forward in the building of future systems: flexible, evolvable systems, conceived by users and designers together, supported by a framework of a conceiving system embedded with engineer-developed, design concept tools, driven by user-architects.

11. SUMMARY AND CONCLUSIONS

The aims set out for the thesis, at the end of the first chapter were as follows:-

> **To review the present Systems Engineering Scene.** It is a conclusion from the thesis that systems engineering is alive and, if not well, is at least showing signs of life in its embracing of soft issues. See Chapter 2, 3 and 4
>
> **To identify shortcomings in that Scene.** Shortcomings abound, not least in the lack of consensus over the substance of systems engineering. "Off-the-shelf" syndrome rules in many spheres The advent of software engineering tools, a bonus in their intended sphere of activities, threaten creativity and innovation if misused at the higher / earlier levels of systems analysis. See Chapters 1 and 2
>
> **To examine the contribution that softer methods might make to systems engineering.** See Chapter 2. The brief overview of SSM and parallel activities suggested that many of them were subjective, relying strongly on the skill of the practitioner to develop the issues in the problem situation. While this characteristic may be inherent in the understanding of soft issues, it leaves the less experienced analyst without guidance or direction. SSM presents its practitioners with a framework that is so loose that, while it might be all-encompassing, it is not always entirely helpful. There are, moreover, no associated methods, formal or informal, of handling the complexity of problem situation generated by the method. The elaboration of Weltanshauungen, seductive in their expansion of understanding, is not matched by a process of reconciliation between the various viewpoints,

which leaves the problem of choosing rationally between them unaddressed. SSM seems, judging by the nature of the method and by the examples presented by the many advocates, to be best applied to problem situations associated with existing systems, as opposed to the creation of new systems *ab initio* SSM has, none-the-less, a considerable potential its ability to unravel extant problems. Other soft methods, notably those of Eden and Janes, while still subjective, do offer means of handling complexity and in the process may identify previously-hidden emergent properties of the SOI; these methods offer more promise for the less experience analyst.

To develop the essentially human and anthropomorphic nature of human-designed IDA systems - see Chapter 4, in which it is contended that homo sapiens sapiens is not as unpredictable as suggested by many scientists and engineers, soft and hard, but that territorial imperative, dominance and submission, and many other reasonably-understood characteristics of human behaviour, make the human animal rather more predictable than less. Lessons in individual and group behaviour are there to be learnt for IDA system designs from social anthropology, psychology, even ethology. Lessons are there to be learnt in the superb compromise that is the human body. Some of those lessons are presented in the body of the thesis, more in the enclosed papers

To introduce the concept of a conceiving system as a bridge between enquiring systems and creating, or design & development, systems - see Chapters 5, 6 and 8.

To present and justify a conceiving system for the conception of Information-Decision-Action systems - see Chapters 7, 8 and 9

To explore the the role of the user / operator in Conceiving Systems - see Chapter 10

To achieve these objectives, in part, by presenting work undertaken in the last ten years and in so-doing to present a view of the confluence of Hard and Soft Schools - see below

The methods presented in previous chapters for the conception and development of system architectures are essentially soft methods in that they address issues and seek solutions which are imprecise. That they are more pragmatic than the methods, or lack of methods, prescribed by such as SSM is both true and should be no surprise, since industry needs results. The Seven Step Continuum, shown in action in previous chapters, is unusual in that, although it seeks a solution, it can show that no sensible solution exists along the path pursued by the practitioner, as with the Channel Tunnel Crisis Management System. Methods, such as the N^2 / CADRAT approach, which introduce balance and symmetry into abstractions of the system cannot be described as "hard", but they are pragmatic. The decision circle (and the many similar paradigms to be found in IDA system design) evident in the body of the thesis and in many of the enclosed papers, is not a "hard" concept, nor are notions of territorial imperative, tribal loyalties, social psychology, ethology and the many other viewpoints being expressed by today's leading-edge designers

It is my contention that the terms "hard" and "soft" are misleading and outdated. The views expressed by SSM devotees concerning hard engineering are passe caricatures of design practice. Equally, the hard engineer recognizes that human issues are at the root of design and, while he or she may not embrace the set of approaches generally called "soft", they generally have their own methods for addressing these issues which, because of the environment , need to be justifiable, traceable and suitable for team design, rather than individual design.

So, the so-called hard engineer uses soft methods and indeed the so-called soft analyst uses hard, reductionist methods too. Let us dispense with the polarization and use the best means at our disposal to meet the need.

REFERENCES

Ashby, W. R., 1985, "Introduction to Cybernetics", Wiley

Alford, M., 1985,"SREM at the Age of Eight; the Distributed Computing Design System", Computer, Vol. 18, No 4, pp. 36-46

Bird, D. F., 1987, "International Standards in Military Communication", from "Advances in Command, Control & Communications Systems" published in book form by Peter Peregrinus Ltd on behalf of the Institution of Electrical Engineers, 1987: ISBN 0 86341 094 4,

Boulding, K.E., 1982, "Review of Systems Thinking, Systems Practice", Journal of Applied Systems Analysis, 9

Bronowski, J., 1973, "The Ascent of Man",BBC Publications, London, ISBN 0 563 10498 8

Burrell, G., "Systems Thinking, Systems Practice: A Review", 1983, Journal of Applied Systems Analysis, Volume 10, pp 121-126

Burrell, G. and Morgan, G., 1979, "Sociological Paradigms and Organizational Analysis", London, Heineman

Coe, G. Q., and Dockery, J. T., "OJCS Initiatives in C2 Analysis and Simulation",from Science of Command and Control,AFCEA International Press Publication, ISBN0-916159-16-7

Checkland, P.B., 1972, "Towards a Systems-Based Method for Real World Problem Solving", Systems Behaviour, Harper and Row

Checkland, P B., 1981, "Systems Thinking, Systems Practice", John Wiley

Checkland, P B., 1971,"A Systems Map of the Universe", Journal of Systems Engineering, 1(2)

Churchman, C.W., Ackoff, R., and Arnoff, E., 1967, "Introduction to Operations Research", John Wiley and Sons Inc, New York

Coyle, R. G., 1977, "Management System Dynamics", Wiley

DeMarco, T., 1978, "Structured Analysis and System Specification", Yourdon de Broglie, L., 1951, "La Cybernetique, Theorie du signal et de l'Information", editions de la Revue d'Optique, Paris

de Neufville and Stafford,1971, "Systems Analysis for Engineers and Managers", McGraw-Hill Book Company, New York.

Engineer, The, 22nd June 1989, "Design looks to Natural Selection", article by Martin Ince

Financial Times, 18th August 1989, "Darwin's Theory Gives Birth to Genetic Computer Programs", article by Gren Manual

Forrester, J. W., 1961, "Industrial Dynamics", MIT Press

Gane, C. and Sarson, T., 1979, "Structured Systems Analysis: Tools and Techniques", Englewood Cliffs, N.J., Prentice Hall,

Guilbaud, G. T., 1959,English translation form "La Cybernetique", Presses Universitaires, Paris 1954

Hall, A. D., 1962, "A Methodology for Systems Engineering", Princeton N. J., Van Norstrand.

Horgan, J., 1989, "Ethnograph as Art" Scientific American, Volume 261 Number 1

Hughes, J, 1986,"Structure and Design of Manufacturing Systems", Open University PT611, Unit 8 - Manufacturing Systems: Detailed Design", ISBN 0 335 10922 5

Jackson, M.A., 1975, "Principles of Program Design", Academic Press

Lammers, G. H., "C3 Effectiveness Studies" from "Advances in Command, Control & Communications Systems" published in book form by Peter Peregrinus Ltd on behalf of the Institution of Electrical Engineers, 1987: ISBN 0 86341 094 4,

Lano, Robert J. A "Technique for Software and Systems Design", North Holland Publishing Co. 1979. ISBN 0-444-85354-5. TRW series on Software Technology, Volume 3

Lano, Robert J. "Operational Concept Formulation", TRW series on Software Technology, unnumbered

Levis H. J., and Athans, M., 1988, "The Quest for a C3 Theory; Dreams and

Realities,from Science of Command and Control,AFCEA International Press Publication, ISBN0-916159-16-7

Macgill, S. M., 1983, "Cluster Analysis and Q-Analysis", Int J. Man-Machine Studies, (1984), 20, 595-604

Manheim, M.L. and Hall, F.L., 1967,Abstract Representation of Goals, "Transportation: A Service", New York Academy of Sciences and American Society of Mechanical Engineers, New York

MAP, Version 2.1, 3 March 1985, General Motors Technical Center, Manufacturing Building, 30300 Mound Rd., Warren, Michigan 48090-9040, USA)

Mayk, I. and Rubin, I, 1988, "Paradigms for Understanding C3, Anyone?",from Science of Command and Control,AFCEA International Press Publication, ISBN0-916159-16-7

M'Pherson, P. K., 1980, "Systems Engineering: an Approach to Whole Systems Design", The Radio and Electronic Engineer, Vol 50 No 11/12

Popper, K.R., 1972, Objective Knowledge, Oxford: University Press

Roberts, N.,1983, "A Systems Dynamics Modelling Approach", Addison Wesley 0-201-06414-6.

Ross, D. T., 1977,"Structured Analysis(SA): A Language for Communicating Ideas":IEEE Transactions on Software Engineering, Vol SE 3, No1, pp 16-34

Ross, D. T., and Schoman, K.E., Jr, 1977, "Structured Analysis for Requirements Definition", IEEE Transactions on Software Engineering, Vol SE-3, No 1, pp 6-15

Rubenstein, Edward, "Stages of Evolution and their Messengers", Scientific American, June 1989 Volume 260 No 6

Shannon, Claude C, and Weaver, W., 1963, "The Mathematical Theory of Communication", Urbana: University of Illinois Press

Sugiyama, K., and Toda, M., 1983, "Visual Q-Analysis(1): a Case Study of future Computer Systems Development in Japan", Cybernetics and Systems, 14:185-228, 1983, Hemisphere Publishing Corporation

Warfield, J. N., 1974, "Towards Interpretation of Complex Structural

Models", IEEE Transactions in Systems Management, Cybernetics, SMC - 4(5): 405 -417

Wilson, B.,1984, "Systems; Concepts, Methodologies and Applications", John Wiley and Sons

Wohl, J. G., 1981, "Force Management Requirements for air force tactical command and control"

ENCLOSED PAPERS, REPORTS AND ARTICLES BY THE AUTHOR

"Air Command and Control", MARCONI 81 SYMPOSIUM Proceedings

"Avionics Systems Architecture", joint paper by Hitchins and D O'Dwyer, Proceedings of the Royal Aeronautical Society Conference on Digital Computer Technology for Airborne Applications, !5 March 1989, ISBN 0 903409 40 2 - Note: paper written in two parts; Hitchins wrote the second part

"BM/C3 Design by Cumulative Selection", Proceedings of the Third IEE International Conference on Command, Control, Communications and Management Information Systems: 2 - 4 May 1989, ISBN0852296380, ISSN 05387-9989

"Building Future Systems", accepted for MILCOMP 89 Conference in London, 26-28 September 89"

"Creative Design by Cumulative Selection", submitted to the IEE for publication in the Proceedings, 22 May 1989

"Design by Cumulative Selection", IEE M&D Digest No:1989/105, 26 April 1989

"General Theory of Command and Control, A", Proceedings of the Third IEE International Conference on Command, Control, Communications and Management Information Systems: 2 - 4 May 1989, ISBN0852296380, ISSN 05387-9989

"Getting SEAMS Straight", IEE M&D Digest No 1988/93, 17 May 1988

"Heuristic Intelligent Threat Assessment System (HITAS)", described in "Automatic Airborne Threat Assessment, an EASAMS Company Publication, 1981

"Human Element in C3I, The", Proceedings of the First IEE International Conference on Advances in Command, Control and Communication Systems: Theory and Applications: 16 - 18 April 1985, ISBN 085296308, ISSN 0537-9989

Initial IDA Engineering Framework, developed as a management tool in Racal Defence Systems Ltd

"Managing Systems Creation", 1986, IEE Proceedings, Vol 133, Pt A, No. 6, September 1986

"MOSAIC Concepts for the Deployment of Air Power in Europe", (Extended Version), part of "Advances in Command, Control & Communications Systems" published in book form by Peter Peregrinus Ltd on behalf of the Institution of Electrical Engineers, 1987: ISBN 0 86341 094 4, extended from "MOSAIC Concepts for the Deployment of Air Power in Europe", Proceedings of the IEE International Conference on Advances in Command, Control and Communication Systems: Theory and Applications: 16 - 18 April 1985, ISBN 085296308, ISSN 0537-9989

"System Survivability Science", EASAMS Ltd Company Publication No: M7101/DH/1, March 1989

"Systems Creativity", 1987, IEE Proceedings Vol. 135,Part A, No 6, July 1988

AIR COMMAND AND CONTROL

D K Hitchins
Manager
Forward-looking Studies
EASAMS Limited

As most people are aware, air power is normally split into four functional groups. Offensive operations are concerned with interdiction and strike, close air support, counter air and air superiority. Defensive operations are self explanatory: fighter interception, surface to air missiles, airborne early warning and the air defence ground environment (the radars). Maritime operations are usually comprised of long range patrol, anti-submarine warfare, tactical aid in support of maritime operations and air defence of the fleet. Air transport operations are usually concerned with tactical transport, reinforcement and resupply and mobile forces. Unfortunately, this rather neat grouping conceals a multitude of problems.

There are quite a large number of functions which are part of air power which cut across the apparently clean boundaries between the four major functions. Some of them are - reconnaissance, aircraft recovery, search and rescue, intelligence, contingency planning and resource management and so on.

You might therefore believe that the functional groupings are not very valuable: that would not be the case. The functional groupings are extremely valuable for many reasons. They concentrate like skills; they enhance motivation and competition; they afford a very simple organisation; they allow simple measures of effectiveness by function; and, not to be overlooked, they allow the separation of the various functions into defence and offence; d'etante favours expenditure on defence as opposed to offence. From which we conclude; firstly, that effective air power will retain functional boundaries in the future and, secondly, during any prolonged periods of peace, defensive operations will receive continual funding and may tend to dominate other elements.

May I turn aside briefly to mention a system design axiom and corollary and show how they relate to air power. The axiom states that a system comprised of elements which have been separately optimised will not itself be optimised and its corollary is that an optimised system will contain elements which are not of themselves optimal. The reason behind this is very simple. A group of elements, separately optimised, can be considered as producing an overall function labelled '4' in fig 1; this function No.4 has not been optimised: nor have the interfaces between the various elements of which it is comprised been considered in the context of the overall task, from which we conclude, in the context of air power, that improving the effectiveness of air power requires a total system approach rather than the function by function development, which is more usual.

If we are to improve the effectiveness of a complex system such as air power we need some measure of effectiveness so as to discern when it has been improved. This raises a problem for which I do not offer a ready solution. As fig 2 purports to show, the blue force, eg, has an overall effectiveness which is a function not only of its offensive and defensive capabilities but also of the orange's offensive and defensive capabilities. Moreover, blue defence and blue offence may not always help each other. They may tend to interfere with each other. The description of blue air power effectiveness is extremely complex. Please note that the diagram even at this level does not include all the supporting functions. Measures of effectiveness are extremely important and they are also, of course, extremely difficult to deduce and measure.

Contemporary thinking on air power suggests that a comprehensive command and control system is needed to overcome the limitations imposed by existing functional demarcation between offence and defence. We must define the system that we are talking about. The system that we are describing as requiring improved effectiveness is the one containing all the elements of air power within a major NATO command or a major subordinate command. I have used the NATO context for the whole of this presentation. The system embraces sensors, weapon system and all three arms; land, sea and air. It is essentially a 'sharp-end' system. The design aim is to optimise the effectiveness of air power as a total system by improving the organisation and flow of timely information between the various elements of that air power.

Fig 3 shows a general schematic diagram indicating the relationship between the various types of command and control systems. You will see at the top a major NATO commander, SACEUR in the case of Europe, and below him he has a major subordinate command structure. He has, in fact, got four major subordinate commands. They tend to communicate with each other, through a CCIS, a command and control information system which also embraces the command level below major subordinate command, called principal subordinate command. The structure tends to be horizontal in contrast to that required for an air command and control system which tends to be a vertical, pyramidical type of structure, grouping together offensive/ defensive maritime, transport and all the other elements to provide a vertical input to the otherwise hgiher level horizontal structure.

Fig 4 shows the air command and control system, or ACCS, capabilities and you will see on there such diverse requirements as orders of battle, contingency planning and resource management, nuclear, biological and chemical, emission control, electronic warfare, sensor fusion and so on.

Air command and control systems are going to be extremely complex. There are a number of widely differing parameters which will contribute towards the eventual architecture design, the more important ones - survivability and succession of command, communications, interoperability, security, life cycle costs, reliability, availability and maintainability (commonly known as the 'ilities'). The existing CCIS and weapon/sensor facilities will also have to be coordinated and brought together in order to provide the system. Not only are these various factors different in their various degrees of emphasis, they are also really not comparable on the same basis. By this I mean they are of different dimensions. Most of them are also non-numerical dimensions. There is a need for specialised analysis techniques if we are to successfully design the organisation and structure of air command and control architecture. I would like to go through one or two of the factors which I have mentioned.

Survivability is conventionally considered under three headings: avoidance of detection, self defence and damage tolerance. Avoidance of detection reallly means not making yourself very obvious, not radiating, having low thermal signatures or, conversely, placing yourself against a background which is very similar. This can mean camouflage; it can also mean, to use Mark Twain's analogy, hiding a book in the library. Thus there are some views that command and control centres should be in urban complexes. Self defence is not a parameter which is readily achieved by a command and control headquarters, which by its very nature is passive. Damage tolerance covers a multitude of factors - hardened structures, electromagnetic pulse resistance, multiple or mobile war headquarters, succession of command, and change of location of command, non nodal organisation and communications, manual fallback to anticipate equipment failures, and standard operating procedures. Damage tolerance and SUC OC are the principal means by which survivability will be attained for command and control systems.

Two interesting developments are shown in Fig. 5. One is the development of JTIDS - Joint Tactical Information Distribution System or MIDS - Multifunction Information Distribution System as it is sometimes called and the NATO Identification System. As the diagram shows, communications will be a complex and very busy affair. One of the major areas of difficulty will be identifying friend from foe when the opposing forces may well have a sophisticated capability for jamming or presenting themselves incorrectly as being friends. The advent of these new digital communications systems does offer very great potential for improving the coordination of air defence and of coordinating offensive and defensive operations.

The wide range of communications systems and communications techniques which are expected to be employed in Europe during the next 10 to 15 years have really only one common feature; eventually they will all tend to be digitally based rather than analogue based as is the current situation.

Figure 6, which appears to be a very busy diagram, is an N^2 interface diagram. I would ask you not to look at the detail particularly but to half-close your eyes and see the patterns that emerge. Down in the bottom right hand corner you will see a complex interface pattern. This is the air defence structure. It is a close knit, mostly airborne element where the various participants communicate principally by radio. Radio links are shown as ovals as opposed to land-line links which are shown as circles. The other leading diagonal boxes are the various participants. You will see the major subordinate commander in NATO at the top left hand box. If you come down to the 5th box, you will find air command and control system headquarters. The ACCS headquarters is at the centre of a cross indicating its nodal nature in the system. The N^2 diagram, which is being used here merely for communication with an audience, is in fact a powerful analysis tool. It has a potential for studying functional and physical partitioning in a non-numerical style which is generally a particularly difficult system design task. You will notice that there are certain interfaces which might reasonably be expected to occur. For example, the maritime operations control has an interface with the headquarters naval force. Similarly, the transport operations control has an interface with the headquarters' land forces. You will observe that the major subordinate commander has interfaces with each of the various headquarters and with the AWACS. AWACS itself, or AEW, also appears as a node in the system indicating both its importance and vulnerability.

An alternative analysis technique is known as data state analysis. An example of that is shown in fig 7. Fighter recovery is presented as a logic dependency diagram showing the two different cases for recovery. One is priority recovery because of incapacity of the crew or the aircraft to continue. The other is a precautionary recovery; although the aircraft is fit to continue its operation, there is actually no need for it so to do. The logic of the diagram is such that it is possible to draw in a dotted line (such as the one that appears on the diagram) separating the tasks into two groups. Where the line goes is a particularly important factor because it represents the physical separation of tasks. That means that there need to be criteria for deciding whether a process should be executed to one side of the line or to the other. One of the major tasks in analysing air command and control is to decide where the task should best be done. This diagram is the beginnings of such an analysis.

A slightly more contentious example is shown in fig 8. This is the organisation of identification information. This diagram shows that identification comes really in two sorts. Firstly, automatic identification by systems or communications. That is an easy category to deal with. Secondly, unidentified tracks which are not so easy to deal with. The output from this organisational logic diagram is an identification message into the indirect sub system of NATO Identification System.

That is all I want to say for the moment on air command and control systems in general. I would like now to turn to some forward looking C^3I concepts. The first is HITAS which is an airborne automatic threat assessment system for single seat aircraft. Next, a commander's central threat assessment system; third, answering the commander's "What - if?" questions. Finally, some advance Human Factor concepts.

Taking Fig. 9 at block level only, this is a system intended to go into high speed, low level single seat aircraft. The top block is a thermal imager and its processor. The bottom block is a radar and its processor. On the left you will see non imaging sensors such as IFF, radar warning, laser warning and so on. The system which we have recently designed, takes the information from all these sensors and produces from them a composite picture in the multisensor recognition block. This uses scene correlation, threat extraction, threat recognition and correlation techniques. The threats, having been recognised, are then ranked and presented on a 3-D cognitive map. The outputs from the system are steer signals to evade the threats or reduce the degree of threat: automatic countermeasures and hazard warnings.

Fig. 10 shows a typical display which HITAS could present to the operator. The operator is at one of the focii of the bottom two ellipses and threats are represented in plan by M7, which is a missile, the Z's which represent anti aircraft artillery and an F for a fighter. Intended target positions are T1 and T2. The cross-hatched area indicates the degree of an impending threat. The system is designed to weave its way between various threats towards the targets.

Threat assessment can be executed at a much higher level than at the aircraft itself. Fig. 11 shows raid threat assessment in a headquarters. This consists essentially of pairing the incoming raid groups with one's own vulnerable target list to assess where the enemy is likely to hit and what damage he is likely to inflict. This is an area where computing can radically assist the commander in assessing the problem with which he is faced.

Another area of great interest is the area of artificial intelligence and expert systems. A command and control information system data base being used for conventional operations can also be used with a separate processor for answering the high level questions which a commander might want to pose. These include, for example, "what would happen if I lost all my area SAM?", or "what would happen if I lost such and such a headquarters?". Such questions are typical of those with which a commander is regularly faced. Current command and control systems do not assist him. We have been working in this area and some of the work is concentrated into the development of macro models. These are large scale but very simple models. One is being developed from ecological modelling which deals with competition for territory and the other is a 2-dimensional adaption from kinetic gas theory which may sound rather strange, but the results are quite reasonable.

Fig. 12 is one of two showing some results from earlier modelling. This particular model is the 2-dimensional adaption from kinetic gas theory applied to helicopter/helicopter battles. The graph shows a strategy in which the blue force has attacked orange-held territory overtly and draws back suffering severe losses.

In the next diagram, Fig. 13, however, blue has been more cunning and he enters orange-held territory covertly, groups his forces, masses them and then attacks orange in a coordinated manner. The result is that he inflicts considerable greater losses on orange forces. These models are at an early stage but it is our intention to develop them to the point where they form the heart of an intelligent system.

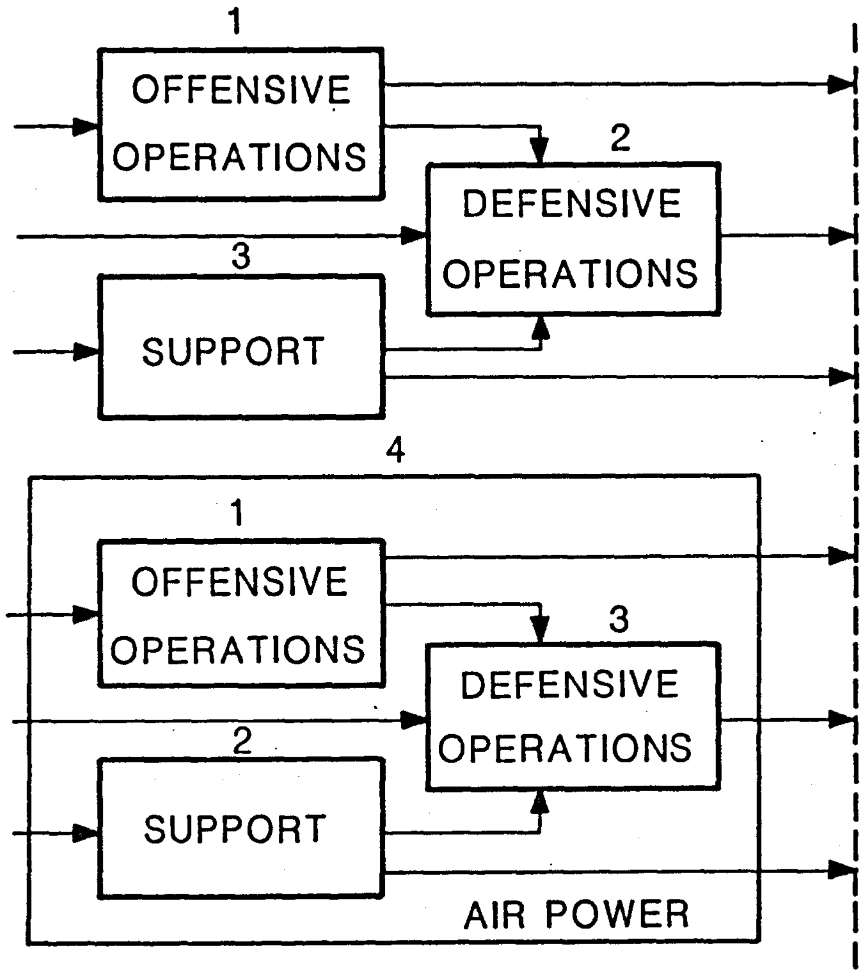
Fig. 14 shows a typical operator sitting at the hub of a number of communications systems. In more tactical systems, it is reasonable to suppose that the operator in a command and control systems spends much of his time interfacing with a computer system, a display system or a radar system or whatever. Our practical research in current projects shows that at higher level command and control, such as an ACCS HQ, this is not the case at all; the operator spends a very large proportion of his time interfacing with other personnel, either directly or through the telephone.

The number of occasions on which he uses ADP displays is relatively small. It certainly could be less than 25% in particular applications. This tends to place a different emphasis on the way in which the man machine interface should be designed. Central displays, maps, wall charts and similar non high technology devices very clearly have their part to play in a modern command and control system and indeed the concept of a CCIS operator is itself invalid in most cases. The CCIS becomes a management tool which may be accessed by any of a number of command team users.

Fig.15 shows a concentrated data display for executive use. The two halves of the circle represent past and present situations, to give trend indications. The circles in this case might represent a summary of 2-300 operational aircraft states, giving at-a-glance data to the commander. Such presentations can be useful for representing war consumables, POL and crews as well as large numbers of aircraft.

In summary, this brief description of various elements of air command and control covers a wide range of activities with which EASAMS concerns itself, from the philosophy, balance and structure of air power through architectures of command and control, down to operational analysis and studies to full systems implementation of weapon systems, with continuing emphasis on the human at the heart of the systems.

SYSTEM DESIGN AXIOM AND COROLLARY



AXIOM

'A SYSTEM COMPRISED OF ELEMENTS WHICH HAVE BEEN SEPARATELY OPTIMISED, WILL NOT ITSELF BE OPTIMISED'

COROLLARY

'AN OPTIMISED SYSTEM WILL CONTAIN ELEMENTS WHICH ARE NOT, OF THEMSELVES, OPTIMAL'

REASONS:

(A) 4 IS A NEW FUNCTION. ITS OPTIMISATION HAS NOT BEEN ATTEMPTED.

(B) THE INTERFACES BETWEEN THE ELEMENTS 1, 2 AND 3 HAVE NOT BEEN CONSIDERED IN THEIR CONTEXT OF 4.

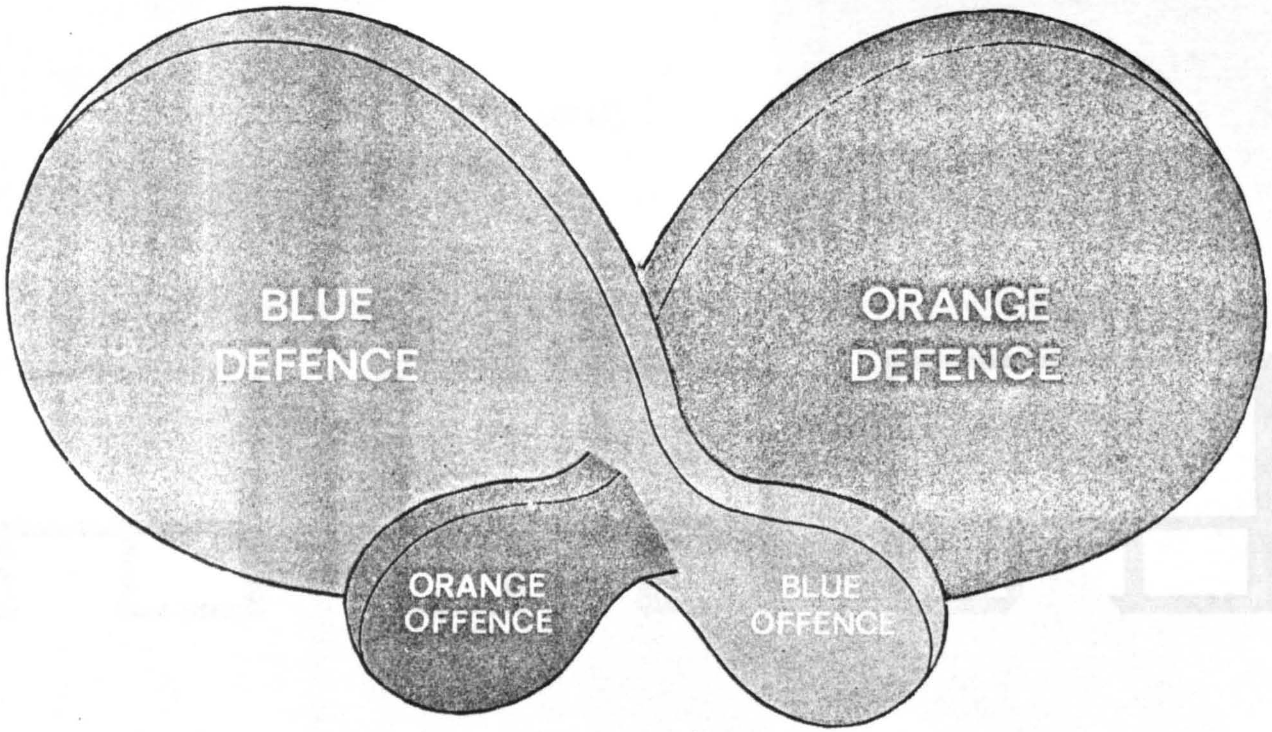
CONCLUSION:

IMPROVING THE EFFECTIVENESS OF AIR POWER REQUIRES A TOTAL SYSTEM APPROACH, RATHER THAN A FUNCTION-BY-FUNCTION DEVELOPMENT

Figure 1

EASAMS

AIR POWER – MEASURES OF EFFECTIVENESS



BLUE AIR POWER EFFECTIVENESS

= f | *BLUE DEFENCE v ORANGE OFFENCE*
BLUE OFFENCE v ORANGE DEFENCE
BLUE DEFENCE v BLUE OFFENCE |

Figure 2

EASAMS

ACCS SCHEMATIC

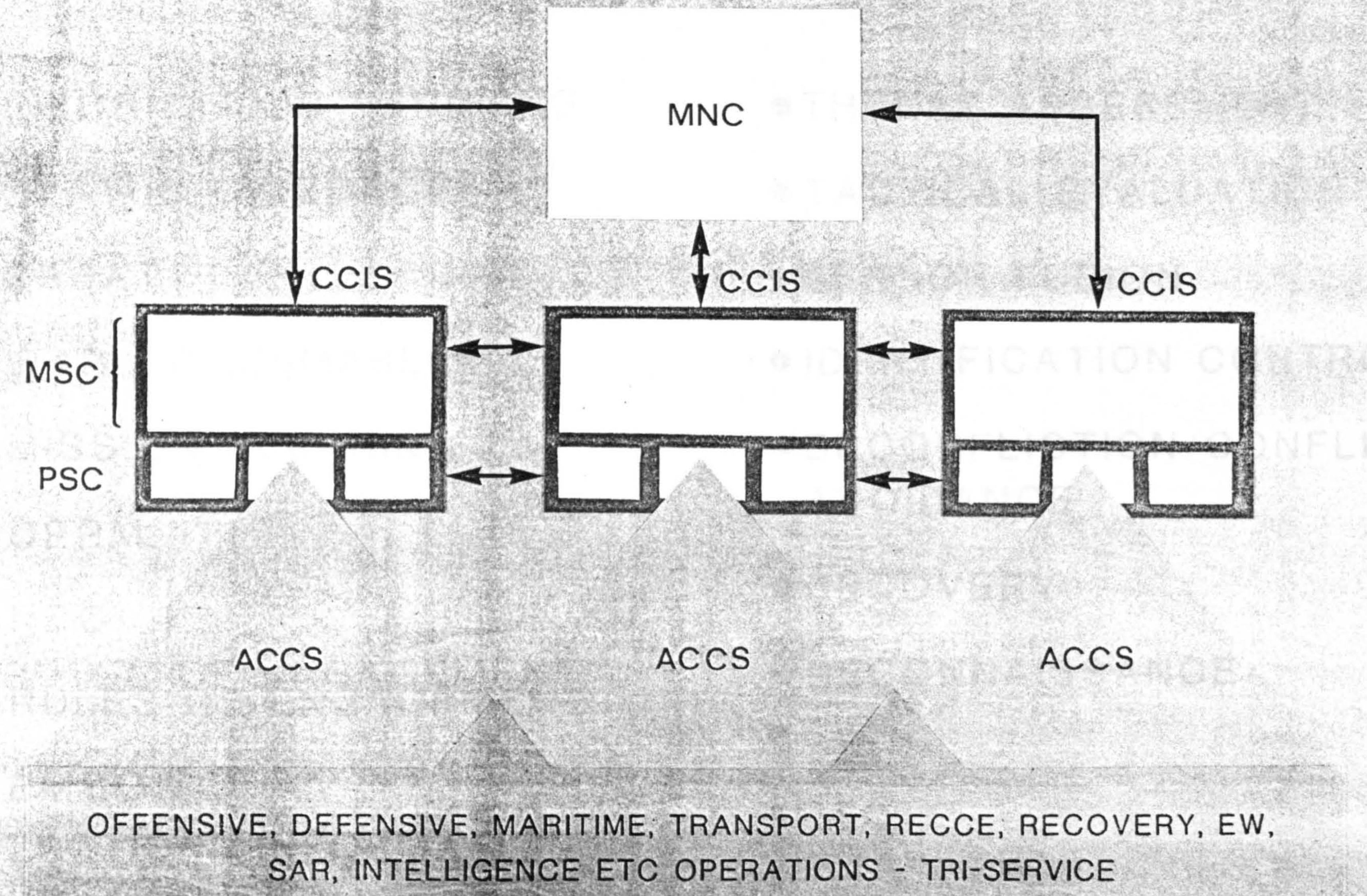


Figure 3

ACCS CAPABILITIES

- INFORMATION HANDLING
- STATUS OF FORCES
- ORBAT
- WAR CONSUMABLES
- MISSION PLANNING
- CPRM
- NBC
- RULES OF ENGAGEMENT
- EMCON
- TACTICAL INFORMATION
- THREAT ASSESSMENT
- TACTICAL EVALUATION
- SENSOR FUSION
- IDENTIFICATION CONTROL
- DECONFLICTION/CONFLICT AVOIDANCE
- RECOVERY
- RECONNAISSANCE
- EW
- ETC.

EASAMS

COMMS IN ACCS

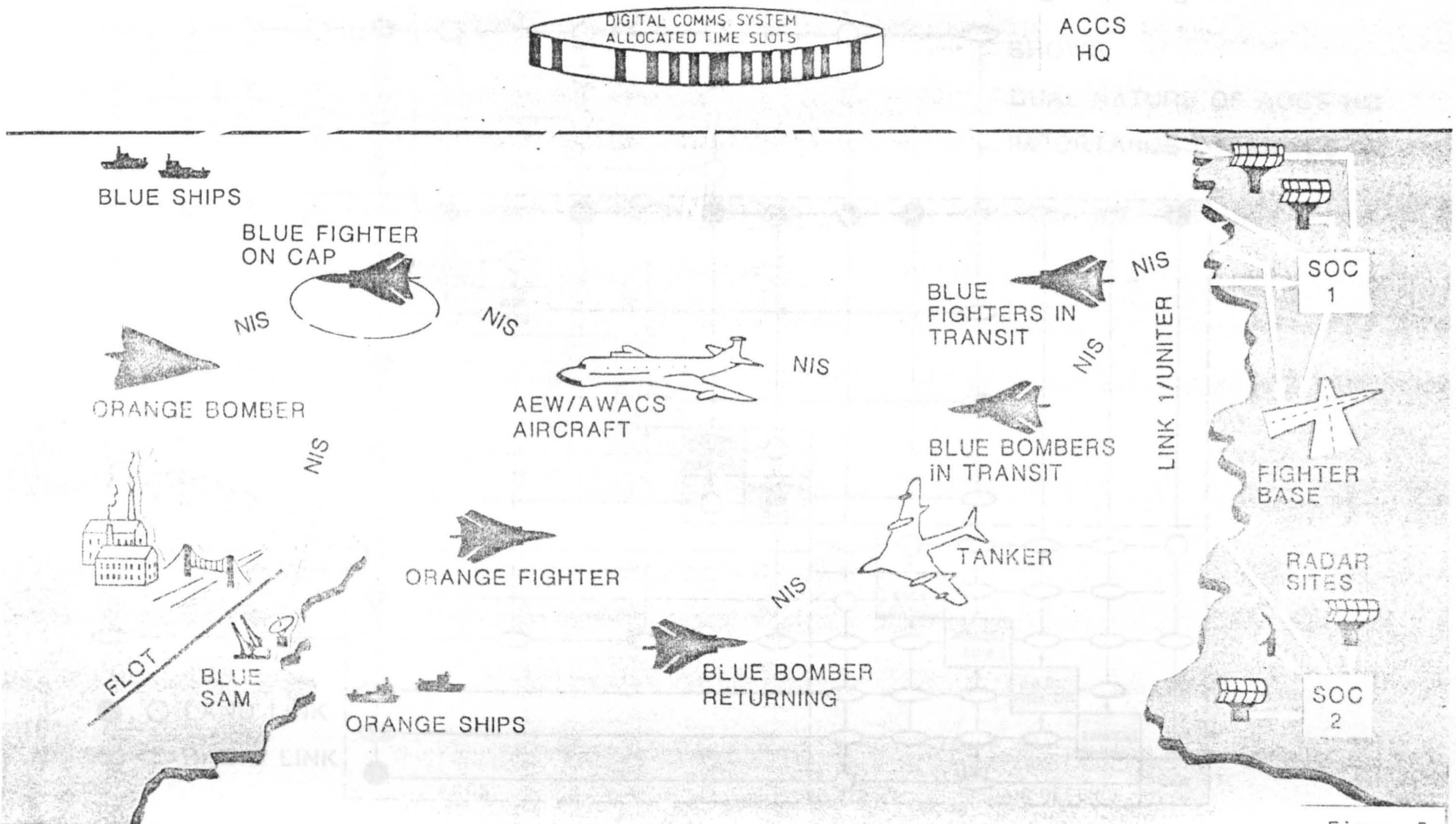


Figure 5

ACCS - PRIMARY INTERFACE DIAGRAM

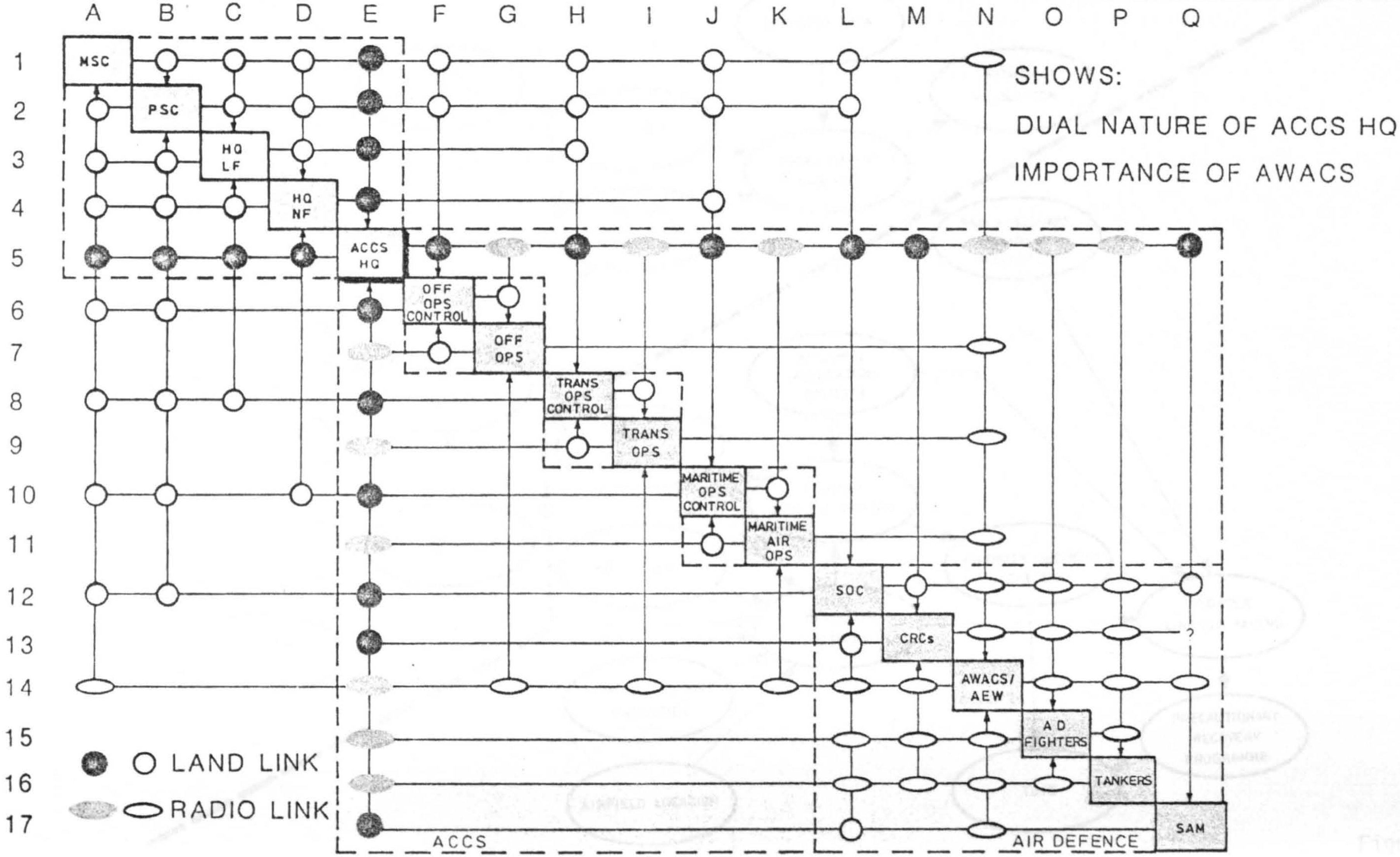


Figure 6

FIGHTER RECOVERY DATA STATE ANALYSIS–2ND LEVEL

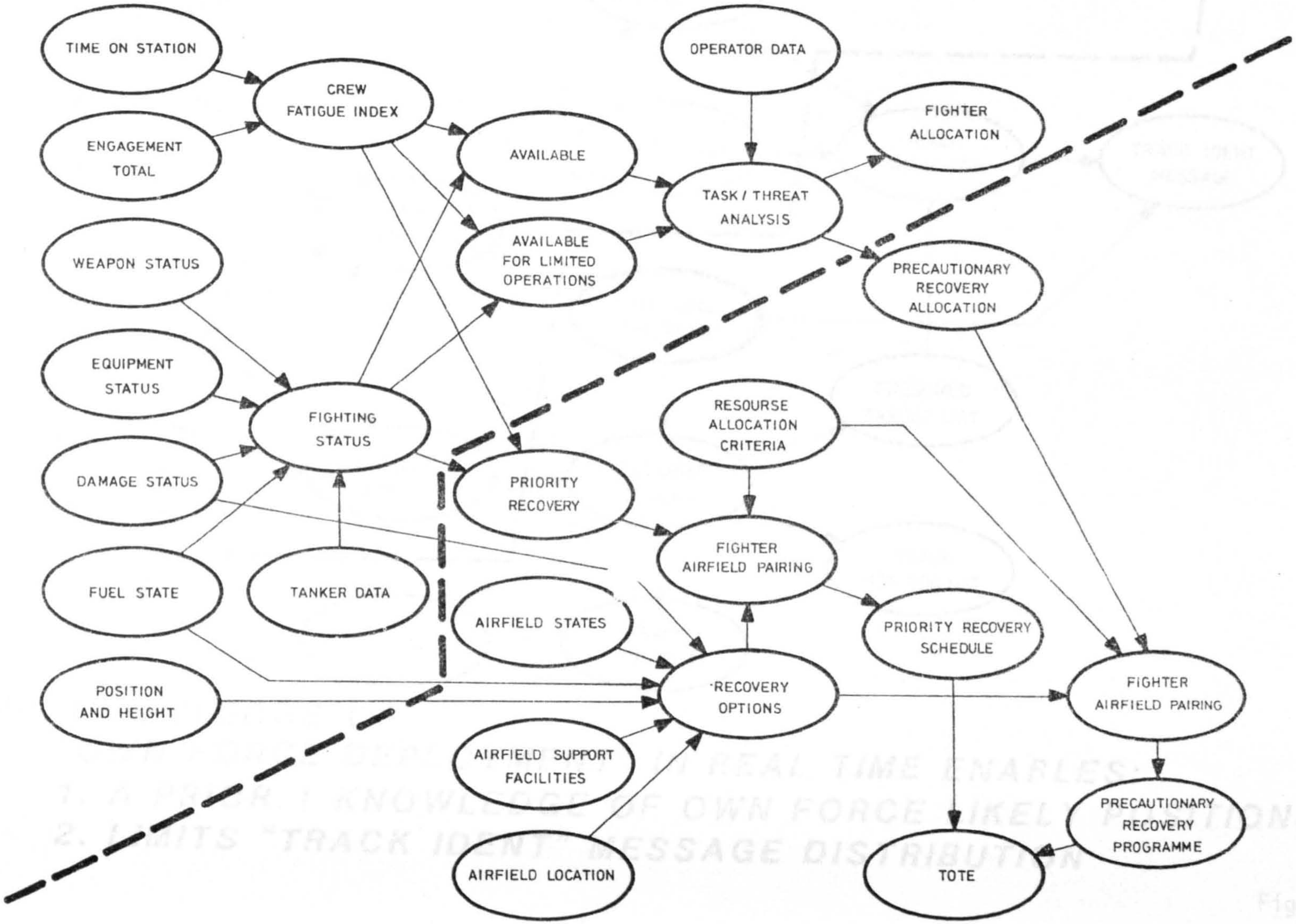
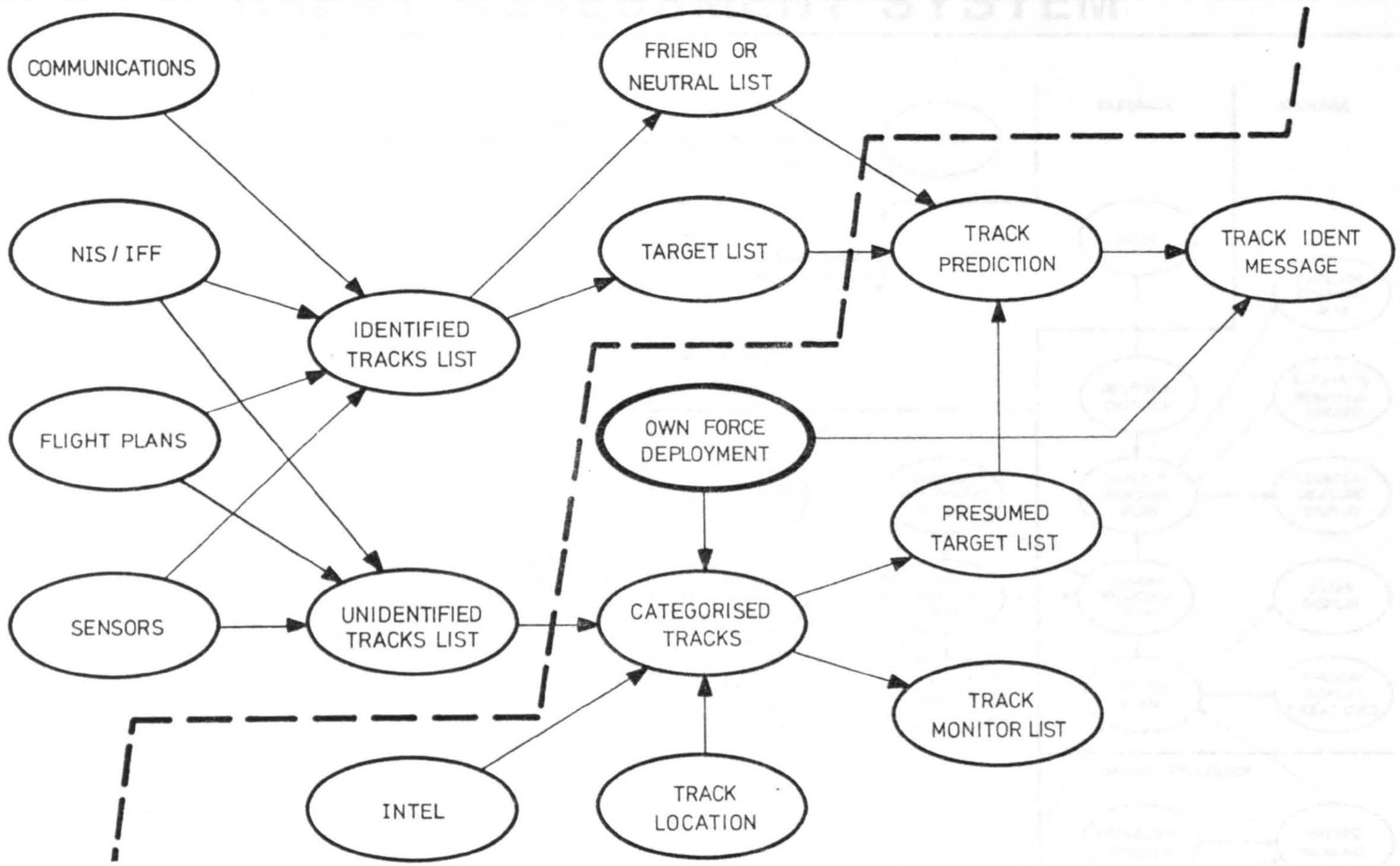


Figure 7

ACCS-ORGANISATION OF IDENTIFICATION



NB. KNOWLEDGE OF
"OWN FORCE DEPLOYMENT" IN REAL TIME ENABLES:
1. A PRIOR 1 KNOWLEDGE OF OWN FORCE LIKELY POSITIONS
2. LIMITS "TRACK IDENT" MESSAGE DISTRIBUTION

Figure 8

EASAMS

HITAS - HEURISTIC INTELLIGENT THREAT ASSESSMENT SYSTEM

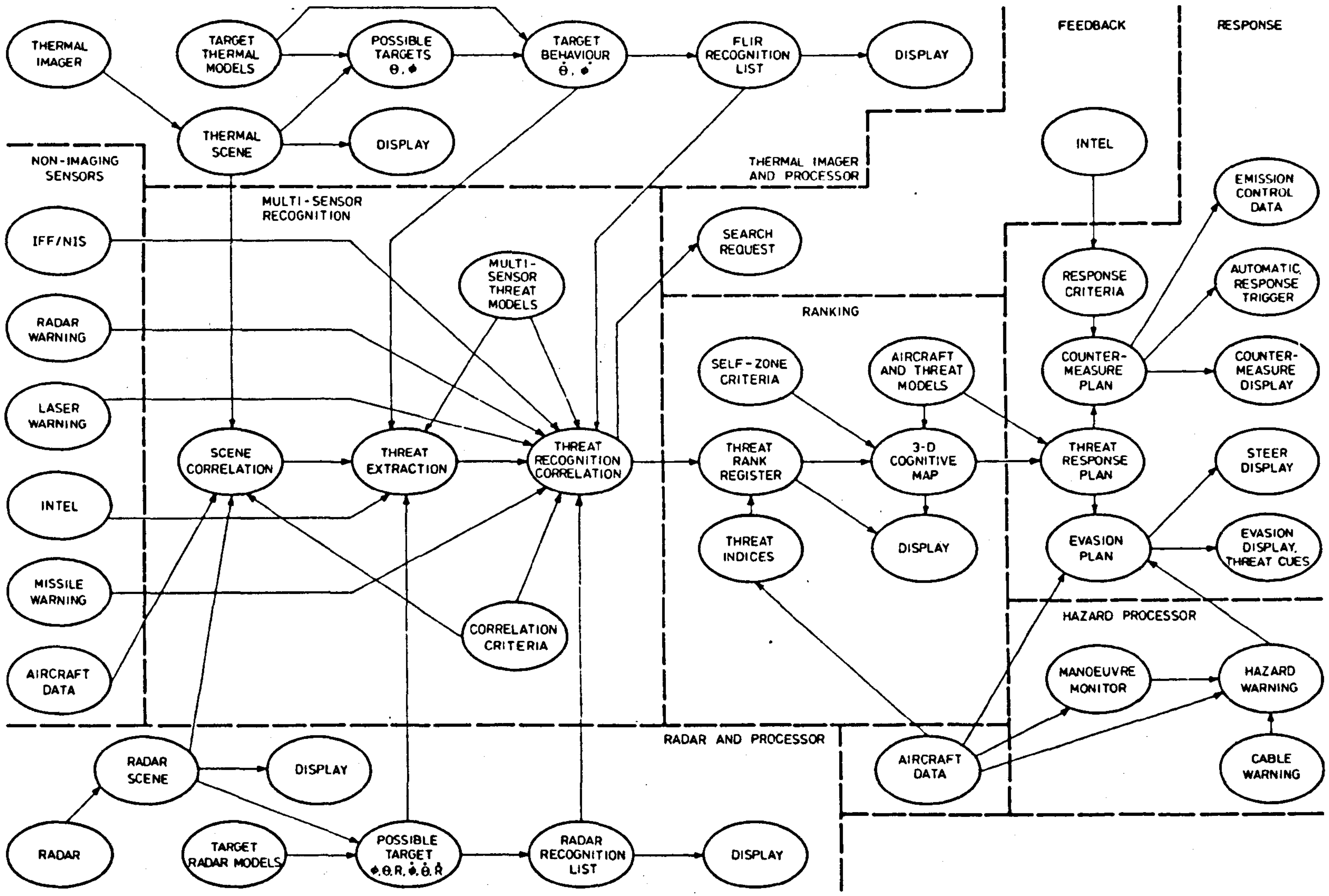


Figure 9

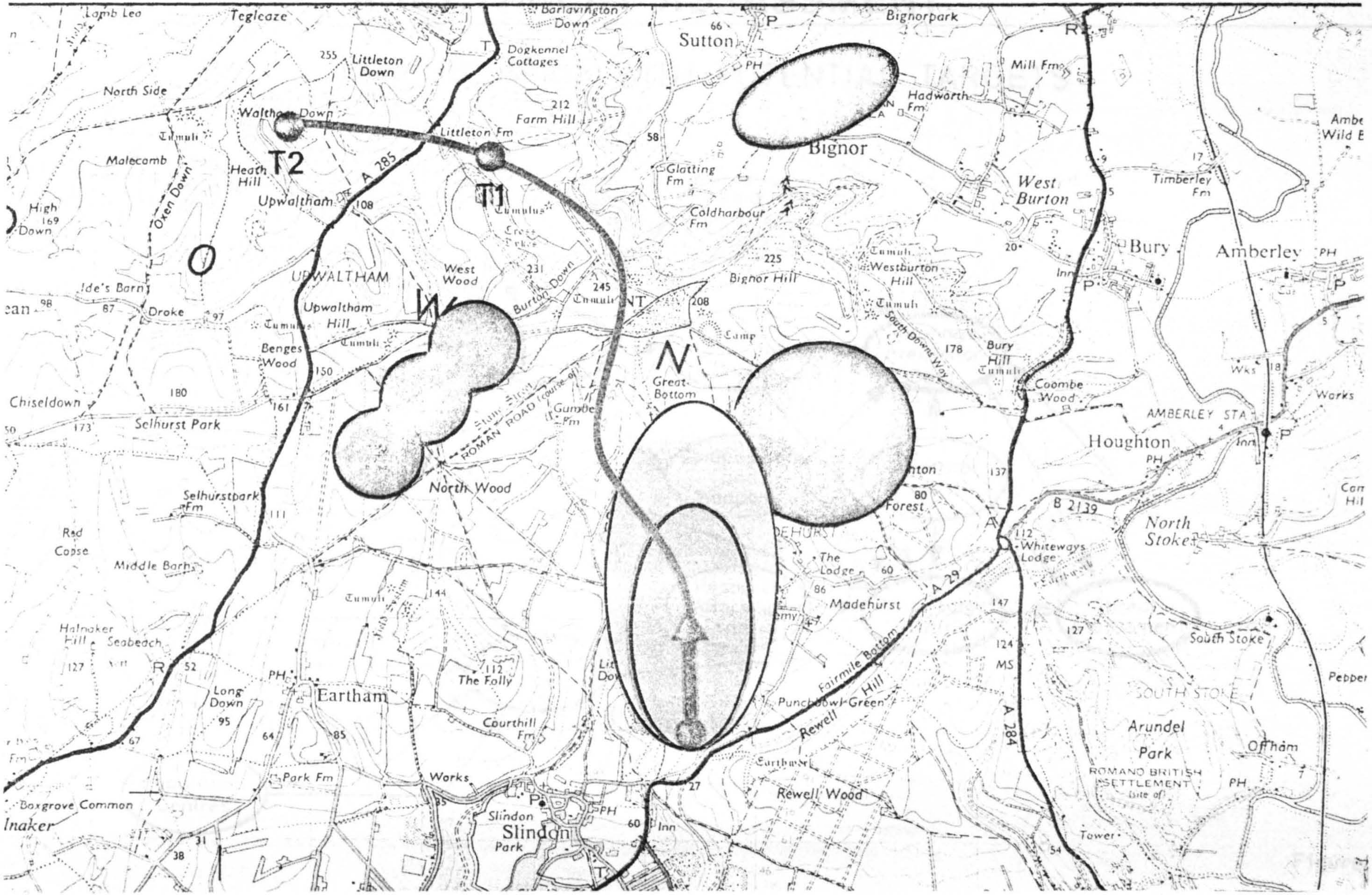

T2
T1
Waltham Down
Littleton Fm
Littleton Down
Tegleaze
Sutton
Bignor
Bignorpark
Hadworth Fm
West Burton
Bury
Amberley
Upwaltham
Upwaltham Hill
North Side
Malecomb
Heath Hill
Benges Wood
Selhurst Park
Chiseldown
Selhurstpark Fm
Middle Barn
Great Bottom
North Wood
Bignor Hill
Westburton Hill
Bury Hill
Coombe Wood
Houghton
AMBERLEY STA
North Stoke
Whiteways Lodge
Madehurst
The Lodge
Eartham
The Folly
Courthill Fm
Long Down
Slindon
Slindon Park
Punchbowl Green
Rewell Hill
Rewell Wood
Fairmile Bottom
Arundel Park
South Stoke
Offham
Boxgrove Common
Park Fm
A 285
A 29
A 284
B 2139
ROMAN ROAD (course of)
Stane Street
South Downs Way

Figure 10

RAID THREAT ASSESSMENT

"PAIRING" RAIDS WITH POTENTIAL TARGETS

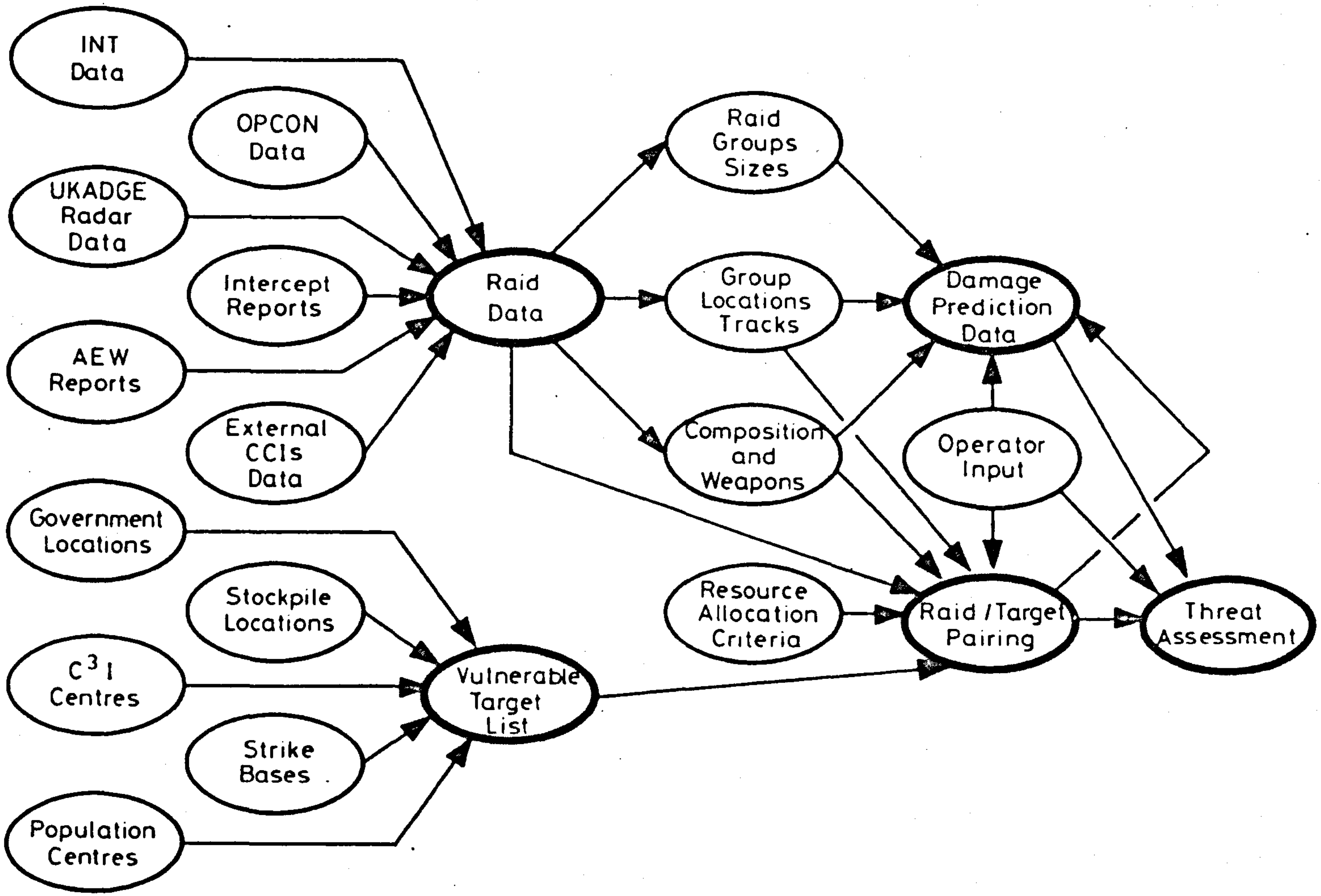


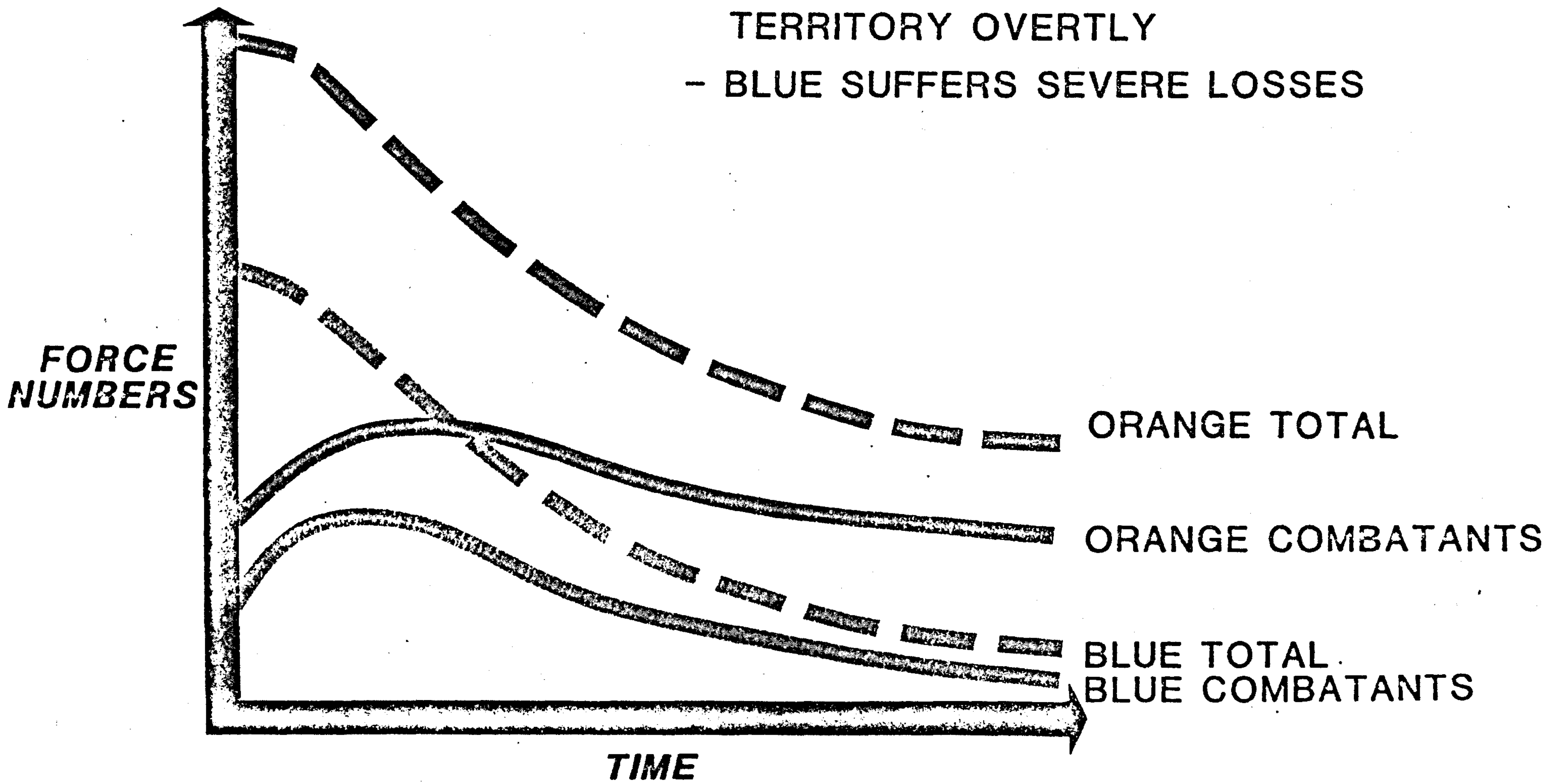
Figure 11

EASAMS

HELICOPTER MACRO BATTLE MODEL

STRATEGY – BLUE ATTACKS ORANGE–HELD TERRITORY OVERTLY

– BLUE SUFFERS SEVERE LOSSES



HELICOPTER MACRO BATTLE MODEL

STRATEGY – BLUE ENTERS ORANGE-HELD TERRITORY COVERTLY AND GROUPS FORCES

– BLUE MOUNTS SUPRISE MASS CO-ORDINATED ATTACK

– ORANGE SUFFERS SEVERE LOSSES

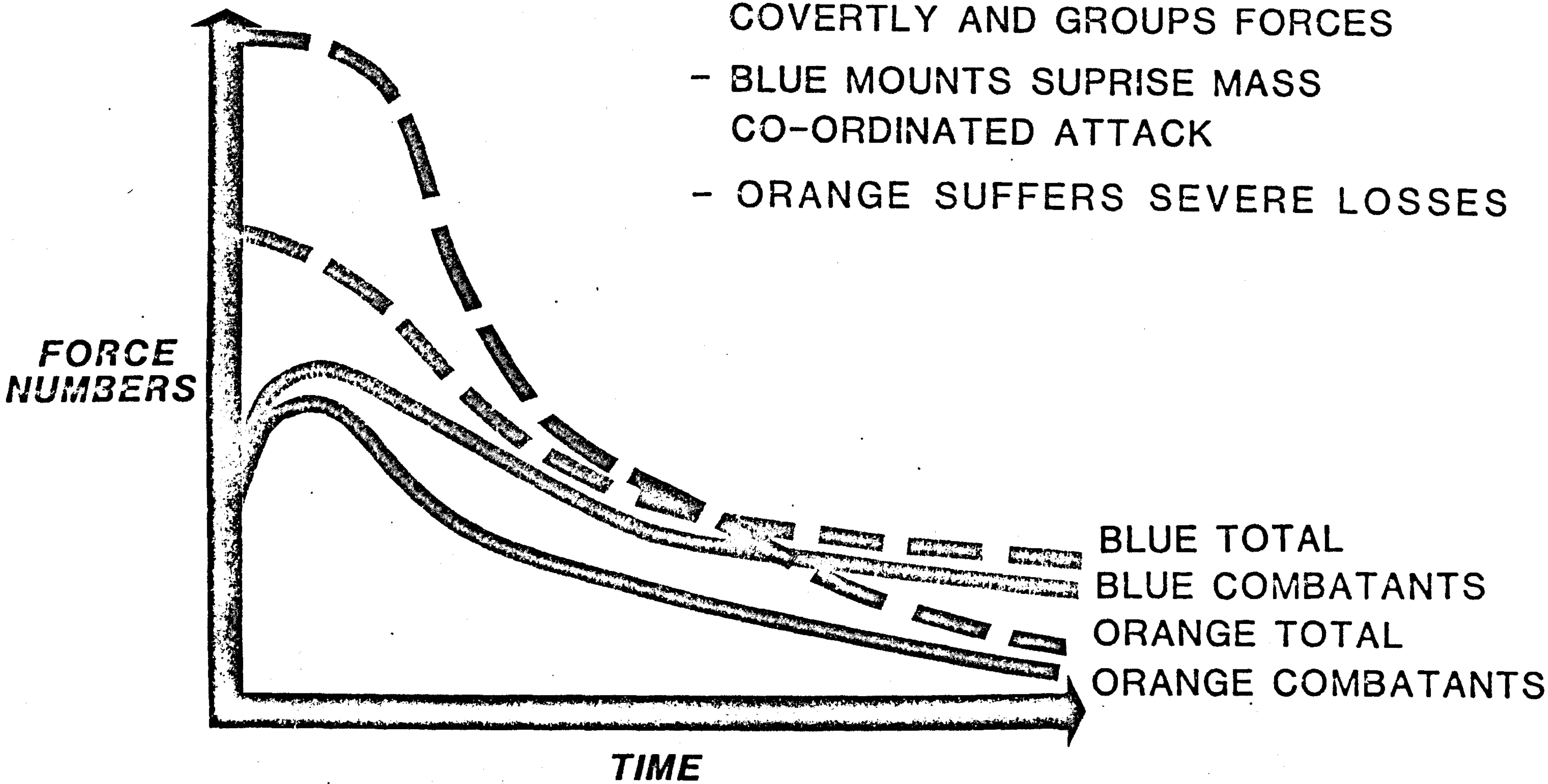


Figure 13

COMMAND TEAM COMMUNICATION ENVIRONMENT

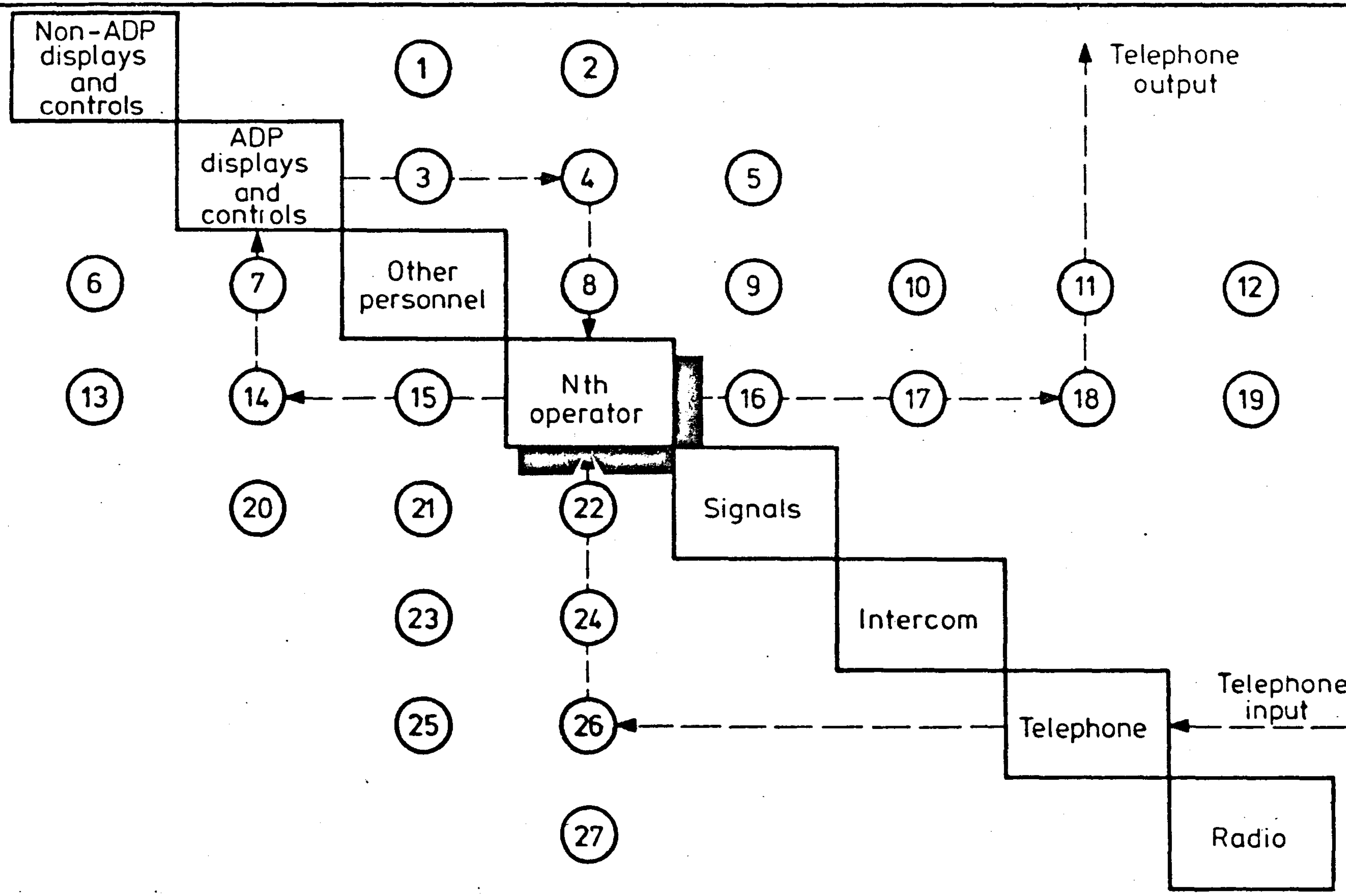
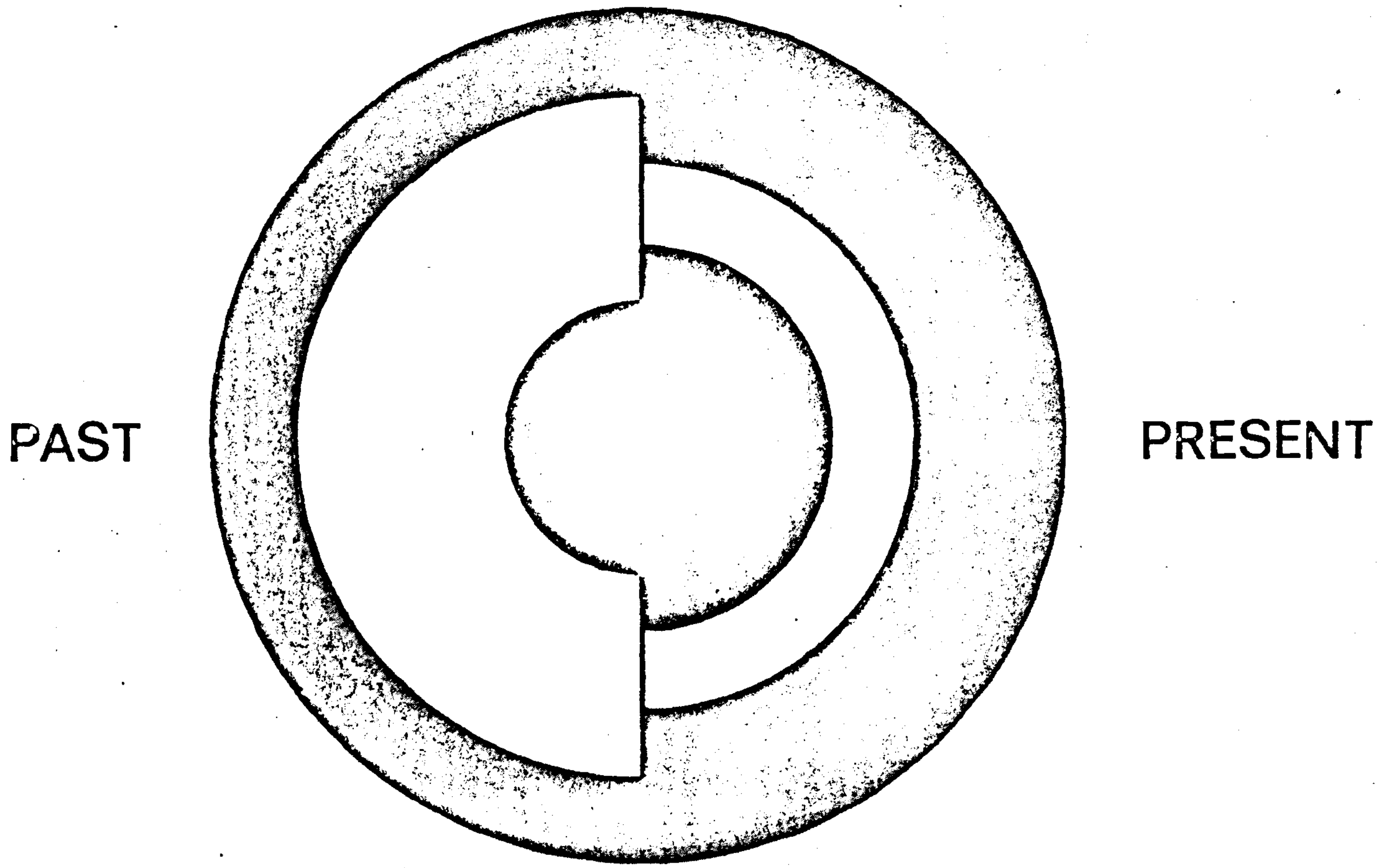


Figure 14

EASAMS

CONCENTRATED DATA DISPLAYS



COMBATANT   AVAILABLE IN x HOURS   NOT AVAILABLE

Fgure 15

AVIONIC SYSTEMS ARCHITECTURE

A paper to be presented at the RAeS Conference

on

Digital Computer Technology for Airborne Applications

15th March 1989

Professor D K Hitchins

City University

D O'Dwyer

EASAMS Limited

AVIONIC SYSTEMS ARCHITECTURE

Professor D K Hitchins
City University

D O'Dwyer
EASAMS Limited

SUMMARY

New avionic architectures are required to help provide cost-effective solutions to current and envisaged requirements for combat aircraft. Architectural concepts like those pioneered in the PAVE PILLAR programme will give greater flexibility in design but pose new problems for designers.

In this paper the challenges facing the avionic systems architect are considered, the need for a powerful design methodology is discussed and a possible approach based on cumulative selection principles is described.

INTRODUCTION

There are three fundamental attributes for the operational effectiveness of a force comprising a number of weapon systems:

- Availability of the systems
- Survivability of the systems in operation
- Performance of the systems in the operational environment.

There is also a vital fourth attribute without which it will not be possible to establish and operate a force:

- Affordability.

Present operational effectiveness is threatened by the impact of increasingly hostile environments on the performance and survivability of weapon systems and by the workload imposed on aircrew in critical mission phases. Effectiveness is further threatened by the difficulty of keeping current systems available within reasonable cost constraints.

More flexible and adaptable system architectures which take advantage of new digital data handling technologies will help designers to provide cost-effective solutions to meet increasingly onerous requirements, provided that a powerful architecture design methodology can be implemented.

NEW CHALLENGES FOR THE AVIONIC SYSTEMS ARCHITECT

Previous technological advances, particularly in data buses, microprocessors and microelectronic components have provided a route for progressing from architectures based on a central computer configuration to more flexible arrangements with multiple buses and distributed system computing. Improved sensors and displays have provided a greater capacity for acquiring data and displaying information to aircrew. However, these bus-based architectures, still with relatively conventional sub-system partitioning, and with system functions realised in equipment Line Replaceable Units (LRUs) do not have the potential flexibility in design, adaptability in operation or economy in-service that are necessary to meet envisaged future requirements.

We have reached the end of the road with conventional architectures. If designers are to provide, with acceptable installation penalties, this flexibility, adaptability and economy, and the capability for through-life improvement that users will demand then new architectural concepts and new design approaches must be introduced.

Advanced Avionic Architectures

Advanced architectures, which take advantage of new technology and are based on common Line Replaceable Modules (LRMs) and high speed data buses, are seen as keys to the development of a new generation of flexible, adaptable airborne weapon systems with an acceptable cost of ownership.

In the US, the PAVE PILLAR programme has been set up to develop the foundation for the next generation of integrated avionic systems (Ref. 1). Implementation strategy is based on the system-wide utilisation of common LRMs with advanced microelectronics, integrated racks and fibre-optic communications. The modules support programmable processing, input/output and memory functions. Reconfiguration capability to meet mission reliability and moding requirements is a feature.

In contrast to current systems with their subsystem groupings, three integrated functional areas have emerged from the PAVE PILLAR architecture as boundaries for resource sharing and reconfiguration. These are sensor management, mission management and vehicle management.

PAVE PILLAR has been a catalyst for change in UK, too (Ref. 2). A PAVE PILLAR collaborative programme has been initiated with the US and an Advanced Avionics Architecture and Packaging (A^3P) programme has been launched by UK MOD.

Potential benefits in effectiveness and affordability from the adoption of the new concepts are considerable. However, it has been emphasised (Ref. 3) that the benefits will only be realisable if many factors other than the purely technical are considered. These include design methodology.

Avionic Architecture Design

For system designers, advanced avionic architectures will bring much greater flexibility and the chance to influence directly many system attributes which are important for effectiveness and against which a possible configuration must be evaluated. These attributes include:

- Adaptability to mission role
- Failure and damage tolerance
- Interoperability within platform and force
- Efficiency in using system resources
- Effectiveness of man-machine interface
- Integrity in safety critical applications
- Invulnerability to electromagnetic hazards
- Fault detection and diagnosis capability
- Expandability for planned improvement.

The penalty for this flexibility is that the designer - the system architect - may lose control of the design process and arrive at a solution which is markedly less than optimum.

To put the search for an effective design method in context it is useful to outline the design approach. Avionic systems design is an iterative process with architecture design at its heart. Six stages can be identified:

Design Boundary. First the design boundary is delineated to define what is within the avionic field of responsibility and, equally important, what is outside it.

Architectural Framework. Within the boundary an architectural framework is set up. This is defined in terms of basic system modules, their possible disposition within the platform and allowable interconnections. It may be based on an existing system or on new but experimentally proven concepts. However, it will always be necessary to establish both the range of viable options for allocation of functions within the framework and the architecture design goals to be met.

Supporting information which will constrain the design includes system interface characteristics and hard system constraints.

Interface characteristics and constraints may be operational or technical and will reflect requirements at many levels. They will be derived from the weapon system requirement itself, from conceptual design documents, from platform and weapon specifications, from force requirements and from fleet-wide considerations.

Requirement Analysis. Analysis of the system requirement leads to a functional design concept and the identification of functional entities and their relationships at the lowest level of interest relevant to architectural design.

Functional Analysis. System rules for functional to physical mapping during detailed design are generated using functional analysis techniques.

Architecture Design. Viable options are generated for the implementation of entities by processes, data structures and data flows within the architectural framework.

Architecture Analysis. The viable configurations are analysed to establish their impact on system availability, survivability, performance and affordability.

With PAVE PILLAR or even with simpler architectural sub-sets the task of establishing optimal avionic configurations for an airborne weapon system in a complex operational environment will only be manageable if tackled in a logical, systematic and efficient way.

Beyond PAVE PILLAR, the proposed PAVE PACE initiative (Ref. 1) gives a glimpse of even more powerful architectures that might be available for the avionics designer in the longer term. These require massive processing power for display graphics, real-time artificial intelligence, advanced signal processing and selective automation. Powerful meta-functions are envisaged. These are integrated affinity groups of functions implemented in families of parallel processing networks supported by a PAVE PILLAR communication framework.

What design approach can the systems architect adopt when faced with such choice and complexity?

THE NEED FOR A POWERFUL ARCHITECTURE DESIGN METHOD

Architecture is burgeoning as a new science, fuelled by advances in processing, software structures, protocols and - in the case of avionics - the quest for major advances in performance, availability and survivability. PAVE PILLAR has shown that the UK, once pre-eminent in design if not production, has been overtaken by more visionary concepts in the US which offer exciting prospects.

British designers find themselves in a 'Catch 22' situation. Innovative architectures could be conceived, and UK designers have the skills. If such architectures were introduced, however, they could threaten the avionics industry status quo; subsystem manufacturers may lose elements from their traditional subsystems into a re-organised architecture. Thus architectures which might benefit the complete aircraft, might detract from company business for the designer; PAVE PILLAR could cause that degree of turbulence. Is PAVE PILLAR the optimum approach, however? Why have the subsystems been partitioned as they have? Why are the internal data bus communications so structured? To address such questions, and to move confidently forward to PAVE PILLAR and beyond, the UK needs to apply new concepts to the design of architectures. This section and the next look at one possible approach.

To avoid the temptation of conventional wisdom, it is necessary first to abstract architectural design to the level of General Systems Theory. At that level of abstraction, an architecture comprises clusters and links. The clusters are composed of related entities. The links connect the entities within each cluster and there are cluster-to-cluster links too.

Entities in an avionics cluster are operational and process functional primitives - a target detector, a tracker, a processor, a convertor, a display, a launcher, an attitude sensor, a velocity/acceleration sensor and so on. Target detectors with different characteristics (spectral coverage, for instance) would be discrete entities. Links are abstractions of the internal communications implicit within the subsystem and the system architecture, and the links can also cluster between themselves.

Clustering

If the design problem starts with an amorphous set of primitive functions, then there must be some 'drivers' or influences which both encourage clustering and which also determine the number of clusters and their constituent functions. Also, some clustering solutions will be better than others. (In PAVE PILLAR, ICNIA may be regarded as a cluster for which one of the drivers was the management of mutual radio interference between communications, navigation and identification.) There is, too, a particular number of clusters for any given number of functions which minimises the sum of internal cluster links and external cluster-to-cluster links; this is of significance with respect to interface, link length and weight reduction.

Clustering algorithms influence the types of cluster. Where the nucleus of a cluster has already formed, unattached entities may be attracted on one of several bases:

a) Strong association with one member of the cluster

b) Association with N out of M of the current cluster members

c) Association with all cluster members.

These different clustering strategies result in different types of cluster - strung-out, clumped and very tight; they also affect the number of clusters and the tendency for entities to be 'left over' because there is no clear association. It is noteworthy that (a), (b) and (c) above form a continuum of relationship, as the value of N goes from 1 to M; this factor may favour a particular solution method, as will be indicated below.

Clustering Drivers

Clustering influences must be real-world related for the function clusters to be useful. An incomplete list of influences is set out in Table 1.

With so many different and often mutually incompatible drivers, the process of cluster optimisation is fraught with difficulty and special methods of accommodating influences, combinations and the mass of information become essential.

Linking Drivers

Links within and between clusters are also subject to real-world influence. Link bandwidth is a function of entity entropy - the classic dilemma of sending all the source information at a cost in bandwidth, or pre-processing to save bandwidth at a cost in reduced information content. Links can also proliferate beyond the basic need for connectivity to provide redundancy. Physical links can be greatly reduced by clustering, not only by the obvious cluster-to-cluster broadband link, but also in developing equally-sized clusters, by reducing the overall cluster infrastructure. These various factors can also be considered as underpinning a linking continuum, from simple trees giving singular point-to-point connections through loops to networks and 3-D networks, giving increasingly redundant connectivity and reconfigurability. Examples of this continuum will be presented below. Sufficient groundwork has been laid at this point to enable the search for a method of architecture design.

Table 1 Clustering Influences

Influence	Example
Physical Juxtaposition	- Displays and controls with crew - Antennae with rf front ends
Physical Location	- Radar and ESM antennae isolation - Communications antennae
Functioning Relationships	- Signal reception and signal conditioning
Installation	- Centre of gravity - Self-armouring (lower priority devices protecting vulnerable devices)
Ease of Development	- Self-contained subsystems
National Fit	- Self-contained EW fit
Number of Clusters	- Infrastructure reduction, moding control, survivability, etc.
Mutual Interference	- JTIDS/GPS/TACAN/NIS competition for bandwidth and time - ECM/ESM lookthrough
Like-to-like Attractors	- Frequency synthesisers - Data processors - Interfaces
Intra-cluster Access	- Diagnosability/Maintainability

Methodologies

The initial attempt at a methodology follows the sequence:

- Decompose operational prime directive into functional primitives
- Cluster primitives into groups - modules, line replaceable units, line replaceable modules, etc.
- Link clusters
- Draw architecture.

This simple approach couples the previously discussed features. Even at high level it has a major problem. The clusters and links are optimised in their own rights; it follows that the system (complete architecture) cannot be optimal since no optimum system can be comprised of separately optimal parts.

A better approach might be:

a) Decompose prime directive into functional primitives

b) Cluster primitives

c) Link clusters

d) Draw architecture options

e) Measure architecture cost-effectiveness

f) Select best architecture option

g) Loop to (b)

h) Continue until peak cost-effectiveness is achieved.

While overcoming the optimisation problem, this new method contains some difficulties to be addressed.

Functional Decomposition. Functional decomposition is not yet a formal process; given a prime directive for, say, a fighter to maximise the ratio of 'shooting opportunities to shot at opportunities', there is no consensus view on the processes which will decompose consistently, completely and without preconception and prejudice. This is an area presently being researched.

Clustering Algorithms. The potential combinatorial explosion implicit in the large number of primitive functions, together with the diversity and mutual incompatibility of the many clustering drivers makes clustering a major issue. A new method is being researched, based on cumulative selection, which might contribute to solving the problem - see below.

Cluster Linkage Generation. Again the structural combinations could be many, although this intraoperability structure may also be resolved in principle by cumulative selection.

Formal Cost-effective Trade-offs. Cost-effectiveness is a much maligned term. Cost-effectiveness Analysis (CEA) is not always undertaken in a formal, justifiable and traceable manner, however. Figure 1 shows the high level structure of the approach. The figure shows Performance, Availability (of Performance) and Survivability (of Performance) combining to provide Effectiveness, which then associates with Cost; the combination and association processes are crucial. Often the combination process is some weighting and scoring system which weights subjectively and adds numbers algebraically even though they represent different dimensions and should be summed vectorially. Given a measure of effectiveness, CEA often resorts to dividing Effectiveness by Cost for each putative solution, selecting the highest ratio as preferred. This approach ignores thresholds and the utility of cost and effectiveness.

CUMULATIVE SELECTION - A POSSIBLE APPROACH

Cumulative selection is the process which supports biological evolution. Parents produce offspring which differ marginally from themselves and from each other. A hostile environment favours some differences and disfavours others; successive generations adapt and evolve towards a more survivable and effective 'design', even as the environment itself changes. 'Design' is an inappropriate word; the process of cumulative selection is blind, it has no distant goal at which to aim. Instead it finds the best solution by gradually balancing a very wide variety of influences so that the being operates successfully within its niche. The process is not random. Successive generations retain the advances made by their forebears where these advances remain valuable.

Developing the Concept

Early research has produced a prototype tool (Ref. 4) which emulates cumulative selection in a hostile environment inter alia for aircraft platforms and for communication networks. The tool uses analogies to genes which code for design features; genes for Aggression, Stealth, Co-operation, Shell and Infrastructure code for weapons, low observables, communications, armour and internal processing architecture. With a drawing routine - which must have its equivalent in nature - the genes can be used to draw a platform. Offspring are generated by varying each gene in turn, for each offspring; each varies from the central parent in one, and only one respect. With five genes there are ten offspring as each gene can be increased or decreased. This process emulates the work of Richard Dawkins (Ref. 5). Dawkins did not include the next step, however.

The offspring are subjected to a hostile environment comprising measures of:

- Performance
- Survivability:
  - Avoidance of Detection
  - Self-defence
  - Damage Tolerance
- Interoperability
- Security
- Availability
- Affordability.

Each measure is the output of an analysis or model. The measures interact with the respective gene settings to produce a CEA score for each offspring. The highest scorer becomes the parent and produces its offspring in turn. The process continues until successive generations produce the same CEA score. The initial SYSMORPH, the seed, can be set at any combination of platform features; choosing a currently favoured design enables the designer, while balancing a large and complex array of different influences, to see which one change would most improve his design cost-effectiveness and which would be the next most beneficial change.

Figure 2 presents the new concept; the cumulative selection process contrasted with the conventional approach. The new approach, still using conventional - and trusted - air combat models, allows a much greater potential to handle complexity within the machine.

The process outlined in Fig. 2 is natural selection, that is, cumulative selection in a hostile environment. It has similarities to the mathematical technique of hill-climbing, but takes a different route to optimisation; this alternative route allows the operator/designer to monitor the developing fighter design step-by-step.

In the real world the environment changes too, and a being must continue to evolve, if it can, to avoid extinction. In the computer-generated world of cumulative selection, that process of environmental change can be pursued too. Figure 3 shows a repeat of Fig. 2 at the left-hand side but now there is a new process. The fighter design of Fig. 2 was developed using a fixed enemy threat as an essential feature of the Many-on-many Combat Simulation. Once that fighter evaluation is stable it is possible in principle to turn the process around and use the evolved fighter as a counter to the enemy threat, which can be evolved in its turn. This process has been dubbed 'Systolic Cumulative Selection' because it beats between the two halves of the process like a heart.

Architectural Evolution

If a fighter and its threat can be evolved, can the process be applied to the fighter's architecture and its threat? It seems reasonable to expect so. After all, the melee of fighters and threat aircraft is essentially a highly dynamic architecture of sorts, with entities, relationships and communications. Figure 4 shows the concept. A seed architecture, probably an initial design of the architecture team, is used as a start point, and offspring are generated. Architecture offspring will vary from their parent in respect of which primitives comprise clusters, capacities, connectivities, physical positioning within the aircraft, redundancy, etc., and 'genes' will code for all the basic characteristics which define 'architecture'. The hostile environment will comprise models and analyses of architectural performance, availability of performance and survivability of performance within the physical framework of the current fighter design and structure. The preferred offspring will be the one with the single change most beneficial to the architecture, and that offspring will breed.

As with the fighter and its threat, it should be practicable to evolve the architecture and its 'hostile' environment which is in effect the aircraft layout, the aircraft skin and whatever may traverse it, together with the '-ilities' including affordability. An evolved architecture might, for instance, weigh less and allow:

- An increase in fuel
- An increase in payload
- Better aerodynamic performance in combat.

So, in principle, there is potential for Systolic Cumulative Selection between the architecture and the fighter. See Fig. 5.

At this point, it is necessary to emphasis that much of the foregoing is speculative. The basic processes have been tested in simple prototypes and Systolic Cumulative Selection has been addressed 'handraulically'. The processes do, on the other hand, seem to provide several invaluable features:

a) The processes emulate those pursued less formally by designers now

b) Many disciplines can contribute separately to the definition of the hostile environment

c) The process is very simple; each step is obvious, visible, methodical and traceable

d) The process is none-the-less very powerful in its ability to accommodate complexity

e) The procedure contains no 'black arts' and is understandable by user, operator, maintainer and designer alike.

Cumulative Selection in Operation

Finally, as a very simple example of cumulative selection, see Figs. 6 and 7 which show how a network is evolved between a fixed set of source/sinks. Nodes or communication switches are randomly scattered between them and, by relocating each node in turn cyclically, the preferred positions are cumulatively selected until the minimum overall path length of the complete network is achieved. The process (which is a little more complex than the outline suggests) may be optimised not only for overall length but for number of nodes, degree of link redundancy (e.g. tree, loop or network), resistance to interference, and so on. As before, genes code for these basic characteristics and the balanced optimal solution is evolved in the simulated hostile environment.

CONCLUSIONS

The introduction of advanced architectures based on common module and high speed data bus concepts should have major benefits, particularly with regard to:

- Efficient use of system processing resources
- Capacity for on-line reconfiguration
- Maintenance costs.

However, avionic architects will need a powerful new design methodology if all potential benefits are to be realised in practice.

At a system theoretical level, an architecture comprises:

- Functional entities
- Clusters of entities
- Links within and between clusters.

Functional decomposition followed by iterative use of clustering and linking strategies and cost-effectiveness analysis provides the basis for a methodology. However, several practical difficulties must be overcome to handle the diversity of architectural options and the many conflicting goals and constraints.

A potentially promising approach to solving some of these problems is the application of cumulative selection principles.

REFERENCES

1. J C Ostgaard and D R Morgan, "PAVE PILLAR and PAVE PACE: Avionics System Architecture for the 21st Century"

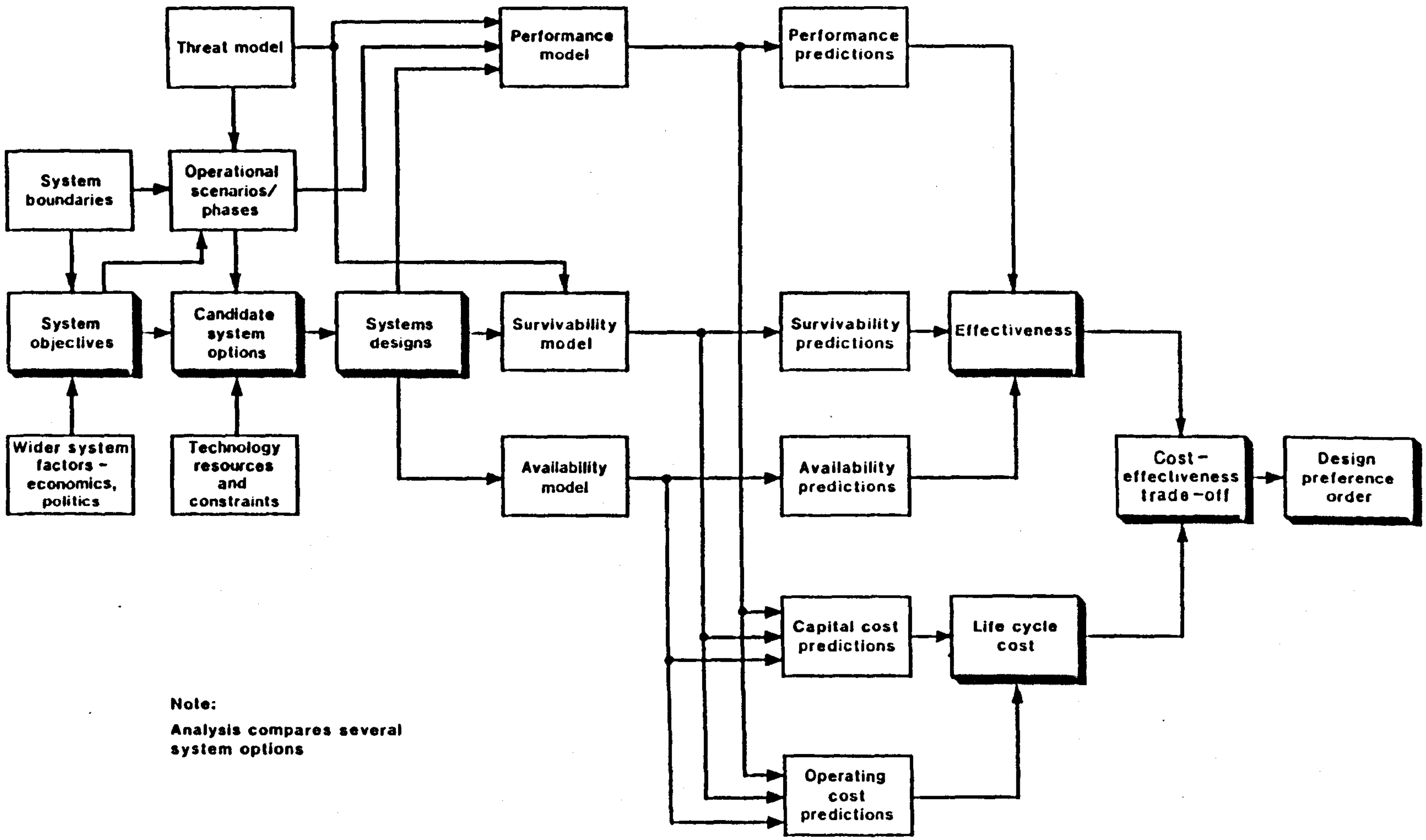
2. J Bishop, "Overview of Advanced Architectures - The User's Viewpoint"

3. A B Close, "A3 Implementation Dimensions"

Refs. 1 to 3 are published in the seminar volume Military Avionics Architecture for Today and Tomorrow (ERA Technology, London, 1988)

4. D K Hitchins, "BM/C^3 Design by Cumulative Selection", (To be published in the Proceedings of the IEE 3rd International Conference on Command, Control, Communications and Management Information Systems, May 1989)

5. R Dawkins, The Blind Watchmaker (Longman, London, 1986).



Note:
Analysis compares several system options

Fig. 1 COST-EFFECTIVENESS ANALYSIS

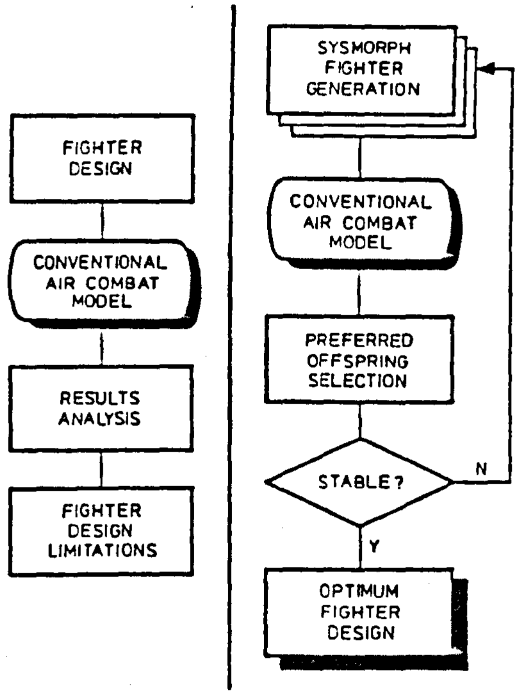


Fig. 2 NEW MODELS FROM OLD?

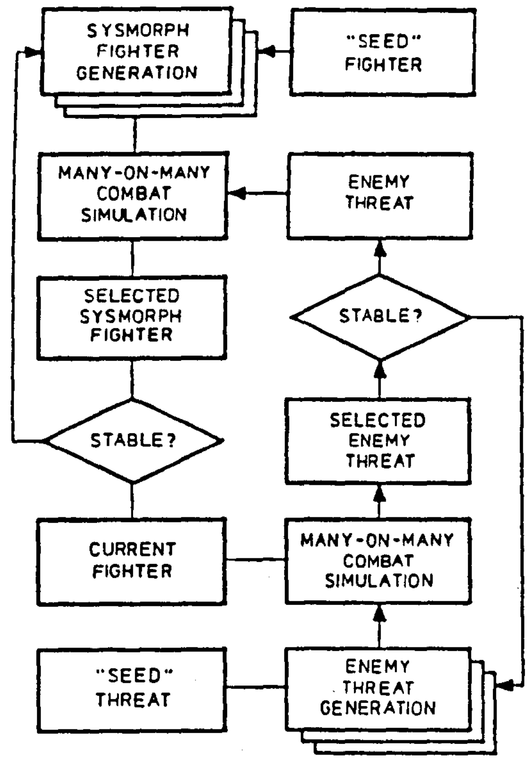


Fig. 3 SYSTOLIC CUMULATIVE SELECTION OF FIGHTER-THREAT PAIR

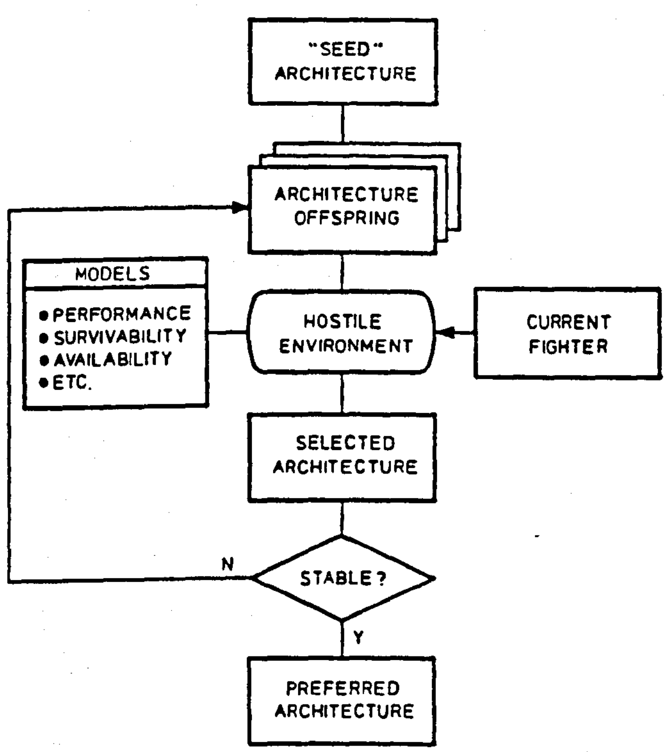


Fig. 4 CUMULATIVE SELECTION OF ARCHITECTURE

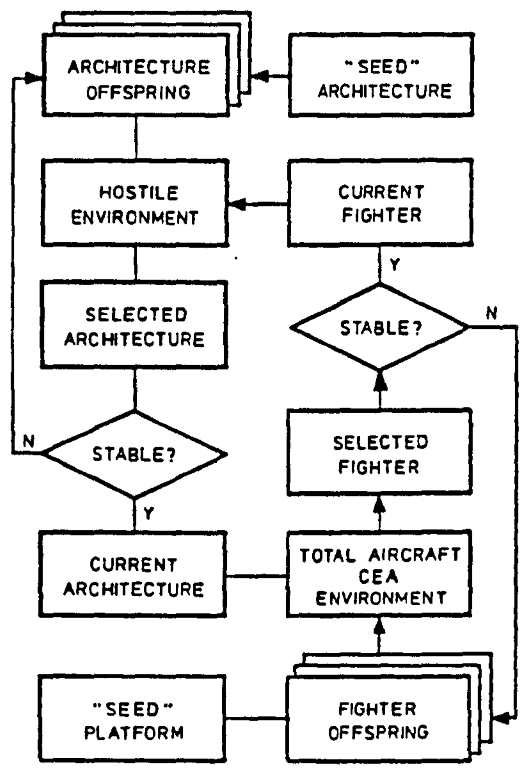


Fig. 5 SYSTOLIC CUMULATIVE SELECTION OF ARCHITECTURE-FIGHTER PAIR

BM/C^3 DESIGN BY CUMULATIVE SELECTION

PROFESSOR DEREK K HITCHINS

CITY UNIVERSITY - LONDON

INTRODUCTION

Command and Control (C^2) systems are becoming more complex, with more processing elements, modules and infrastructure all needing to operate in reducing time as enemy performance improves. Our C^2 systems are consequently taking longer to design and develop, during which interval threat, technologies, user procedures and policies will inevitably change.

Present concentration on ever-more detailed requirements capture can be seen as simultaneously essential to control the complexity, but futile in its end result of delivering a carefully outmoded capability. Systems are arriving in service only to commence more-or-less immediate update, trying to evolve in pursuit of an accelerating threat.

Design, development and in service processes are strained; our techniques for accommodating these difficulties are presently inadequate. CAD/CAE/CAM has concentrated on hardware and software and there is no real system CAD, where system might refer to a complete aircraft, ship, or BM/C3 set-up. In the absence of proper tools, fundamental design issues can be taken "on the nod". Should the C2 HQ be hard and static or soft and mobile? Should the new fighter weigh 9 tonnes or 10 tonnes? Should the communications network have four switching centres or five? And should all of these questions be treated in isolation or should they be intimately related?

Perhaps we can approach complex system design in a different way. Nature does not seem to share our difficulties. After all, the most complex organism on the planet is ineffective for some ten years after delivery - we expect that to be so and nurture new arrivals accordingly as they evolve socially to meet an environment markedly different from that to which nature originally responded. The new arrival is provided with the <u>ability</u> to evolve socially and not, as with lower order animals, imprinted with immutable social behaviour patterns.

Present approaches to BM/C3 design tend to adopt the "lower animal" approach by examining the way current systems and organisations operate and then encapsulating procedures, rules and methods into structure and software. Perhaps we can learn to adopt the higher animal approach and produce systems which are both imprinted <u>and</u> able to evolve in service.

The aim of this paper is to introduce and explore the application of cumulative selection as a promising and exciting new way to approach the creation, design, development, management and operation of complex BM/C^3 systems.

CUMULATIVE SELECTION

In his book, The Blind Watchmaker (1), which was shown as part of the Horizon TV series under the same title, Dr Richard Dawkins, an evolutionary zoologist at Cambridge University, developed the concept of cumulative selection. Dawkins was concerned with the argument between the Darwinian Evolutionists and the Creationists, with the latter propounding that the creatures of the world were all created at time zero and, while some might have died out since, none had evolved. Creationist argue persuasively that brains and eyes, for example, are far too complex to have resulted from sheer chance; the odds against are truly astronomical. Richard Dawkins showed that such awesome complexity could come about quite simply by cumulative selection.

Cumulative selection arises because progeny inherit characteristics from both parents and are imperfect copies of either. Some imperfections adapt offspring to their (changing) environment either by enhancing their performance or their survivability. While the generation-to-generation changes might be quite small, their accumulation over twenty or thirty generations can be spectacular.

Richard Dawkins developed a computer analogy to cumulative selection which he called BIOMORPH in which he could draw on a graphic screen an "organism" using a recipe into which he stirred analogues of genes to represent colour, size, numbers of leg branches, branching angles, etc.

One parent BIOMORPH bred a set of offspring with, in each case, one and only one "gene" nudged either up a notch or down a notch. The operator sits in front of the screen, selects the offspring with the characteristic he wishes to develop and that offspring then breeds its set of progeny in turn. A truly incredible variety of shapes and patterns can be produced in this way, but - as Richard Dawkins noted - this is not what Nature does.

Nature is blind to the future - it has no target design that it wishes to develop. Instead, offspring survive and replicate according to their ability in the evolving environment. The process of change is cumulative selection, the drive for change is survival in a hostile environment. Each successive generation survives and breeds because it has features - marginal differences, perhaps - which improve the probability that it will replicate. A 1% improvement in camouflage may sound insignificant, but it will make the creature a little less detectable, a little more likely to survive and replicate. Successive generations are likely to sustain a greater proportion of adapted offspring,. and so on. This process is called, of course, Natural

Selection.

A simple prototype design tool, called SYSMORPH in deference to Richard Dawkins' precursor, has been developed to prove the principal of design by cumulative selection and to explore its value, limitations and basis for further exploration. A further tool is in development, called COMMORPH, which seeks to apply the technique experimentally in communication network designs, with the same intent. Following paragraphs will describe both these tools.

SYSMORPH

SYSMORPH Gene Selection

The choice of "genes" for SYSMORPH is far from obvious. Genes do not act like a blueprint, in the sense that obtaining half the blueprint would describe half the structure. Instead, genes permeate throughout every part of the design, resulting in quite different features in different situations. A colour gene for example may affect colour over a whole range of internal and external body features. A size gene may differentially affect different parts of the body's limbs. And so on. Dawkins compares the genes with parts of a recipe - each instruction may be evident in some or all parts of the result.

The concept arises of applying the process of cumulative selection in the design of BM/C3 systems. There are two stages in such an approach;

* Develop a parallel to Dawkins' BIOMORPH to enable design by cumulative selection under operator guidance - I call this "Selective Breeding"

* Wrap a hostile environment about the cumulative selection process such that favoured offspring are selected by their fitness, so eliminating the operator - I call this "Natural Selection"

After considerable trial and error, five genes were selected for the prototype SYSMORPH. The number was a reconcilliation between simplicity, orthogonality and the ability to both manually cross-check the calculations and to display the results on a graphic screen. The chosen genes were:-

* Aggression, the degree of overt hostility

* Stealth or Discretion, the desire to remain undetected

* Co-operation, the desire to act in concert with others

* Shell, the self protection gene to represent the degree of hardening or armouring

* Delegation, or the degree of node replication in the internal system strucures

The objective of orthogonality between genes is important; without orthogonality it would be possible to over-emphasise attributes resulting from cumulative selection by accidentally accounting for the same factor twice over.

The five genes taken together provide a robust set, adequate for the prototype but clearly short for a full-blown system. It is none-the-less surprising how much can be achieved with only five genes, suggesting perhaps that cumulative selection may not require a vast number of source genes to be effective. In view of the potential combinatorial explosion inherent in the technique, this is a suggestion to be welcomed.

SYSMORPH - Selective Breeding

Reproducing Dawkins' work was fairly straightforward once the genes had been selected. The process is based on a standard drawing routine which determines the basic shape of a SYSMORPH -see Fig 1 - with the following features:-

* An external structure, analogous to a Shell or Carapace.

* A set of weapons to be used for the achievement of the SYSMORPH's mission; these overt weapons reveal the presence of the platform to hostile elements in its environment and symbollically pierce the shell so that they are visible from outside the shell in consequence.

* A set of sensors to sense both the external environment and the threats it may contain. Since sensors may be either active (self-revealing in use) or passive (covert), active sensors are shown piercing the Shell while passive sensors rest underneath , inside the Shell and symbolically hidden from observation.

* A set of communication facilities which, in use, reveal their presence to others in the environment and which also pierce the shell.

* An internal structure represented by the three contained boxes. These boxes vary both in individual size and in aggregate area to indicate the degree of autonomy and the capacity of the internal processing and communications architecture.

It is straightforward, using the stylised drawing interpreter -which must have its counterpart in nature - to draw the shape of Fig 1, which is referred to as a SYSMORPH. Since a SYSMORPH is constructed uniquely from five genes, it is possible to generate 10 offspring SYSMORPHs with, in each case, only one gene either increased or decreased per offspring. Figure 2 shows such a generation, with the parent SYSMORPH in the centre and the offspring SYSMORPHs surrounding it marked A to J. An operator sitting at the SYSMORPH tool can identify any characteristic he wishes to emphasise from among the various offspring, select the appropriate button, A to J, and that offspring will appear in the centre to be followed swiftly by its own offspring. And so on. The tool is simple to learn and fun to use but its value is limited, since essentially it is just a rather smart drawing tool, responding to the operators chosen target design.

SYSMORPH - Natural Selection

Natural Selection presents each of the SYSMORPH offspring with a hostile environment in which it has both to perform and survive. Each SYSMORPH has different gene values which respectively promote its abilities. However performance, survivability, etc, are each affected by more than one gene. Performance for example is related to aggression, to mobility, to co-operation, and so on. So a complex pattern emerges.

The complexity is resolved by scoring each design driver (e.g. Interoperability, Affordability,etc.) in turn, weighting the driver score and then aggregating the weighted scores to give an overall figure of merit for each SYSMORPH. The SYSMORPH with the highest score reproduces.

The process of setting up all these gene values, scoring arrays and weighting vectors is the task of EXECSYS, one of the programs in the SYSMORPH suite. The structure of EXECSYS, together with typical formats, is shown in Figure 3. The operator is presented with a model menu, presently containing:-

* Fighter Aircraft

* Tank

* Frigate

* C2 HQ

* Android

On selecting the required model he is presented with two arrays, with sample figures already inserted, which he may change. The first array defines the limits between which gene values may move, and the starting conditions of the original parent SYSMORPH. As the prototype is used to experiment, three start points have been incorporated so that the original parent may have few, median or many genes, the purpose being to confirm that the final outcome is the same no matter what the start point.

The second array, the gene contribution matrix, defines the interaction between each gene and its environment, represented by the design drivers shown in the table. The matrix represents the inputs that, in a full system, would be derived from comprehensive simulations of combat, survivability, etc.

Natural Selection - Analysis Methods

Three methods are used to aggregate the weighted scores for each SYSMORPH. The reasons for using several methods are twofold:-

* There is no firm concencus on a correct method

* Having three methods allows sensitivities to be analysed

The first method is simple algebraic addition. To enable this, all cost figures are negative since their sum, by whatever means, militates against system conception and development. The second method is similar, but the weighted scores are summed algebraically and then divided by the affordability sum to give an effectiveness: cost ratio. The third, and least suspect method, is to sum all weighted scores using root mean square and to divide by the modulus of Affordability. Using RMS is less discomforting because it is a reasoned way of aggregating mutually-orthogonal parameters which, to a significant degree, the design drivers should be. This process, together with a coherent weighting strategy, are subjects of ongoing research.

Natural Selection - Affordability

Affordability is a key issue in SYSMORPH, since it balances the tendencies of other drivers to increase complexity, weight, etc. Cost is accrued by each SYSMORPH not only by the addition of weapons, sensors, communications, processing and armour, but also by the increasing size of the platform/vehicle/HQ needed to accommodate the load, and by the increasing complexity of interfaces and infrastructure.

The prototype SYSMORPH employs simplified cost models. For example, correlation can be shown between the total number of missiles and guns (WPNS) carried by a non-delta fighter and its All Up Weight (AUW):-

$$AUW(kg) = 4632 * (WPNS)^{0.7645}$$

There is also a correlation between AUW and fighter cost which allows cost to be related to total number of weapons in the parametric cost model.

The cost of systems engineering can be assessed broadly by noting that it increases with the number of interfaces (I), and that they increase as:-

$$I = k * P(P-1)$$

where k is a proportionality constant
P is the number of connected peripherals in the system

For a fighter, then, the overall cost for each offspring can be formulated by summing the cost due to weight, the cost due to weapons and the cost due to systems engineering. Absolute costs are generally unnecessary, since the program is seeking to find the offspring with the highest effectiveness: cost ratio; so long as relative costs are properly maintained, the correct selection will result.

SYSMORPH - the Program

Figure 4 shows the main SYSMORPH program structure. At the centre are the features common to both modes, Selective Breeding and Natural Selection. Features to the right of centre are particular to the more complex Natural Selection process. The figure, which is a simplified version of the full prototype, is self explanatory.

SYSMORPH in Operation

Testing SYSMORPH is an interesting concept. When I had built the program, I proceeded to test it by inserting gene start-points and hostile environment conditions that were well understood and to which I knew the answer - a simple test that SYSMORPH would provide the obvious answer to the obvious question. It did not. It certainly produced the eventual design characteristics which I expected, but the sequence of evolutionary steps which it took were not what I had expected.

For example, see Fig 5. The figure is a typical printout showing the start point and the sequence of evolutionary steps that SYSMORPH followed in deriving a high-performance fighter interceptor. One can argue that, by Iteration 15, it has produced a reasonable description, given the initial conditions, which started the aircraft with 4 missiles, mainly passive sensors, little communications, medium armouring and partly distributed processing. By the end it has two missiles, only one passive sensor (which the program interprets as also meaning 5 active sensors), a plethora of communications, minimum armour and fully distibuted processing.It is not the end point that surprises, but that the first step chosen by SYSMORPH was to increase communications/ cooperation.

The reason for surprise is that the start conditions and, particularly, the gene contributions, are an expression of the operator's (my) prejudices. In the absence of full combat models, I selected and weighted according to my experience. So, the program should have matched my intuition (which was to remove armour and to add active sensors first). The program, which I wrote, initialised and supplied with data, was disagreeing with me.

The program was correct. On checking, there were no errors. What SYSMORPH was achieving was to balance all the factors which I had supplied more effectively than I was able to do mentally. SYSMORPH, even as an elementary prototype, seemed to be providing a stimulus to lateral and creative thought.

Further testing of the prototype showed this deduction to be the case. SYSMORPH is good at taking a wide variety of different factors, each with some predictable impact on eventual design, and balancing their contributions objectively. The eventual design parameters produced by prototype SYSMORPH are sometimes less interesting than the route by which it reaches the end point, in the sense that the final outcome is often obvious from the bias introduced by the operator. So, conceptually we could introduce the parent genes of a current design for, say, a frigate and SYSMORPH (or at least a full-blown version) would tell us the first, most important upgrade we should introduce to enhance its performance/survivability/cost effectiveness, then the second, and so on.

SYSMORPH - Limitations and Promise

The present prototype is very limited - it was intended only to establish a principle. It has done so, rather unexpectedly well. But with only five genes, and without full simulation models representing threats and environments, its results are only a guide to future value. The principle does appear to be incredibly powerful, however, and the technique holds promise of revolutionising our approach not only to design, but to development and operations, as later paragraphs will show.

SYSMORPH Potential

The SYSMORPH idea can be applied in some highly creative ways in conjunction with conventional models. Consider the process of assessing a fighter aircraft's performance - see left half of Figure 6. Conventionally we develop, perhaps over many years, rather good many-on-many (MOM) combat models and we then fly simulated new aircraft designs in the model, to assess likely performance. The results are analysed, limitations identified, and perhaps the fighter model is revised off-line.

Replace the fighter model with a SYSMORPH program, wrapped around the conventional simulation as shown in the right of Figure 6. A complete generation of slightly-differing SYSMORPH fighters is produced; each is flown in turn agains the threat, in an MOM simulation, and the best offspring is selected to reproduce. Cumulative selection will evolve the ideal fighter within the limits of the model representations. To be sure, there is a tremendous use of the existing, conventional model - but that is good, since it is available and trusted as a reference.

But that is only the tip of the iceberg! Consider Figure 7, the left half of which repeats Figure 6 in condensed form. The response by an enemy to a better fighter will be to produce a better counter-threat - that is the way of the world. So, having evolved our SYSMORPH fighter, what is to prevent us evolving the counter-threat? Having evolved the threat to see what our SYSMORPH fighter, presently only on the drawing board, may face when it matures, we can then further evolve our SYSMORPH fighter to accommodate this new environment. I call this Systolic Cumulative Selection since it operates like a heart, beating between the two parts of the model.

Evolving a fighter in the way just described is far from the ultimate approach. It is axiomatic that a system comprised of separately optimised parts will not itself be optimal. While evolving a fighter might be very attractive, it pales beside the prospect of optimising, for example, a complete Air Defence System, comprising interceptors, missiles, surveillance, identification, Command and Control, communications, etc. Yet that is the challenge which we should address since it will undoubtedly prove much more cost-effective than our present piece-meal approach. Can cumulative selection help?

Probably, is the prudent answer. Not because the technique will not work - it will. The question is more one of computability, since the number of genes or rules needed to describe such a complex system in adequate detail may be prohibitive when it comes to the computational task. Maintaining the level of abstraction is the key. And such presently intractable tasks as simultaneously optimising for cost, performance and architectural survivability of a C^2 system do seem to be practical targets.

COMMUNICATIONS SYSTEM MORPHOLOGY - COMMORPH

A classic problem facing wide-area communication system designers and tactical planners concerns the optimum siting and numbers of switching nodes. Microwave links are expensive and it is essential therefore to limit the number of "shots" (links) and switching centres, consistent always with performance. Unfortunately, the problem can prove mathematically intractable for all but the simplest of configurations of source/sinks (the C2 centres).

One analogue solution (literally) to the optimisation problem invokes the phenomenon of surface tension. In this approach, pioneered some years ago at Southampton University, two sheets of perspex are drilled with holes corresponding in plan to the geographic sites of the C2 centres. The two sheets are then separated by inch-long studs screwed through the drill holes, to form a sandwich, and the whole is dipped in soap solution. On withdrawal, the soap solution forms into planar bubbles which connect the vertical studs by the overall shortest length. The technique is both ingenious and inconvenient.

COMMORPH tackles the problem in an analogous manner. Consider Figure 8. It comprises ten source/sinks, C2 HQ in this instance. Their locations are determined, and it is desired to interconnect them by radio links. Eight switching/repeater nodes are scattered randomly into the area of the source/sinks. The COMMORPH core routine then manoeuvres the nodes gradually into an optimal configuration which might be a tree, a loop, or a fully-connected network according to operator pre-selection. The tree result is shown in Figure 9; the solution, once seen, is obvious of course. But notice how some nodes space themselves evenly into straight lines (relays) while others form junctions (switches).

The COMMORPH core process is essentially very straightforward. Each node is allowed to associate with the nearest source/sink; the nodes are then interconnected into a tree/loop/network configuration and the overall length of all the links is recorded. Next, each node is stepped in eight directions, each at 45 degrees from the rest-point. At each step the node re-associates with the (then) nearest source/sink and the overall link length associated with that step is recorded. Eight overall link lengths result; the shortest is selected and the node is moved to that new position. The process is repeated for each node in sequence, until all nodes have been stepped, and the complete cycle is repeated until successive iterations show no reduction in overall link length. The process is simple cumulative selection.

Hypnotic though the process described above appears on the graphic screen, it is only the core of COMMORPH. The original number of nodes was chosen arbitrarily, and what has been described was simply the development of the parent COMMORPH. Offspring are generated from this parent next, with differing numbers of nodes and the offspring which produces the minimum overall link length is evolved, again by cumulative selection.

Nodal step-length is a crucial parameter. Were it too small, not only would the program take inordinately long to run, but sub-optimal sticking points might be encountered. (These may prove interesting in the communications context and are the subject of present investigation). Step length is therefore varied automatically under program control, starting large and reducing with each cycle until an optimum appears to have been found. Step length is then increase to see if the configuration switches to a new mode. (This is analogous to shaking the bubble formed between the perspex plates, so giving it sufficient film/lattice energy to switch to the lowest energy mode).

At this point we have simply emulated the soap bubble method. With the new approach it is possible to introduce a variety of factors in addition to minimum path length. In a manner similar to SYSMORPH, the network may be created from genes which code for node characteristics such as power, sensitivity, capacity, damage tolerance, camouflage, mobility and even mast height. A hostile environment might include terrain, electronic warfare, physical warfare, interoperability, security, affordability, and so on. The performance measure would generally be signal to noise ratio, which would account automatically for line-of-sight limitations imposed by terrain screening.

COMMORPH Applications

The applications to which such a flexible tool might be put range from circuit board design to satellite communications and are limited only by imagination. Two examples only will be presented:-

Mission Route Planning. When planning - or when actually flying - a route through hostile territory, there are many factors to consider. Takeoff, destination and return base will be known, and there should be good knowledge of terrain if not of enemy en-route defences. The problem is to derive a 3-D route plan to the target which minimises risk from being seen or attacked, which reaches the target, and which recovers with equally minimal risk. With COMMORPH, cumulative selection will find the route, either completely on its own or from some preferred start-point inserted by the operator. A full-blooded capability would need to be coupled with good knowledge of the enemy's defensive lethalities, and of the attacking aircraft's characteristics so that "balooning" over a projection was avoided, for instance. Lack of definition of enemy characteristics such as lethality ranges, for example, can be accommodated by employing systolic cumulative selection and evolving the enemy's defences within sensible limits and then returning to the original mode to fine-tune the flight path in the light of the evolved threat. There is potential for a real-time airborne planner here.

Naval ASW Screening. A sonar contact may require a naval battle group to reform into a defensive configuration, or screen, in response to the presumed enemy position. The rules for forming such a screen can be extremely complex in execution; ships could have been anywhere at time of contact, and could be involved in a variety of essential activities, ranging from picket to helicopter operations to defending against air, surface or previous sub-surface attack. A COMMORPH-type of program will employ cumulative selection to determine the best means of establishing the optimum screen from the present starting point taking all of the relevant factors into account. Again, systolic cumulative selection can be employed to anticipate enemy counter moves to some degree.

EVOLVING RULES

Rules in a program can be thought of in much the same way as genes; if we can evolve a gene-based SYSMORPH/BIOMORPH can we evolve rules? It seems likely that we can. Looking back at Figure 7, Systolic Cumulative Selection, might give a clue. The evolving SYSMORPH aircraft was indeed adapting to a changing threat, which is no different to responding to changing rules and is one measure of intelligence. It seems possible that, while cumulative selection has resulted in our own natural intelligence, it might also form a basis for us to develop artificial intelligence.

The idea of evolving rules is seductive. Consider the applications. Conflict is an obvious one, but what about, say, project management rules as set in a development logic/bar chart. PERT and Critical Path Analysis would be no match for a replanning process based on cumulative selection. Research in the US at present is evolving manufacturing robot software rules by cumulative selection so that the robots no longer lock-up when faced with an unforeseen situation, but instead adapt to the new circumstance, within prescribed limits of freedom. Cumulative selection is being used to design aerofoils, low-drag car profiles and turbines in Germany and these advances may well generate new rules too.

Using in-built cumulative selection would enable us to deliver BM/C3 systems with the ability to evolve from delivery, using perfectly conventional, well-established rules and procedures as the initial provision to the user/customer. So, the system would probably be dated at the time of delivery, but the new users would provide contemporary threat and situation data to the system, which would adapt the delivered capability to the new circumstances. The concept is difficult, of course, and would need to be carefully controlled, but present performance is not sufficiently satisfactory that a new and promising approach should be overlooked.

SUMMARY AND CONCLUSION

The paper started by considering prospects for planning, designing and developing complex BM/C3 systems, which are presently often disappointing in their realisation, having been designed and developed over a protracted period and proving to be outdated and inflexible in use. The lack of creative, adaptive system tools which might reduce the impact of this shortfall was observed, where system referred to the major elements in a complex BM/C3 system.

The process of cumulative selection was introduced, and its employment in SYSMORPH and COMMORPH described, both being primitive prototype design tools. The approach to, and limitations of, design by cumulative selection were assessed. It can be concluded that the approach has potential for encouraging insight, by allowing tradeoffs between a wide and incompatible set of design influences at a time during the creative process when such a need is most keenly felt. In effect, the prototype SYSMORPH could be seen as a lateral thinker, while a full-blown version could be seen as a valuable design tool. Indeed, systolic cumulative selection may be a basis for developing artificial intelligence, much as Nature's cumulative selection has produced natural intelligence.

Applications for SYSMORPH and COMMORPH included not only design, but also mission route planning, defensive screen planning, system architecture design and perhaps even the evolution of a complete defence system, subject to computability.

Finally the paper looked at the potential application of cumulative selection as a means of evolving rules intelligently so that systems may adapt their performance in response to changing environment, with the conclusion that the achievement of such capability looked feasible, was difficult but stood out as an exciting, unchallenged prospect. Application of rule evolving systems included project management systems to develop complex BM/C3 systems as well as within the systems themselves for evolving procedures, tactics and the prediction of outcomes.

REFERENCES

1."The Blind Watchmaker" by Richard Dawkins, Longman ISBN 0582 446949

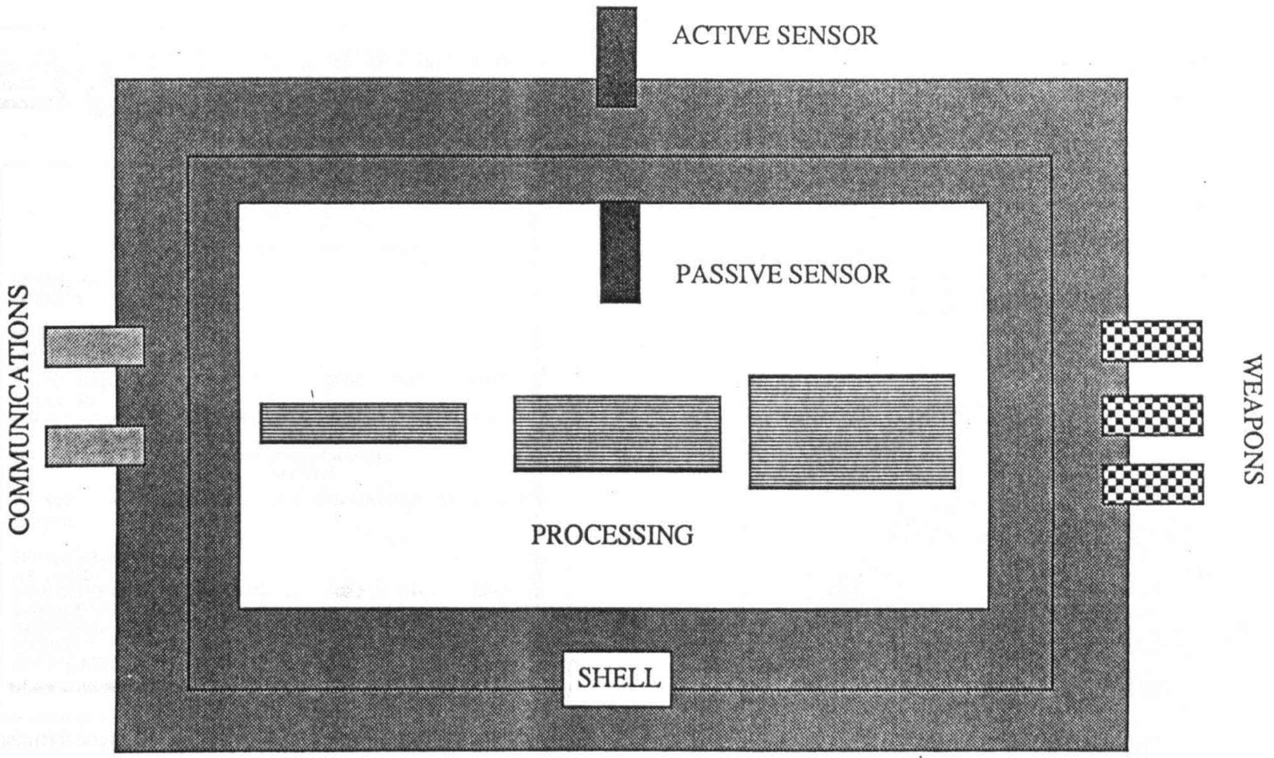


Figure1 - PARENT SYSMORPH

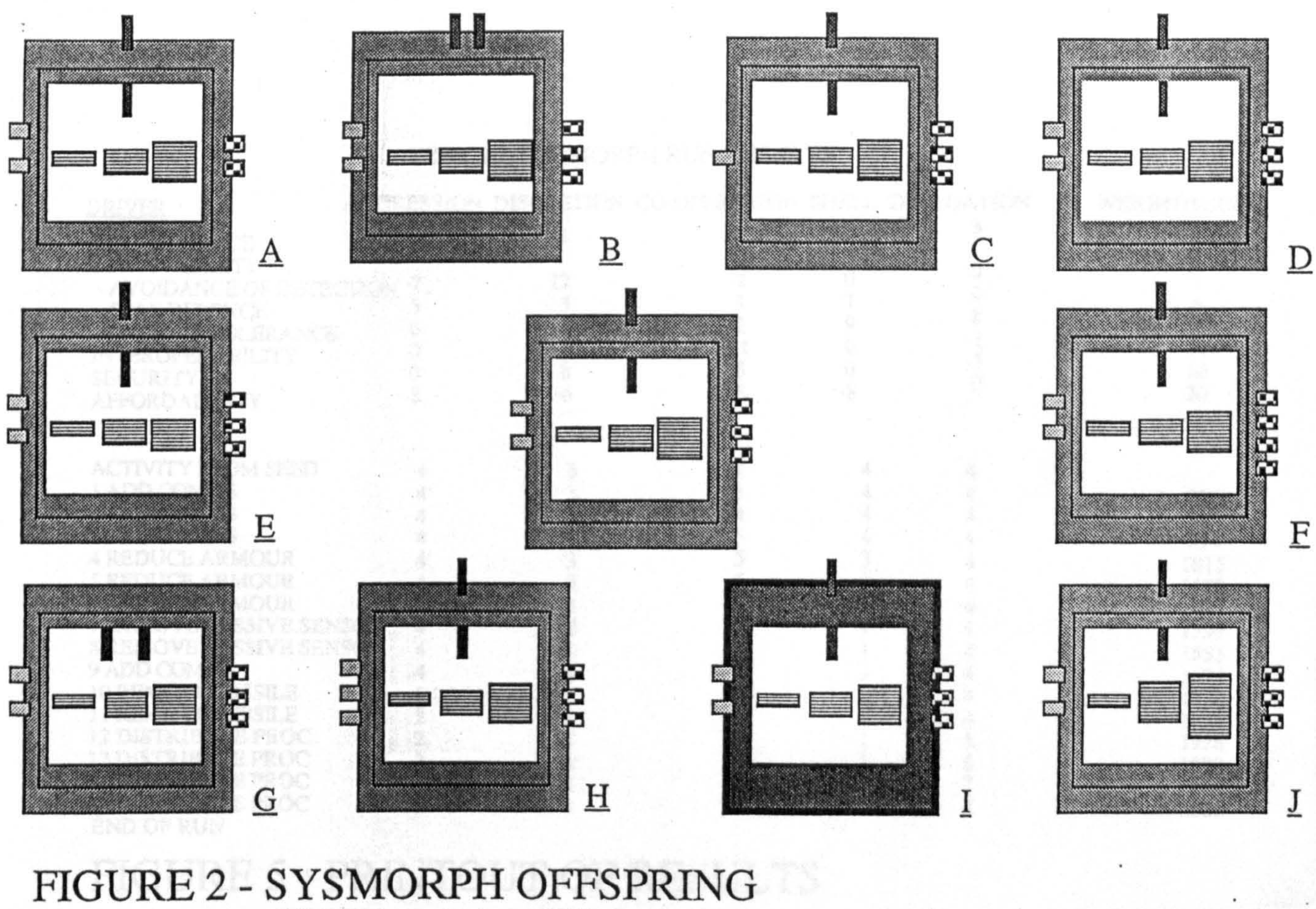


FIGURE 2 - SYSMORPH OFFSPRING

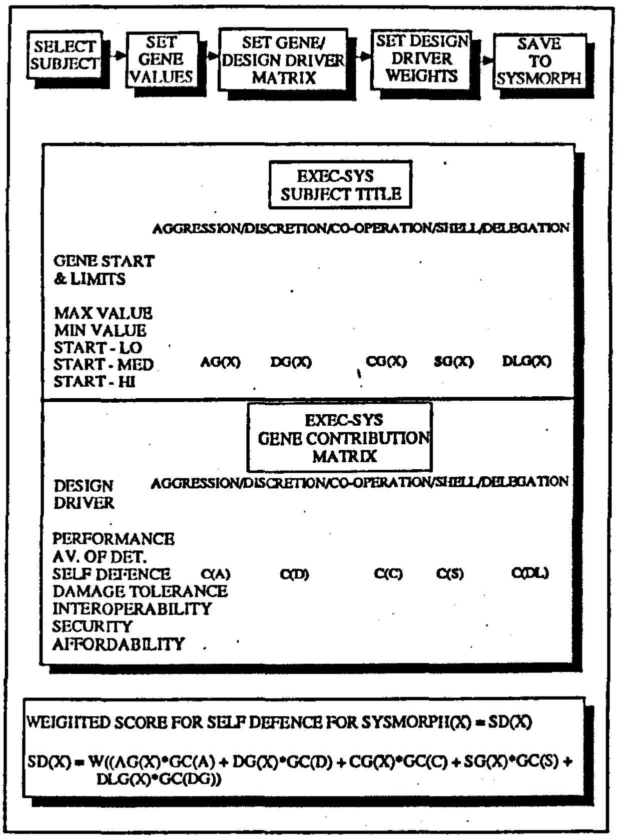


FIGURE 3: EXECSYS CONTROL DATA

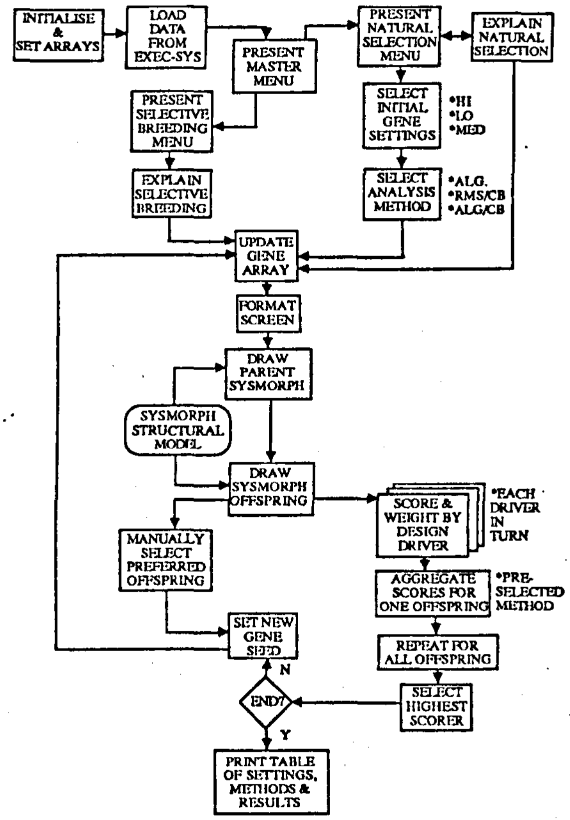


FIGURE 4: SYSMORPH PROGRAM STRUCTURE

TYPICAL SYSMORPH RUN RESULTS

DRIVER	AGGRESSION	DISCRETION	CO-OPERATION	SHELL	DELEGATION	WEIGHT/SCORE
PERFORMANCE	12	-4	10	-5	3	30
SURVIVABILITY						
- AVOIDANCE OF DETECTION	-2	12	-2	0	0	7
- SELF DEFENCE	5	3	3	7	0	6
- DAMAGE TOLERANCE	0	-3	-2	9	8	7
INTEROPERABILITY	0	-5	11	0	-1	20
SECURITY	0	8	-15	0	-3	10
AFFORDABILITY	-8	-6	-5	-8	0	20
ACTIVITY FROM SEED	4	3	2	4	4	
1 ADD COMMS	4	3	3	4	4	714
2 ADD COMMS	4	3	4	4	4	815
3 ADD COMMS	4	3	5	4	4	914
4 REDUCE ARMOUR	4	3	5	3	4	1015
5 REDUCE ARMOUR	4	3	5	2	4	1130
6 REDUCE ARMOUR	4	3	5	1	4	1263
7 REMOVE PASSIVE SENSOR	4	2	5	1	4	1399
8 REMOVE PASSIVE SENSOR	4	1	5	1	4	1553
9 ADD COMMS	4	1	6	1	4	1654
10 REMOVE MISSILE	3	1	6	1	4	1739
11 REMOVE MISSILE	2	1	6	1	4	1874
12 DISTRIBUTE PROC.	2	1	6	1	5	1926
13 DISTRIBUTE PROC	2	1	6	1	6	1980
14 DISTRIBUTE PROC	2	1	6	1	7	2034
15 DISTRIBUTE PROC	2	1	6	1	8	2089
END OF RUN						

FIGURE 5 : PRINTOUT OF RESULTS

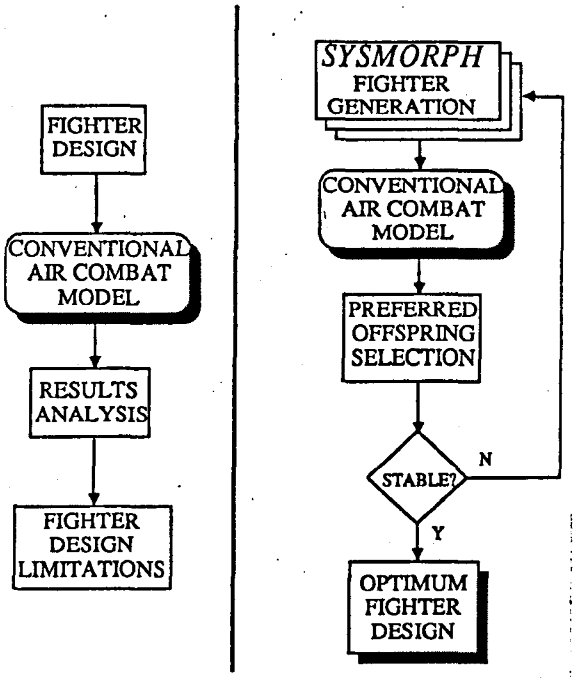


FIGURE 6: NEW MODELS FROM OLD?

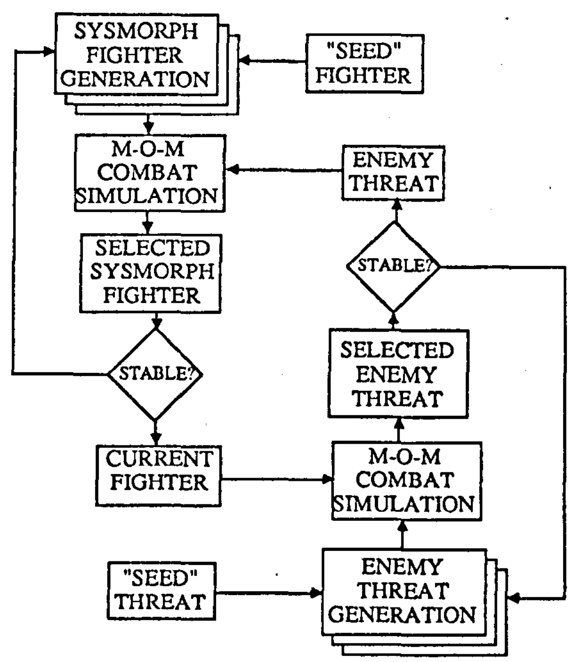


FIGURE 7: SYSTOLIC CUMULATIVE SELECTION

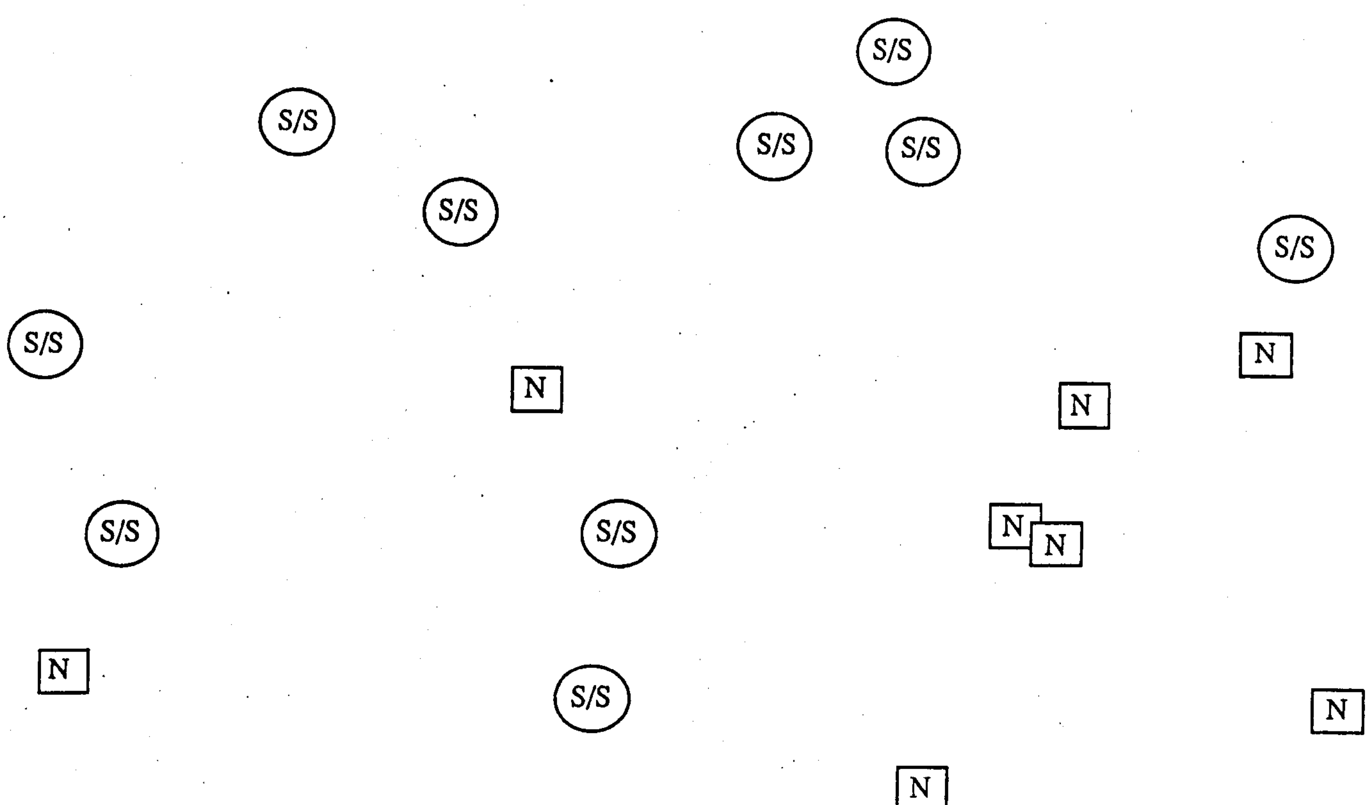


FIGURE 8: COMMORPH START-POINT

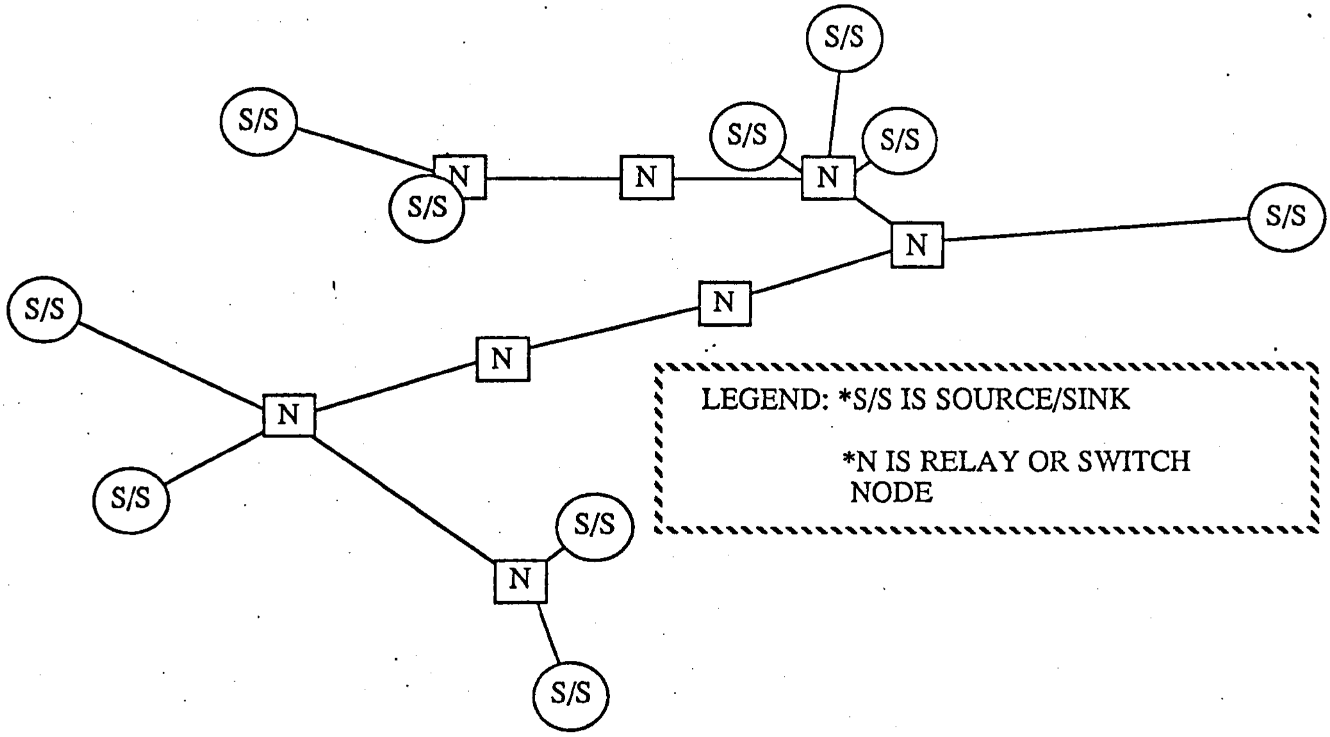


FIGURE 9: EVOLVING COMMORPH

BUILDING FUTURE SYSTEMS

Professor Derek K HITCHINS, City University

1. INTRODUCTION

There is a problem with the way in which defence systems are being procured. The problem is not that insufficient care is being taken, nor that the various parties, specifications and standards are in any way deficient. Quite the contrary. It may be that the very diligence is contributing to the situation.

That the problem exists, however, can be seen by standing back and looking at any number of sizeable procurements for all three services. To find a program that is not late and over-budget is unusual. To find a system delivered into service that does not almost immediately enter a modification/update programme is equally unusual.

This paper seeks some of the underlying causes of procurement shortcomings and offers some innovative, even provocative remedies. Seen as a theme common to all systems, including human activity systems, is their architecture, structure or organisation. Seen as as a major hurdle to be overcome in designing good architectures is the involvement of the users as part of, perhaps as leaders of, architecture design teams. The paper seeks to validate the role of the so-called user-architect, and in so doing unveils a new methodology for creating innovative design concepts, the Seven-Step Continuum, which progresses from "soft" through "firm" to "hard" system concepts.

2. CAREFULLY-SPECIFIED OBSOLESCENCE

Figure 1 outlines the two predominant procurement strategies of recent times: at the top is cost-plus; at the bottom, fixed-price. Cost-plus recognises the imprecision in requirement understanding and capture, and engages in an evolving development process which all-too-often has proved to be prolonged and expensive. Although the system, when eventually delivered, may be satisfactory, the customer may have paid more than anticipated for development and production, and the relevant arm of the services may have had to accommodate operational deficiencies while waiting.

Fixed price procurement engages in a fixed timescale and cost development. To achieve this goal, a very tight specification is sought and strict control is imposed on changes during development. Systems are expected to be delivered to time and budget. As Figure 1 illustrates, the principal difference between the strategies is in the transference of risk and expenditure. In either event, shortfall is inevitable. Both approaches have to recognise that the operational and technological environments are changing during development, leaving fixed-price, fixed-specification solutions particularly vulnerable.

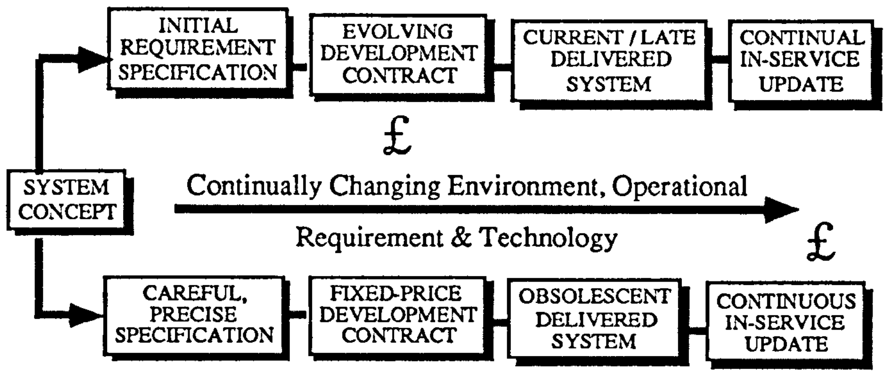


Fig 1: CAREFULLY - SPECIFIED OBSOLESCENCE

It is very probably impossible to specify any complex system with precision. To attempt accurate specification 3 - 5 years in advance of delivery is patent nonsense; this concept is, however, at the heart of present procurement strategy in an attempt to curb runaway costs.

3. ISSUES

The major procurement issues may thus be summarised as follows:-

* Procurement strategies seek to avoid late and costly systems but seem set to do little more than transfer the onus between purses of different procuring partners
* Precise specification may not be possible but the very attempt at precision is extending the timescales of fixed-price systems, invoking rigid systems and must eventually be self-defeating as it extends the development cycle.
* Knowledge elicitation is vicarious, imprecise and suspect - not to mention jargon-ridden.
* System architectures are becoming more complex, resulting in extending timescales and threatened project predictability.
* Punishing "late" fixed-price contractors is self-defeating; they will be unable / unwilling to provide in-service through-life support.

4. WHERE IS ALL THIS ARCHITECTURE COMING FROM?

The development problem seems to be getting worse, too. Figure 2 shows why. As we advance socially, we come to expect more of technology-based systems. International telephone systems are a classic example. Systems become both more capable in themselves and more interwoven. Defence systems in particular face increasing threat and reducing time-lines.

One thing has not changed - the human frailties of designers. They still possess only one person's mental perception and capacity, requiring them to breakdown a complex system into understandable parts. The size of an "understandable part" having remained substantially constant, it follows that more complex systems must have more such parts, more communications between the parts, more infrastructure and indeed more architecture.

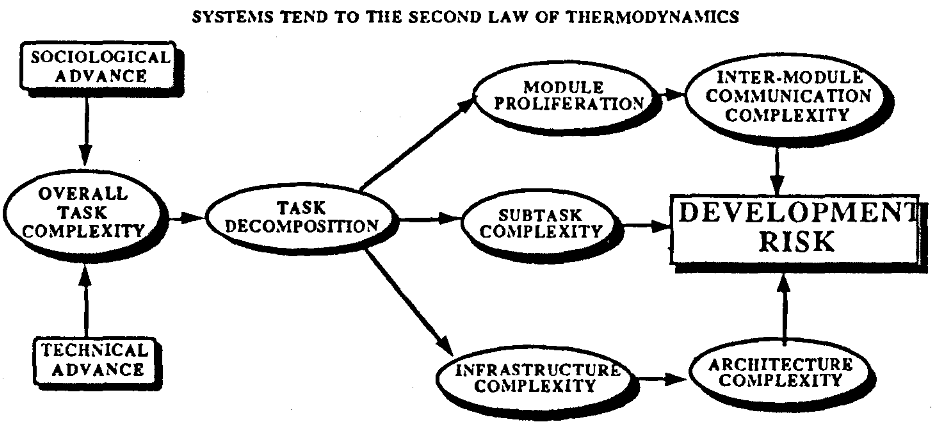


Fig 2: "ARCHITECTURE" GENERATION

Until as little as 10 years ago, the architecture of most Information-Decision-Action (IDA) Systems was a crew of operators, a central processor and star-connected sensors and action-elements (weapons). Such simple structures are vulnerable and of low performance, but essentially simple. Technology has enabled us to distribute, to avoid singularities and to co-operate to enhance performance. Technology has not supplied system-architects, however, nor have we been blessed with a science of system architectonics. Presently there is no discernable engineering framework for designing such complex systems which owe as little to software as human anatomy owes to brain neurons.

5. WHAT IS ARCHITECTURE?

To understand architecture it is helpful to return to basics. Figure 3 is basic. At the left hand side are a number of entities; they could be people, processors, avionic modules, ships, some mixture of all these, etc. There are no groups; there are no links; there is no architecture.

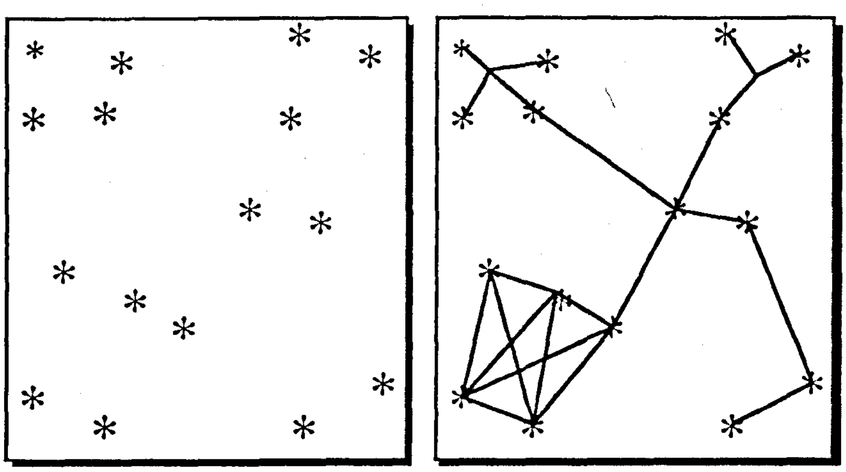


Fig 3: WHAT IS ARCHITECTURE ?

At the right, the same entities are clustered, infra-connected and inter-connected. Now there is architecture, evidenced by several features:-

* Clusters of Entities
* Clustering "drivers" which result in purposeful clustering of some entities and equally purposeful repulsion between others
* Connections within and between clusters
* Connectivity drivers which result in linking, link redundancy, protocols to facilitate linking and security to protect and isolate.

Architecture brings structure and order to chaos; it underpins all system design, no matter how simple or complex. A proper understanding of architecture is essential for would-be user-architects of whatever system.
NB A return to the left hand side of Figure 3 will find the reader unable to "unsee" the architecture of the right hand diagram; this mind-set is a particular pitfall for architects.

6. ARCHITECTURE ISOMORPHISMS

System architectonics is both an art and a science. As an art, perhaps it owes its roots to the human facility for pattern recognition. Figure 4 shows several disparate system types, which share common features. Using anatomical terminology, consider a manned-aircraft. It has:-

* A Brain - crew, giving fight or flight initiatives
* A Central Nervous System - processing, communications, etc
* Visual / Aural Sensors- radar, Infra Red, L3TV, ESM, eye-sight
* Energy reserves - fuel, accumulators, kinetic energy
* Energy Converters / Muscles - generators, controls, motors
* Pain Sensors - Built-in test, central warning panels, attention-getters, etc

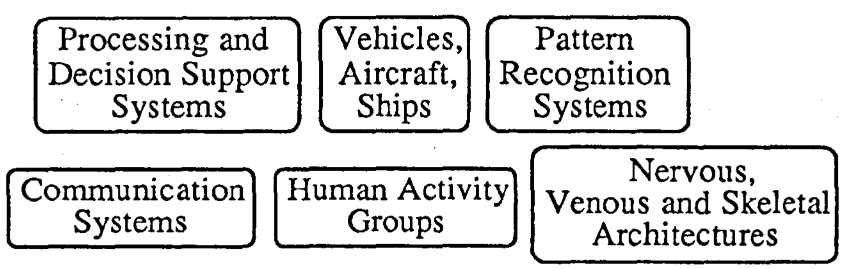


Fig 4: ARCHITECTURE ISOMORPHISMS

The crew themselves, as individuals, display the same anatomical features of course, but less obviously perhaps so does the fleet of aircraft engaged in some airborne exercise, with fighters under control of some AWACS aircraft. Here the architecture is spatially dynamic but perfectly evident, with central nervous system connections being completed by visual and radio links.

Architecture isomorphisms are often to be seen so, embedded within each other like so many

Russian dolls. Much can be learnt by an appreciation of human "architecture", much more sophisticated than any we presently seek to design. It is perhaps axiomatic that we design anthropomorphically - given the suggestion, we could do worse than learn from the best design on the planet.

7. BRIDGING THE USER-BUILDER DIVIDE

If we accept the concept of architectural isomorphisms at systems level, it follows that perception of isomorphisms, far from being an engineer's prerogative, is more likely by users and operators of IDA systems, owing to their everyday immersion in the relevant domain. At present there is something of a divide between users and system designers, due in part to their differing domains, but also to their differing training and particularly to a distinct lack of a common language. Users understand their objectives and aspirations; developers understand technology and its application.

Knowledge elicitation is an attempt to cross the divide, but what price two or three weeks elicitation against twenty / thirty years field experience? Expert systems can prove limited too. What value encapsulated rules for a new system elicited from experts who never went to war, or who did but using out-of-date equipments against a long-past threat?

The solution is evident:-

A. Teach users to be their own system architects
B. Develop user-architect tools to help them work at system level, injecting user domain understanding, doctrine and knowledge into a formal but friendly engineering framework
(C. Develop systems that evolve in service - but that is for another paper)

8. USER-ARCHITECT C.A.D.

User-architect tools are with us now in prototype, if we can recognise them - and they do not resemble software design tools. Figure 5 is an N2 (N-squared) chart, an incidence matrix showing the strength of relationship between eight entities, A - H. Were these entities to represent, say, the activities undertaken by Air Traffic Control at an airfield, or the different activities in a new RN frigate, operators would be ideally placed to understand and present the relationships between the activities and to prioritise those relationships.

As entered, the relationships merely record that some coherent numbering scheme has been used to express priority, that some relationships do not exist and that entity "C" has a relationship with all other entities and is in some way special. The entities can be clustered to associate more closely those entities having the strongest relationships. In the process, the interfaces create different patterns which are comprehensible to both man and processor.

9. CADRAT - CLUSTERING, LINKING & ENHANCING

Figure 6 shows the result. There are two tightly-bound blocks - all interfaces filled - which inter-relate strongly; these are the precursors to two physical architecture blocks, they exhibit purposeful clustering. There is also the previously-noted node, connecting two halves of the system of entities which is a candidate for hardening or replication in anticipation of failure or attack. Finally, "D" and "E" - the "waterfall" pair - represent hierarchical control and reporting, revealing authority lines.

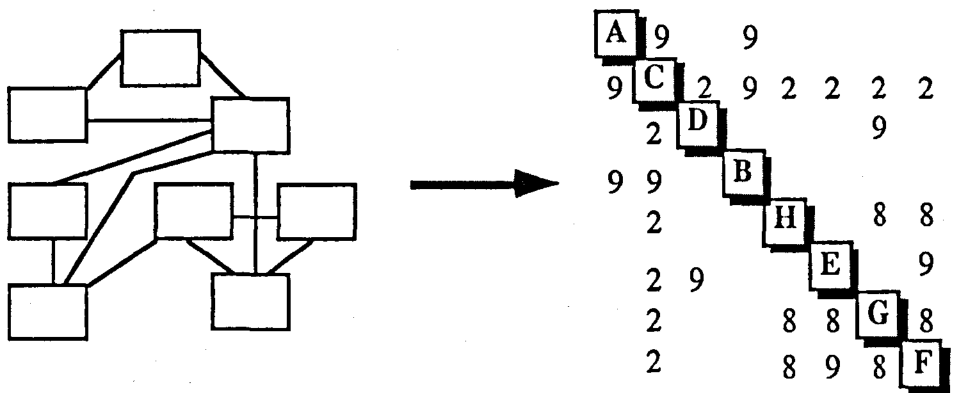


FIG 5. CADRAT - INITIAL INCIDENCE MATRIX

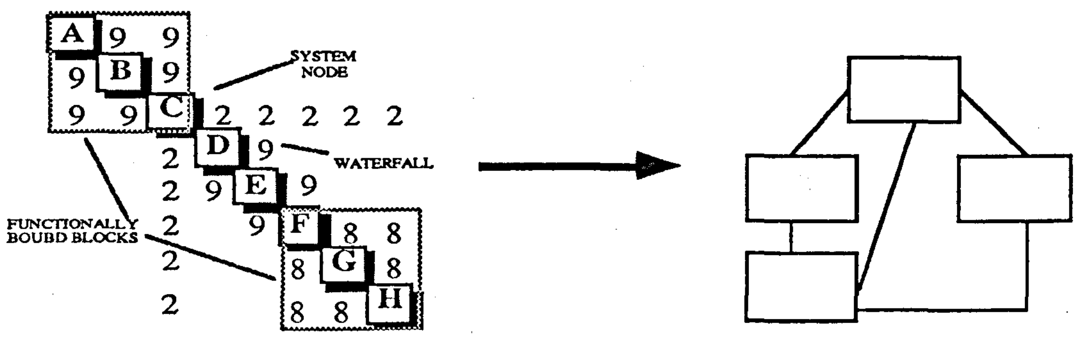


Fig 6. CADRAT - CLUSTERED MATRIX

While this example is trivial, more complex examples can be addressed by user-architects given appropriate training and tools. Figure 6 was clustered by a tool called CADRAT, Computer Aided Design - Relationship Analysis Tool. CADRAT will repel as well as cluster, reveals missing interfaces, enables clustered entities to be aggregated to simplify architecture, and so on. In a more complex situation such as advanced avionic architecture, clustering and repelling drivers might include, respectively:-

Crew Accessibility	C of G
Spectral Overlap	Space
Crew Armouring	Fuselage window Space
Minimal Signal Loss	E-M Isolation / EMC
Resource Sharing	No. of Connections.

Temporal Competition	Damage Tolerance
Functional Similarity	Heat Dissipation
Physical Similarity	Local Environment

A 25-by-25 matrix with all interfaces filled to accommodate all the drivers is too complex to cluster by eye. User-architects can co-operate with engineers in establishing the relationships and priorities - CADRAT will then produce the architecture to their mutual satisfaction.

10. SOFT-TO-HARD CONTINUUM CONCEPTS

Soft methodologies have gained in popularity recently. They recognise that human activity systems are fuzzy, lack repeatability and cannot be specified or understood quantitatively. Attempts to use these methods as a basis for specifying IDA systems (systems in which user/operators are supported by technological systems, e.g. Air Traffic Control, aircraft, ships, tanks etc) have been fraught with difficulty. While SSM helps to unravel the problem, subsequent transition to a solution is more elusive.

The soft-to-hard Seven-Step Continuum seeks to provide a smooth, seamless transition framework. What, then, should be the features of such a continuum?

* Middle-Out Design - Early, exploratory work can be focused effectively using a "Template" which categorises essential information. By prescribing the information needed further on in the continuum, nugatory early work - characteristic of some SSM exponents - can be avoided.
* Many Small Steps - For a continuum to merit the title, each step along the path should be small and easily understood. Small steps imply many steps; some steps may need tools and techniques to accommodate difficult, time-consuming or data-intensive tasks
* Creative Entropy - This is a key concept, developing creative tension and completeness. The technique increases the net information content first by looking both at objectives and the threats to their achievement, and second by introducing a series of nested Necessary and Sufficient (N&S) Sets. An objective might be the strategy for achieving the Prime Directive of a System -of Interest (SOI); the threat would be to the achievement of that strategy. An N&S Set might be Performance, Availability (of Performance) and Survivability (of Performance); as a set, these three ensure Performance.

So the continuum is underpinned by unobtrusive formality which permeates the processes and seeks both to expand and focus information in a creative, innovative framework.

11. THE SEVEN-STEP CONTINUUM

At the highest level, the Continuum comprises seven steps which progress from soft, through firm or mixed to hard system design concepts. The Seven-Step Continuum can be used iteratively, first to develop concepts and then to develop designs. The Seven Steps are shown at Figure 7 - of course.

Step 1 - Understand the Issues — Softer
Step 2 - Establish the S-O-I Requirement Template
Step 3 - Develop Process and Structure
Step 4 - Assign Man-Machine Roles
Step 5 - Develop S-O-I Performance
Step 6 - Develop S-O-I Effectiveness
Step 7 - Develop the Preferred Design Concept — Harder

FIG 7: SEVEN-STEP CONTINUUM FROM SOFT TO HARD

There being no concensus of what constitutes a completed design concept, Steps 5,6 and 7 develop increasing detail and concept robustness. For some, the first four steps will suffice; for others steps 5, 6 and 7 can be added as needed. Being a continuum, it follows that the steps comprise an extendable ladder; Step 6 cannot be undertaken if Step 5 has not been addressed, and so on.

The role and limitations of the user-architect are not clear at the high level of Figure 7. Following topics will unfold the Seven-Step Continuum, will decompose steps to aid understanding and will explore the limits of user-architect applicability.

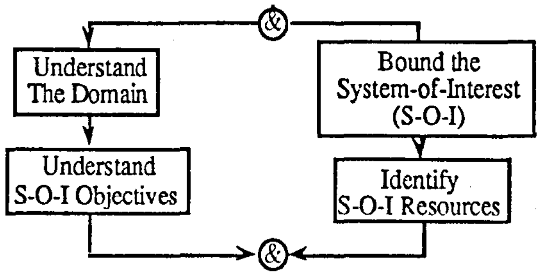


FIG 8: STEP 1, LEVEL 1 - UNDERSTAND THE ISSUES

12 STEP 1, LEVEL 1 - UNDERSTAND THE ISSUES

Figure 8 shows the first level of breakdown, or decomposition, for Step 1.The decomposition shows four related tasks which, it would seem, are appropriate to user-architects except, perhaps, for "Bound the System-Of-Interest (SOI)". Even experienced analysts find this task to be excrutiatingly difficult, the trap being to identify the SOI boundary with some convenient physical or organisational boundary. Bounding systems is difficult and requires training and method; neither is beyond the scope of either engineer or user-architect, but the latter may have the edge due to his domain understanding. To appreciate Step 1 further, we must decompose it to level 2

13. STEP 1, LEVEL 2 - UNDERSTAND THE ISSUES

The second level of Step 1 decomposition is at Figure 9, where the numbers for each box are Work Breakdown indices. With more detail revealed, the role of the user-architect and the essentially soft nature of the tasks becomes evident. Notice the use of Creative Entropy in 1/6

15. SOI REQUIREMENT TEMPLATE

The culmination of Step 2 is the completion, to a depth appropriate to purpose, of the SOI Requirements Template of Figure 11; the figure shows only headings. All earlier work should have been focused towards one or more elements in this template; all headings should be supported by comparable levels of detail. In this way breadth of understanding appropriate to purpose is also assured.

Note the use of Creative Entropy and of N&S Sets:-

* Creative Entropy - Strategy for Achieving PD vs. Threat to Achieving Strategy: Measures of Effectiveness vs. Negative Contribution Factors: Semantic Analysis
* N&S Sets - Management Set: two Threat Sets: MOE Set: Behaviour Set.

It is the aim, in developing the Seven-Step Continuum, to extend these complementary concepts throughout the approach, providing a balanced development of design information with minimum wastage.

S-O I PRIME DIRECTIVE
Semantic Analysis
of
Prime Directive

STRATEGY FOR ACHIEVING PRIME DIRECTIVE
Unifying Concepts
Innovative Approaches
Strategy Options

Behaviour
Aggression
Discretion
Co-operation

Management Set
Mission Management
Vehicle Management
Resource Management

Measures of S-O-I Effectiveness
Performance
Availability of Performance
Survivability of Performance

Negative Contribution Factors
Cost
Parent System Degradation
- Consumption / Dissipation
- Opportunity Cost
- Adverse Mission Effectiveness

Threat to Achieving PD
External
Internal
Environmental

Political / Economic / Technology

Fig 11: S-O-I REQUIREMENT TEMPLATE

16 THE MANAGEMENT SET

An example of one of the N&S Sets, the Management Set, will help to explain the role of the SOI Requirements Template. Figure 12 presents an outline of the Management Set for an advanced avionics system for a futuristic fighter. The three elements, Mission Management, Vehicle (Platform in this instance) Management and Resource Management focus attention in turn on different aspects of the avionics systems' activities in support of the Prime Directive. Mission Management looks ahead and plans to achieve the strategic objective of the mission. Vehicle Management preserves the means (aircraft and crew) of pursuing the Mission in a near-term, tactical sense. Resource Management husbands and apportions resources in pursuance of the other two sets of activities. The Management Set thus can be seen as another variant of

Performance, Survivability and Availability, in respect of attaining the mission. The elaboration of the Management Set can be seen as appropriate to the user-architect in particular, since it is related to the "do" activities of the aircraft / crew combination, rather than to technology.

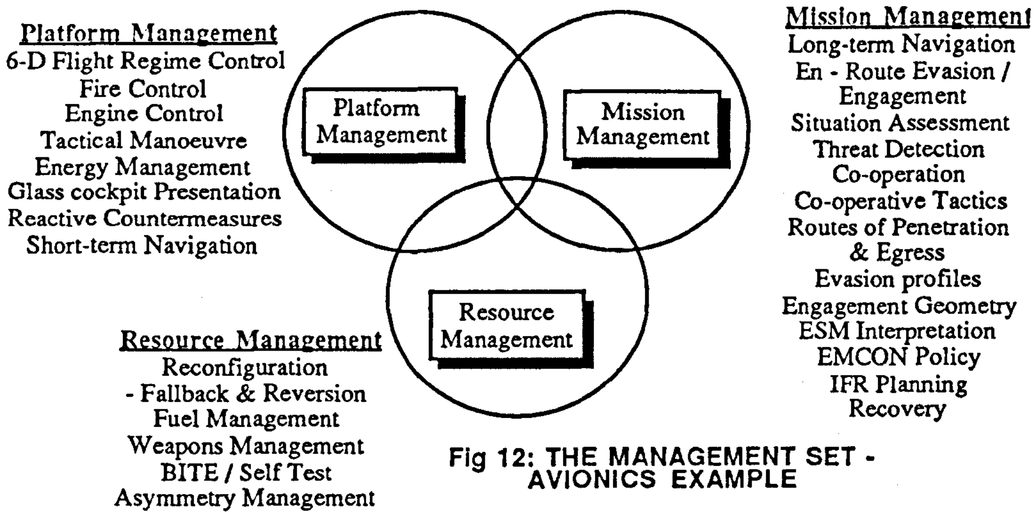


Fig 12: THE MANAGEMENT SET - AVIONICS EXAMPLE

17. STEP 3, LEVEL 2 - DEVELOP PROCESS AND STRUCTURE

Step 3 is elaborated in Figure 13. The tasks fall into two groups: at the left develop organisational architecture by clustering internal system activities and their external relationships and by linking those sets internally and externally. The internal SOI activities are developed on the right hand side, which is clearly the pre-cursor to functionality and to the purpose, intent and high-level design of software. Step 3 thus lays the foundation for architecture in the functional, physical and organisational senses.

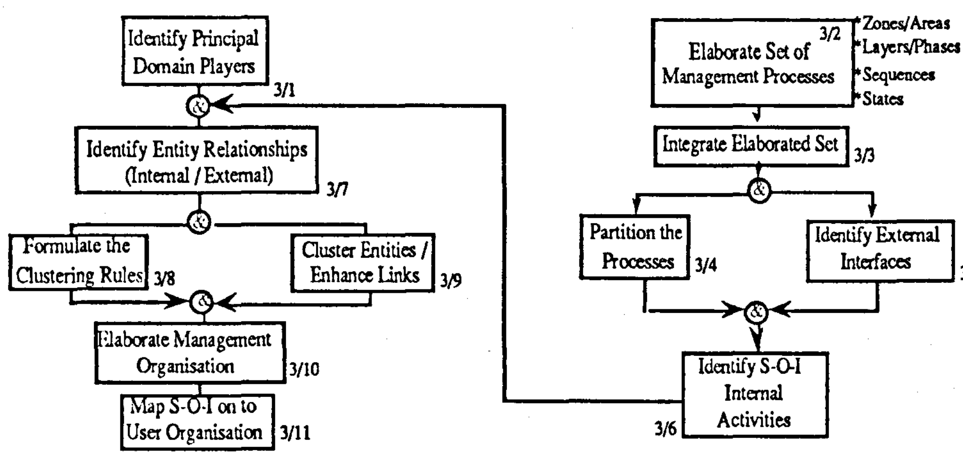


Fig 13: STEP 3, LEVEL 2 - DEVELOP PROCESS & STRUCTURE

In both blocks, at left and right hand sides of Figure 13, techniques and tools are necessary. None of these is exceptional; CADRAT has already been introduced as one means of handling the complexity and of clustering according to rules (Tasks 3/8 and 3/9) and sound techniques exist to enable the intelligent user-architect to undertake the whole of Step 3.

18. STEP 4, LEVEL 2 - ASSIGN MAN / MACHINE ROLES

Step 4, however, moves towards a distinct need for specialised capability found presently in system design engineers. User-architect roles are associated with setting up the SOI Solution Template, the SOI Information Template and in the most difficult task of apportioning functions / tasks between man and machine. Clearly user-architects, by virtue of their domain knowledge, can contribute to all activities in Step 4, but principally to the apportionment of tasks, (4/9). This most difficult, perhaps, of all tasks, must have deep user involvement and is best handled by developing a variety of solutions and trading between them.

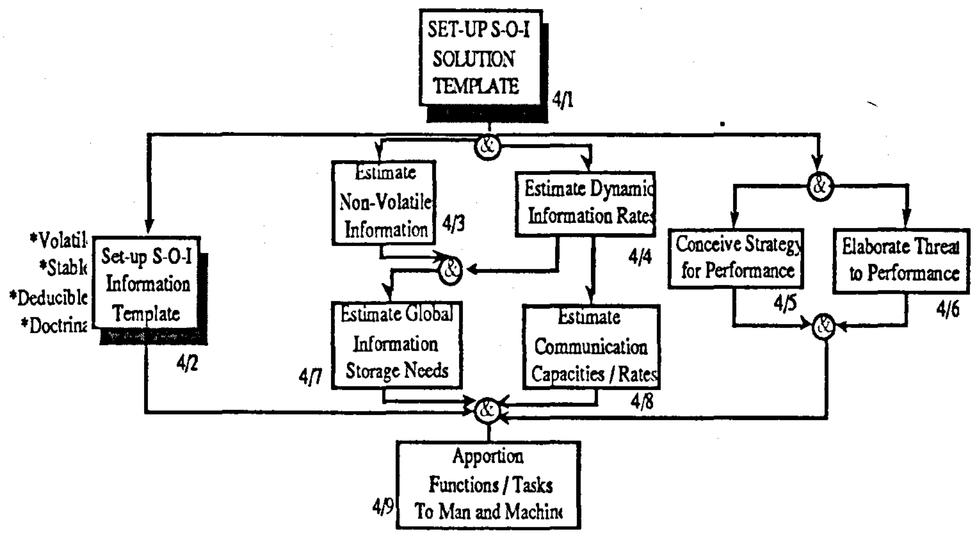


Fig 14: STEP 4, LEVEL 2 - ASSIGN MAN-MACHINE ROLES

19. THE SOI SOLUTION TEMPLATE

The SOI Solution Template is designed to facilitate tradeoffs between optional design concepts - see Figure 15, which contains columns for differentiating judgemental criteria and for solution options. Options are ranked or scored against each criterion and the preferred option selected accordingly. The SOI Solution Template contains no specifics about each design concept option; these are held separately. Instead, it carries forward the measures established in the SOI Requirements Template. Some factors from this prior template do not appear explicitly but are evident in the scoring / weighting process. Survivability and Availability are judged in relation to the Threat: Behaviour influences Performance, Security, Interoperability and others.

The Solution Template is arranged to highlight cost-effectiveness; the effectiveness scores may be sub-totalled in each column to appear immediately above Cost, the first Negative Contribution. Net Contribution may be assessed comparitively by trading Effectiveness (Positive Factors) against Negative Contribution Factors. While a numerical scheme is useful to clear the thinking processes, the final assessment of Net Contribution is better effected using qualitative measures. The user-architect has much to offer in this selection Process.

		OPTIONS 1	2	3	4
EFFECTIVENESS	PERFORMANCE - Mission Management - Vehicle Management - Resource Management AVAILABILITY - Reliability - Maintainability - Re-Configurability SURVIVABILITY - Avoidance of Detection - Self Defence - Damage Tolerance INTEROPERABILITY USABILITY SECURITY, ETC				
NEGATIVE CONTRIBUTION	COST ADVERSE PARENT MISSION FACTORS FAILURE PATTERNS CONSUMPTION DISSIPATION MASS / WEIGHT VOLUME / SHAPE, ETC				
	NET CONTRIBUTION				

Fig 15: S-O-I SOLUTION TEMPLATE

20. USER'S ROLE IN THE SEVEN-STEP CONTINUUM

The Seven-Step Continuum is being developed by applying it to real problems, so as to test its utility to system designers and user-architects alike. City University offers a BSc Degree in Management and Systems which attracts a cross-section of students from advertising, police, the forces, leisure, management consultancy, finance and many more. Some students come straight from school - others are mature. The Course includes a section on Technical Innovation during which the students have been introduced to the Seven Step Continuum in developing design concepts for Air Traffic Control and a notional Channel Tunnel Crisis Management System, with surprising - indeed, very surprising success.

The Seven-Step Continuum is also being applied to the design of concepts for an advanced avionics architecture, in this instance in support of a multi-disciplinary, multi-company design team of professional engineers, analysts and systems designers, facing a severe design challenge. The Continuum serves as a route map to keep the team heading in the right direction and to co-ordinate the various disciplines.

Figure 16 shows a rough estimate of user-architect potential, based on the present, limited experience, to undertake each of the steps in the Seven Step Continuum, given training and tools which have been made available to the City University students. The potential for user architects is clear; while later tasks need more specialist support, the first three steps - which set the scene for the following activities and design - are more suited to user-architects than to traditional engineers and even in later steps the user-architect has much to offer at key points. Instead of the present dis-jointed process of declaring operational requirements and then letting designers, with their knowledge elicitators, loose in competing teams, there is a sound prospect of integrated teams, led by professionally trained user-architects and supported by system designers, tools and techniques, developing sound concepts with high-integrity and applicability in much less time and cost.

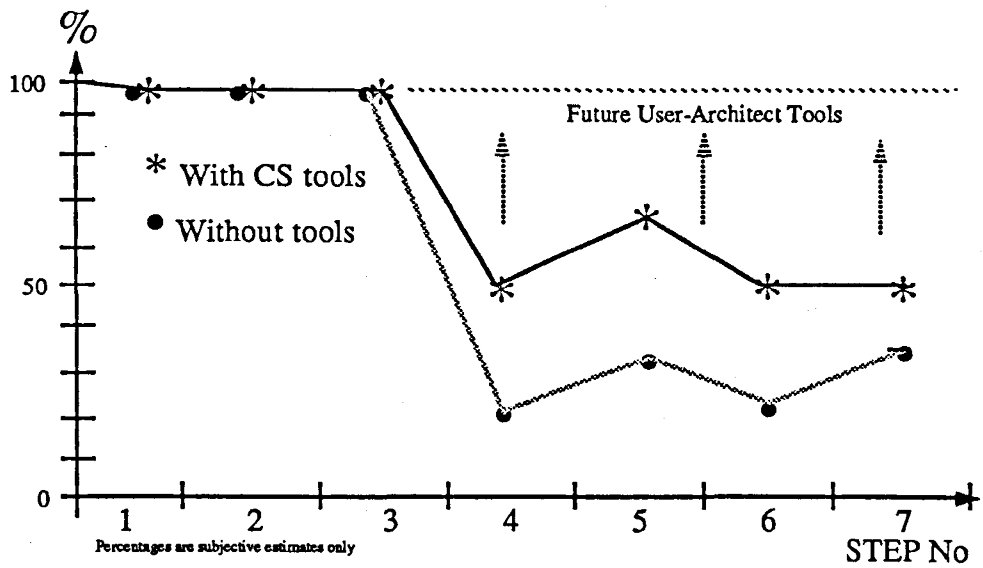


Fig 16: USER'S ROLE IN THE SOFT-TO-HARD CONTINUUM

21. SUMMARY

The paper started by elaborating the Catch-22 situation facing defence procurers today. Fixed price contracting is driving for more detailed specifications which attempt to define soft issues in hard terms. These complex specifications are taking longer and proving less satisfactory as systems become more complex and their development cycles extend, resulting in systems which are increasingly obsolescent at delivery. Techniques such as Knowledge Elicitation are vicarious and offer, at best, a limited interpretation of the user's knowledge and understanding of systems practice, doctrine and relationships in his domain.

How much better if users could bridge the gap presently separating them from system builders. The paper postulates three ways of bridging the gap: user-architects; intelligent tools designed by engineers to support those user-architects; and systems which evolve in-service to accommodate changing environments and threats unlike today's static systems. Only the first two approaches only were addressed in the paper. (For an introduction to the latter approach, see"BM / C3 Design by Cumulative Selection",a paper in the C3MIS Conference of the IEE, April 1989, by the same author, ISBN 0537-9989).

To explore the potential both of user architects and of prototype user-architect tools, a Seven-Step Continuum for developing design concepts from soft to hard was unveiled and the user-architect's role was assessed step-by-step to find his / her limitations

CREATIVE DESIGN BY CUMULATIVE SELECTION

By Professor Derek K Hitchins

City University, Department of Systems Science

INTRODUCTION

Cumulative selection is the process which, in Nature, is responsible for the development of the Earth's profusion of life-forms, animal and vegetable. It has been suggested that such complex physical features as eyes could not have occurred by random chance. The probability that an offspring could, in one gigantic evolutionary bound, develop an eye when neither parent had any such feature, must be vanishingly small. Cumulative selection is not random chance, however - see Figure 1. It arises because offspring are marginally different from either parent. It is the imperfection in reproduction that is at the heart of the matter; some offspring are inevitably better adapted than others, albeit marginally, to their changing environment. Those better-adapted offspring are able to survive or perform more effectively than their siblings and peers, and so they breed preferentially in turn. From generation to generation, beneficial features are retained and enhanced while disadvantageous features are suppressed and discarded. There is, however, no magical template representing some future ideal design. As Richard Dawkins states in his book"The Blind Watchmaker" (Reference 1), Nature is blind to the future.

This paper addresses the prospect of emulating in some very small, but none-the-less useful, way the natural process of cumulative selection as a new means both of designing systems ab initio and of creating systems able to adapt sensibly to their changing environment - perhaps the simplest definition of intelligence.

APPLICATIONS

Applications for cumulative selection abound in principle, but are less evident in practice. The process is becoming more widely used for the design of systems which are not readily susceptible to conventional, mathematically-based analysis, such as the design of low-drag vehicle silhouettes and turbine inlet guide-vanes - see figure 2. The paper will discuss the other applications shown in the figure, which have in common that they potentially generate a combinatorial explosion of design options which have to be reconciled against a plethora of mutually-incompatible measures of effectiveness. Cumulative selection can thus be seen as a simple way of handling complexity.

CAD TOOLS AND THEIR APPLICATIONS

Four prototype system-level CAD tools have been developed to explore design issues using cumulative selection in a variety of guises.The four are introduced at figure 3, and they will be explained and described in following paragraphs.First, however, it is useful to review the way in which the cumulative selection principle might be put to work. The process is shown in figure 4. First, genes are postulated which code for pervasive system features or characteristics. These genes are used to draw or describe a parent "organism" using a stylised drawing routine which must have its equivalent in Nature. At the same time, a hostile environment is usually generated in which offspring are going to be compared and evaluated. Offspring are generated and the best peformer / surviver is selected to sire the next generation. The process continues until some end point is reached where successive generations show no improvement; alternatively an intermediate end-point may be reached at which time the environment may be altered.

PARENT SYSMORPH

The first CAD tool to be examined is called SYSMORPH. Figure 5 shows a parent SYSMORPH, which is drawn using three behavioural genes and two structural genes;

* Behavioural Genes. The Aggression Gene codes, obviously, for aggression; this is evidenced in the weapons to the right of the SYSMORPH which symbolically pierce the shell, since they are visible in use to an enemy. The Stealth Gene codes for passive sensors and against active sensors in the simple prototype. The Co-operation Gene codes for communications

* Structural Genes. The Shell Gene codes for armour and for self defence features - it also implicitly codes against agility, of course, as for a tortoise or a tank. The Delegation Gene codes for internal structure, processing, interconnections, etc.

SYSMORPH OFFSPRING

Figure 6 shows the generated offspring, A to J, with the parent SYSMORPH repeated in the centre. Careful examination of each offspring will reveal that it differs in one, and only one, respect from the parent. "A" has one less missile."G" has on more passive sensor and one less active sensor."I" has increased armour. And so on.

The hostile environment is a series of measures as follows:-

* Performance

* Survivability - Avoidance of Detection
  - Self Defence
  - Damage Tolerance

* Interoperability

* Availability

* Security

* Affordability

In a full version of SYSMORPH, each of these environmental factors would be represented by its own model, which would provide a score or measure for each offspring in turn. In the simple prototype, these various models are represented by look-up tables, serving the same purpose but in a much simplified manner. In either event, the factor-by-factor evaluation is followed by an aggregation process which results in a cost-effectiveness ratio being formed for each offspring. It is this ratio which is used to differentiate between offspring; the highest ratio selects the parent of the next generation. Affordability provides the ratio denominator; it is comprised off platform or structure cost, carried weapons costs and systems engineering costs, in order to give a sensible differentiation in cost terms. Reference 2 contains further details of the evaluation process.

NEW MODELS FROM OLD?

The approach employed in the SYSMORPH tool can now be seen in perspective. As Figure 7 shows, the conventional process uses (for a fighter aircraft as ana example) a "trusted" combat model to represent the hostile environment, into which is placed a model of a proposed new fighter.The results are analysed, often off-line, and adjustments made to the fighter model prior to the next run. SYSMORPH does much the same, but the process is formalised and a rich variety of different embryo fighters is examined at each stage, with new generations being propagated until successive generations show no further benefit or until the operator is satisfied. Unlike more deterministic techniques, SYSMORPH is extremely easy to follow and understand, and every generation can be checked for sense; sometimes the first steps taken by the program are more interesting than its end point.

Two important features emerge; in both the conventional and the SYMORPH approach, the reference is a trusted model representing the hostile environment, and the amount of computing (number of generations) depends on how near the start-point is to a good solution. With the SYSMORPH approach it is necessary to start with a working system which can then be optimised; SYSMORPH will not create a working system.

SYSTOLIC CUMULATIVE SELECTION

Having seen the essential similarity to conventional procedures, the simplicity of the approach suggests that one further step may be practicable. If it is possible to evolve a fighter in a hostile - but stable - environment, is it possible to evolve the environment? Figure 8 shows the approach, called systolic cumulative selection because of the similarity to a heart-beat, with first the fighter evolving, then the environment, then the fighter again, and so on. Evolving the environment amounts to evolving the hostile fighters and ground defences which it contains and is essentially the same process as basic SYSMORPH. There is, of course, a potential for runaway if the process is allowed to free-run; systolic cumulative selection needs to be carefully controlled by competent users, but the concept is exciting, creative, and promising.

ADVANCED AVIONIC ARCHITECTURES

There is much interest in the avionics world resulting from recent step advances in avionics architecture design which promise remarkable improvements in reliability, performance and weight reduction.It is possible to derive such architectures using systolic cumulative selection first to evolve the fighter and then - remembering that the fighter's envelope constitutes the boundary to the avionics architecture - to evolve the architecture. Later topics in this paper will indicate the "genes" that code for architecture and more detail is given in Reference 3. On a grander scale, the full air defence system of which the fighter is an alement is itself a dynamic architecture and may be amenable to design by cumulative selection too.

RAIL CARRIAGE DESIGN

The SYSMORPH approach is potentially useful in a wide variety of situations; one such, in principal if not yet in practice, is the process of rail carriage design. This is a sensitive and important subject and one requiring great expertise; such expertise would come not from SYSMORPH, but from the models and experts who created the hostile environment. Figure10 presents some of the many factors, often mutually incompatible, which must be reconciled in the design of safe, efficient , effective and affordable rail carriages.The role of SYSMORPH would be to aggregate the wisdom of the experts in their different disciplines and to evolve the optimum design solution in a formal procedural framework.

NETWORKS AND TREES

The second CAD tool is called COMMORPH and, as the name implies, it concerns itself with the shape of communications (or any other) networks. Figures 11 and 12 show COMMORPH in action on a simple network problem. Figure 11 presents the problem; to connect the source / sinks using four network switches using the shortest overall route. The process is simple; four switches are scattered randomly among the source sinks and allowed to connect first to their nearest source / sink and then to each other. Each switch is then moved in small steps from its start point in eight different directions, with re-association occuring each time. The step incurring the least total-network length is selected, the switch relocated, the procedure steps on to the next switch, and so on. The four switches slowly migrate across the screen to take up the positions of Figure 12 - not perhaps what might have been predicted. Figures 13 and 14 show a similar example with the program required to connect by a tree rather than a network. Note how the switch nodes (N) space themselves equally along a straight line, a natural consequence of the technique.

The process is made much more interesting when terrain is included and the links are provided by radio. Instead of seeking simply a shortest overall length, that concept can be combined with exceeding a threshold signal path loss. Switches will now migrate to avoid obscuring hills and low-lying ground in order to achieve line of sight.

The number of switches can now be seen as an important variable; in fact, the figures shown for networks and trees are effectively parents, since the number of switch nodes and the connective redundancy are architectural "genes". Offspring would present architectures with variations in these and other pervasive characteristics.

THE CLASSIC STEINER TREE PROBLEM

The Steiner Tree problem is a classic example of the difficulty facing analysts trying to calculate solutions to seemingly simple problems. It concerns the non-redundant connection of a number of major cities in the US by telephone; the task is to calculate the single, shortest route which does not connect any two points by more than one route. Reference 4 gives an excellent description of the problem, its difficulty and ways of solving it. Figure 15 shows a simple, regular test case comprising eight points in a rectangle to be connected by the shortest route. Figure 16 shows the solution as worked out by TREEMORPH using cumulative selection. The TREEMORPH program is very similar to COMMORPH, differing principally in the way connections are made between the migrating nodes and the fixed points-to-be-connected. Steiner points are characterised by being at the centre of three lines mutually at 120 degrees. It is thus possible to be fairly certain that the solution at Figure 16 is correct. For more complex and irregular patterns of points, it is possible to derive sub-optimal trees containing Steiner points without necessarily realising that a globally minimal set has not been reached. Various techniques are available for "shaking" the tree out of such suboptimal configurations.

CADRAT

The fourth CAD tool explores the relationships between system entities; these might represent subsystems, organisational groups, artifacts or - most usefully - functions of a system. The relations are described on an N-squared (N2) chart; an archetypal N2 chart is shown at Figure17. Functions and their inter-relationships in the top left diagram are mapped into the N2 chart, an incidence matrix, in such a way that outputs from any function appear in the row containing that function while inputs to any function appear in the column containing that function. A two-way relationship, represented conventionally by a two-headed arrow, thus appears in two separate squares of the chart; each square not on the diagonal thus represents a potential unidirectional interface between the functions on the diagonal.

Interfaces form patterns. Where a function has interfaces to and from every other function it appears at the centre of a cross; it is a node. F3 is such a system node, and reference to the top left diagram confirms that the loss of F3 would indeed cut the system in two; system nodes are candidates either for protection or replication to enhance system robustness. Hierarchies are represented as waterfalls. Functionally bound blocks appear as blocks of interfaces with all the squares filled in - there are two examples on the figure. Functionally bound blocks are interesting because they represent an inherent grouping to be fostered, either by strengthening or protecting the interfaces or - more usually in organisational terms - by grouping the appropriate functions together as a single physical group. This procedure results in the bottom left figure, simpler than the starting diagram, but topologically the same.

Unfortunately when tackling a real-world problem it is indeed rare to collect relationship data and to find it nicely grouped as is Figure 17. The example of Figure 18 is more likely to be met. In this example, there are 25 military units scattered randomly on the diagram, giving 0.5xFactorial 25 (=7.76E24) permutations of the figure, suggesting strongly that crude approaches such as simply generating every option before selecting the best are not very practicable. Cumulative selection is one of several ways to cluster the interface patterns economically into meaningful groups by rearranging the units; the result is shown at Figure 19, from which it can be seen that the clusters do indeed form sensibly.

Valuable although this result is, it is perhaps surprising to find that the decimal values in the interface squares have little effect. Much the same results are obtained by entering a "1" for interface present and a "0" for interface absent. What, then, is the value of the numbers? They can be used to represent a variety of characteristics, including strength of relationship, urgency or immediacy, interface medium or protocol, and so on. If the numbers are used to represent some hierarchical relationship, then that hierarchy can be examined layer by layer simply by selecting a band of numbers. This is tantamount to "slicing" horizontally through the pyramidal hierarchy and looking at the cross-section. In most organisations, while authority moves vertically up and down the hierarchy, work and co-operation are achieved across the structure. Being able to take horizontal slices at various levels makes it possible not only to see this horizontal structure in bold relief but, by clustering this cross-section, the underlying organisation can be revealed.

BUILDING LAYOUT

Networks and relationships can be used together in the design of intelligent and functional buildings. Figure 20 shows two stories of a typical airport arrivals and departures building. There is advantage to both user and builder in grouping the various elements of the building in three dimensions. For example baggage in and out, passenger entry and exit, and services ducts should be near each other vertically. At the same time there are advantages to be gained by reducing walking distances, minimising congestion, reducing queuing delays and maximising passenger / baggage throughput, not to mention aesthetics of layout and appearance. COMMORPH and TREEMORPH tackle the nuts and bolts of shortest distances; there is even a right-angle variant of the Steiner Tree problem, useful for locating lifts, for example, so as to minimise the average walking distance to the nearest elevator. CADRAT can be used also to group and arrange the internal elements into sets that optimise their functional relationships. Aesthetics are presently beyond the scope of cumulative selection as described here.

CURRENT RESEARCH

Figure 21 shows some areas currently being researched. SYSMORPH faces the problem of combining measures of survivability, performance, interoperability and so on. It is unreasonable simply to add such measures algebraically, even if correctly scaled, since they are effectively in different dimensions. This concern leads to cost-effectiveness analysis, a term with little substance since many people claim to use it as a measure but few seem able to justify their method.

One reason, perhaps, is that effectiveness embraces the concept of utility - how useful something might be presents a subjective view of its effectiveness. In the economics of marketing, the concept of utility dimishing with increasing numbers is prevalent. With no shirts, a man may find the present of a shirt of great utility. With fifty shirts, a man may find the present of a shirt of little utility. These concepts, enshrined in the logarithmic Fechner's Law, are potentially useful in SYSMORPH programs to prevent evolutionary "runaway".

Shake algorithms have been mentioned; they seek to jolt a program out of a sub-optimal set. In the event, shake algorithms are perhaps more of a challenge than the basic programs they seek to supervise, since different sub-optimal patterns respond to different shake strategies. A battery of shake techniques may become necessary to resolve large scale problems.

Last in the list, and perhaps the most interesting of all, is the concept of rule evolution, being researched in the US for robotic machines. Programs controlling robots exhibit behaviour principally as a result of "IF THEN ELSE" statements which form break-points. If such statements are considered as genes, it is possible to conceive programs which evolve towards an optimum way of achieving their purpose. More importantly, such programs are able to follow a changing environment provided that the change is reasonably continuous and that the rate of change does not outstrip the program's inherent rate of evolution. This is perhaps the first glimmerings of true artificial intelligence - the ability to adapt sensibly to a changing environment.

References:

1. "The Blind Watchmaker", Richard Dawkins, Longmans, ISBN 0582 446949

2. "BM / C3 Design by Cumulative Selection", Professor D K Hitchins, Proceedings of the 3rd International Conference on Command, Control, Communications and Management Information Systems", 2-4 May 1989, published by the IEE. ISBN 0 85296380 7 ISSN 0537-9989

3. "Advanced Avionics Architectures", Professor D K Hitchins and D O'Dwyer, Proceedings of the Royal Aeronautical Society Conference "Digital Computer Technology for Airborne Applications" 15 March 1989, RAeS. ISBN 0 903409 2

CUMULATIVE SELECTION

CUMULATIVE SELECTION IS NOT RANDOM CHANCE.

PROGENY ARE IMPERFECT COPIES OF PARENTS.

IMPERFECTIONS ADAPT OFFSPRING TO THEIR (CHANGING) ENVIRONMENT

SUCCESSIVE GENERATIONS RETAIN EVEN PARTIAL BENEFITS ACCRUED BY THEIR FOREBEARS.

CUMULATIVE SELECTION IS BLIND; IT HAS NO TEMPLATE OF THE FUTURE............*BUT*

CUMULATIVE SELECTION HAS PRODUCED ALL OF THE EARTH'S BIOLOGICAL DIVERSITY AND INTELLIGENCE

Figure 1

APPLICATIONS

Low Cd car profiles	Turbine IGV design
Intelligent Robotic software	Ab initio design of - aircraft, ships,etc - rail carriages
System Architecture - clustering - linking	Building layout Network design
Management Organisation	Building Layout,etc

Common feaatures :- Potential combinatorial explosion
Many mutually incompatible measures of effectiveness

Ideal for initial design of complex problems

Figure 2

CAD TOOLS & THEIR APPLICATIONS

*SYSMORPH - morphological design of aircraft, ships, rail carriages,etc

*COMMORPH - physical layout of networks, communication systems, oil pipe-lines, routes of penetration and egress, corridors and pathways, circuit board layout, etc

*TREEMORPH - shortest non-redundant path connecting a set of points

* CAD-RAT - functional-to-physical mapping for organisational and management structures, systems engineering of advanced avionics architectures,etc

Figure 3

USING THE C.S. PRINCIPLE

Identify "genes" which code for system characteristics

Use the genes with a stylised drawing routine to draw the system

Establish this first drawing as the parent.

Create an environment hostile to the parent in which it has to
- perform
- survive
- propagate

Generate different offspring by varying the "genes" slightly

Subject all offspring to the hostile environment

Select the "best" offspring to parent the next generation

Figure 4

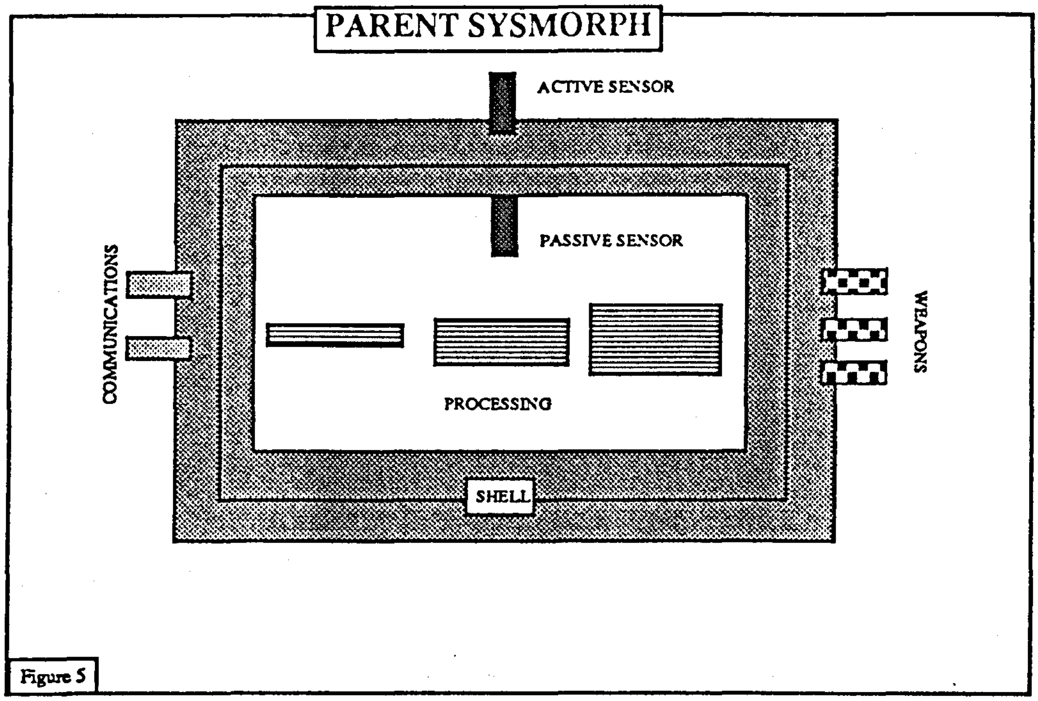


Figure 5

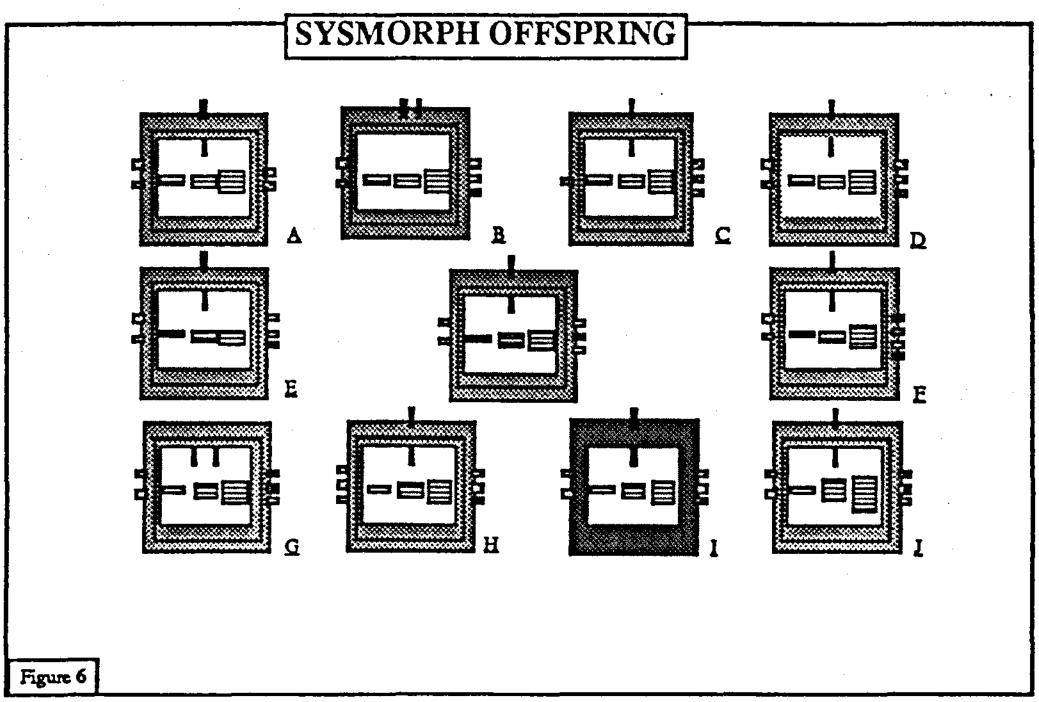


Figure 6

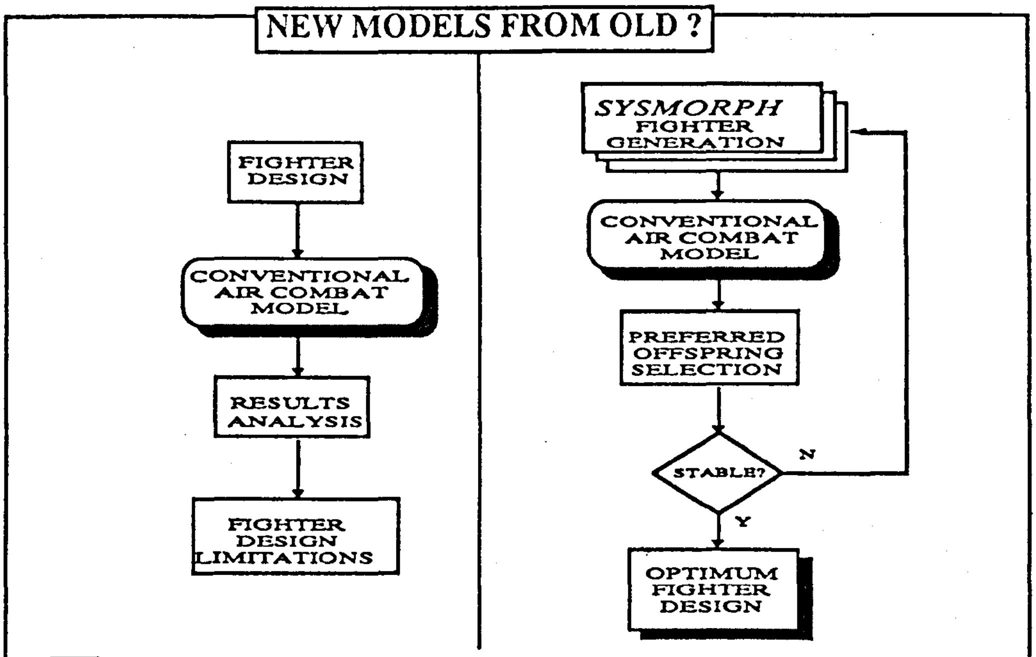


Figure 7

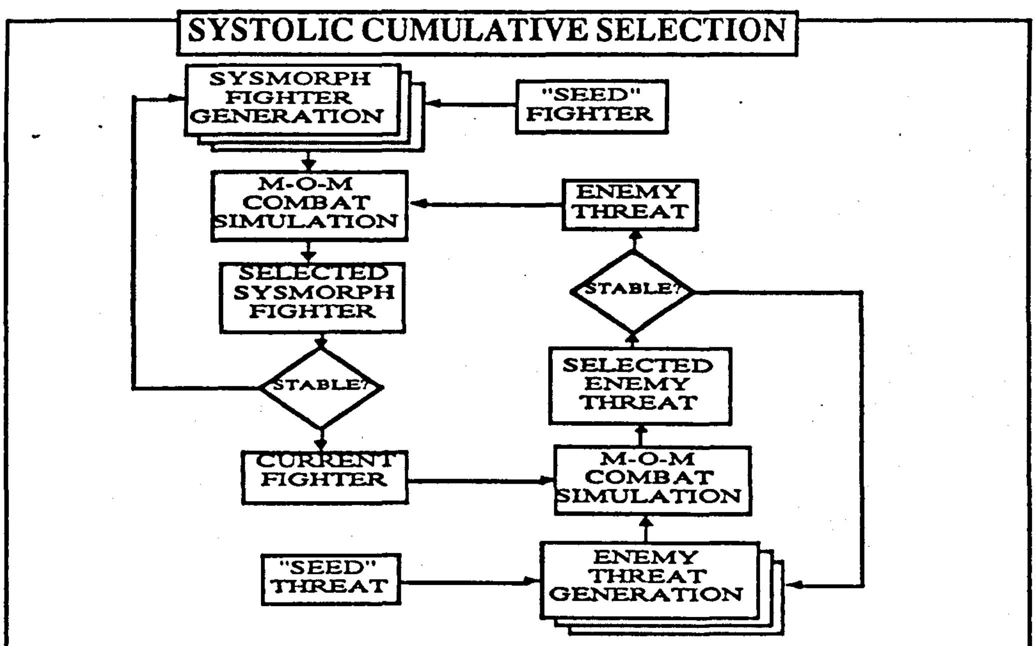
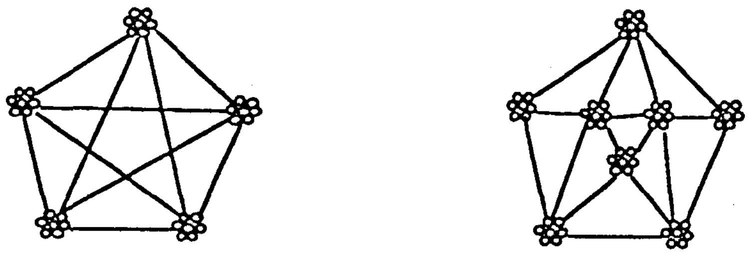


Figure 8

ADVANCED AVIONIC ARCHITECTURES



<u>OPTIMAL ARCHITECTURES</u>

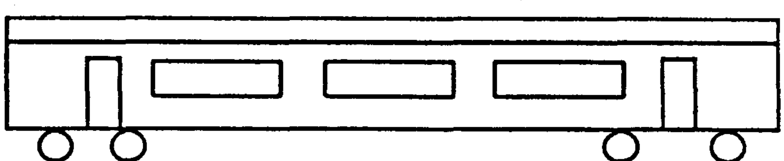
GOALS - RELIABILITY, WEIGHT REDUCTION , LOW COST

OPTIMUM NUMBERS OF NODES, LINKS AND LEVELS OF HIERARCHY FOR :-

PERFORMANCE, SURVIVABILITY, AVAILABILITY, EASE OF UPGRADE, LCC OF OWNERSHIP

Figure 9

RAIL CARRIAGE DESIGN



<u>OPTIMISE DESIGN</u>

* PASSENGER COMFORT * EASE OF EXIT
*PASSENGER CAPACITY * PASSENGER ACCEPTABILITY
* COST OF CONSTRUCTION * PRICE OF FARES
*BUCKLING PERFORMANCE * SHEAR RESISTANCE
*RIDE COMFORT * TOP SPEED

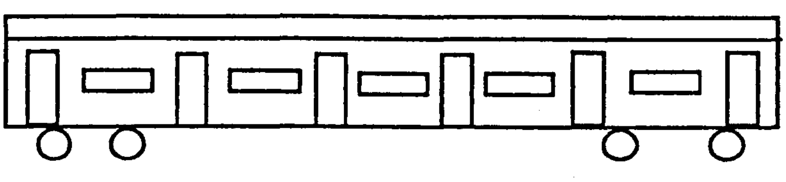


Figure 10

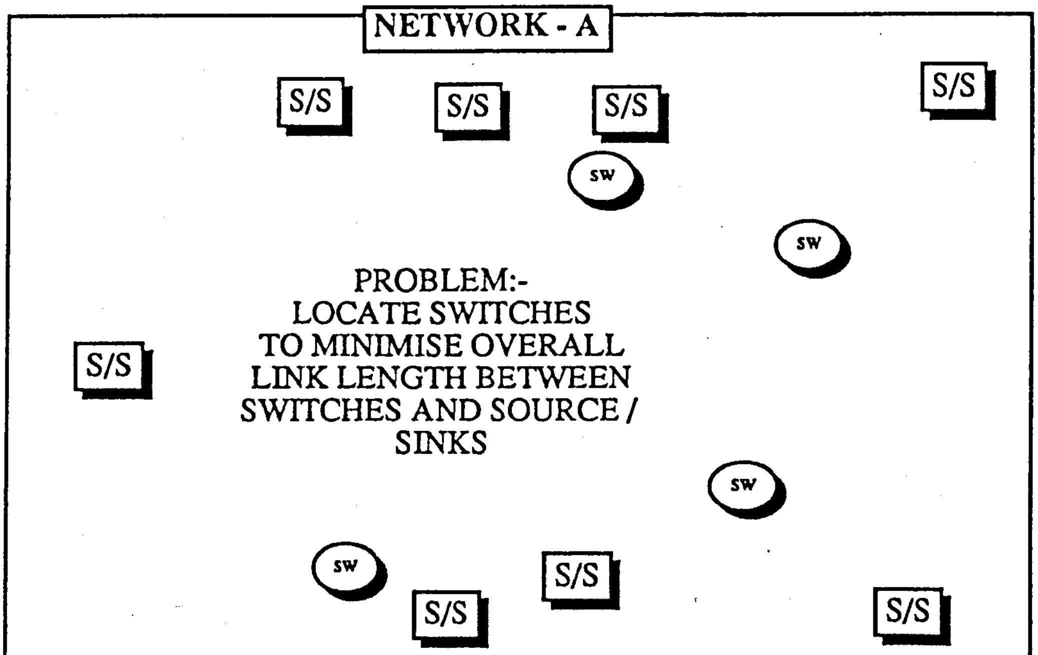


Figure 11

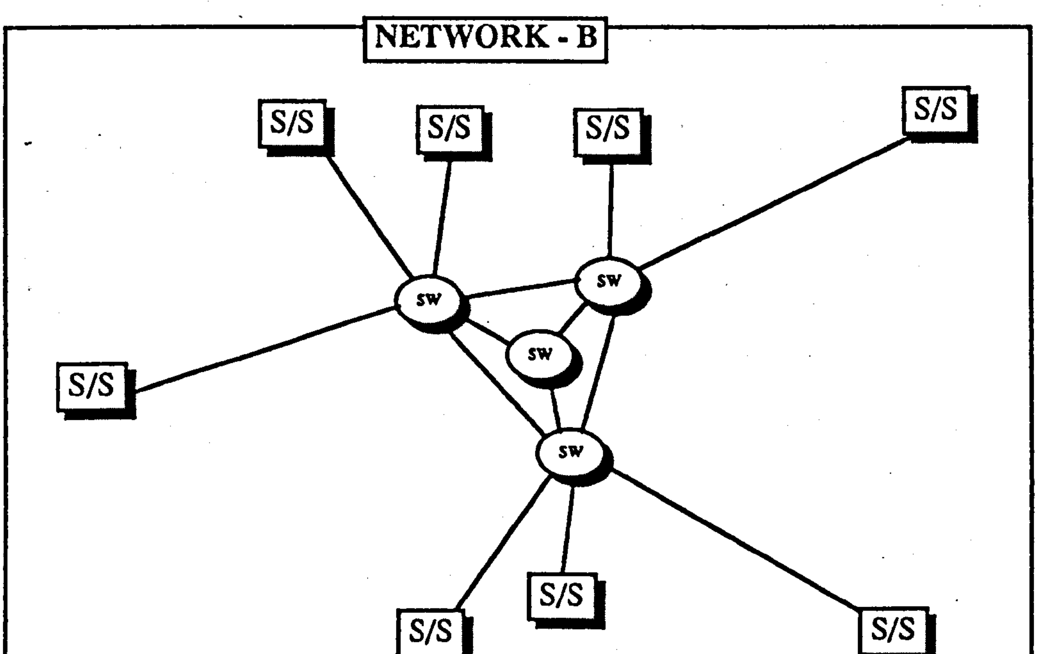


Figure 12

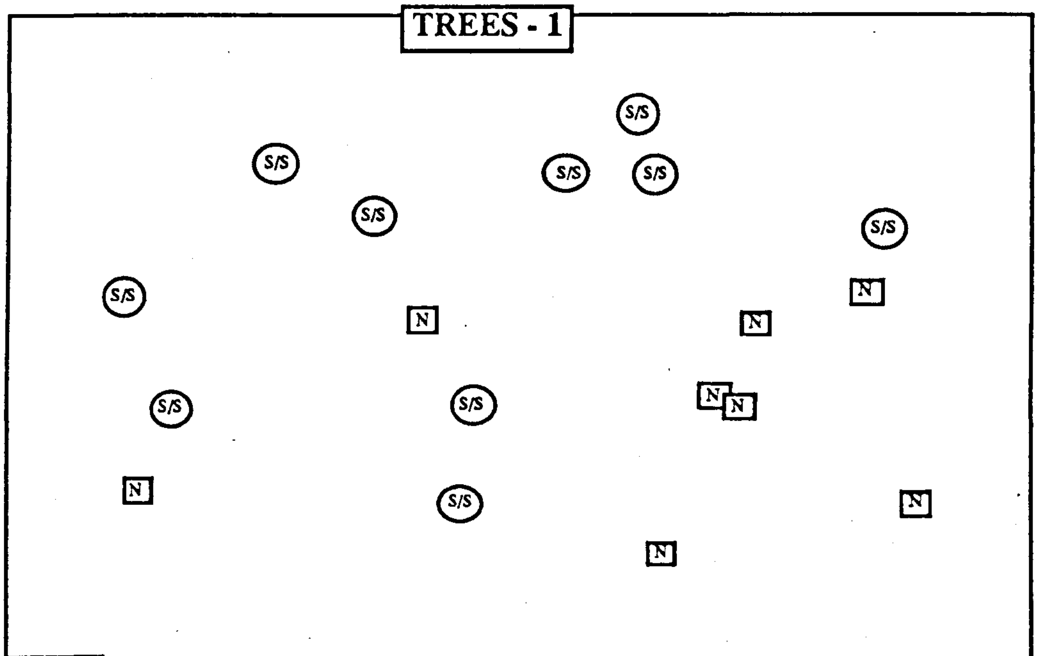


Figure 13

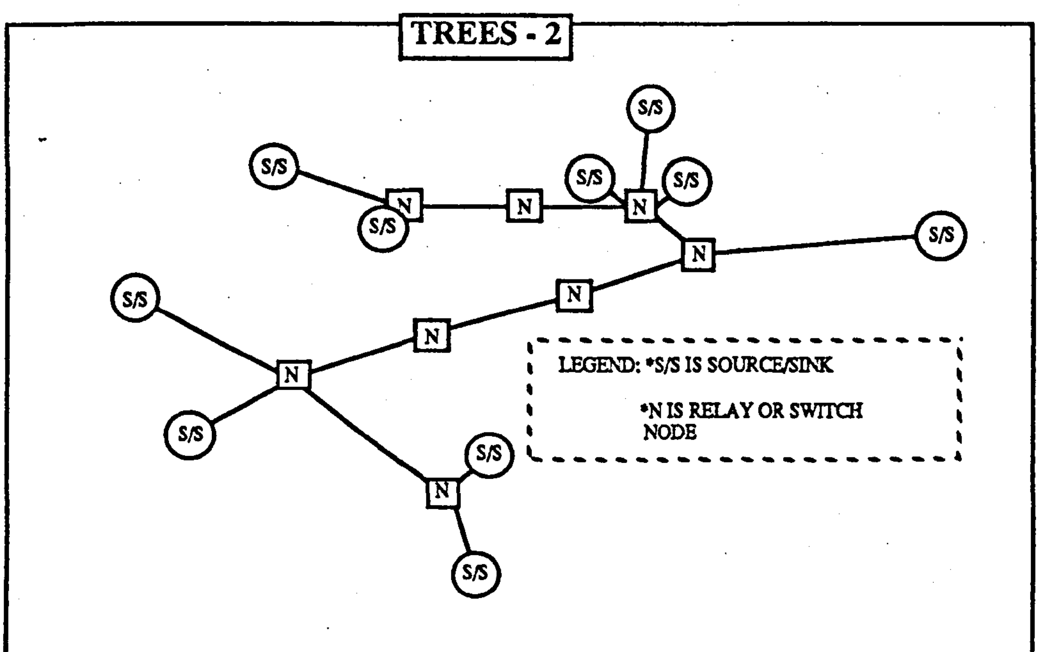


Figure 14

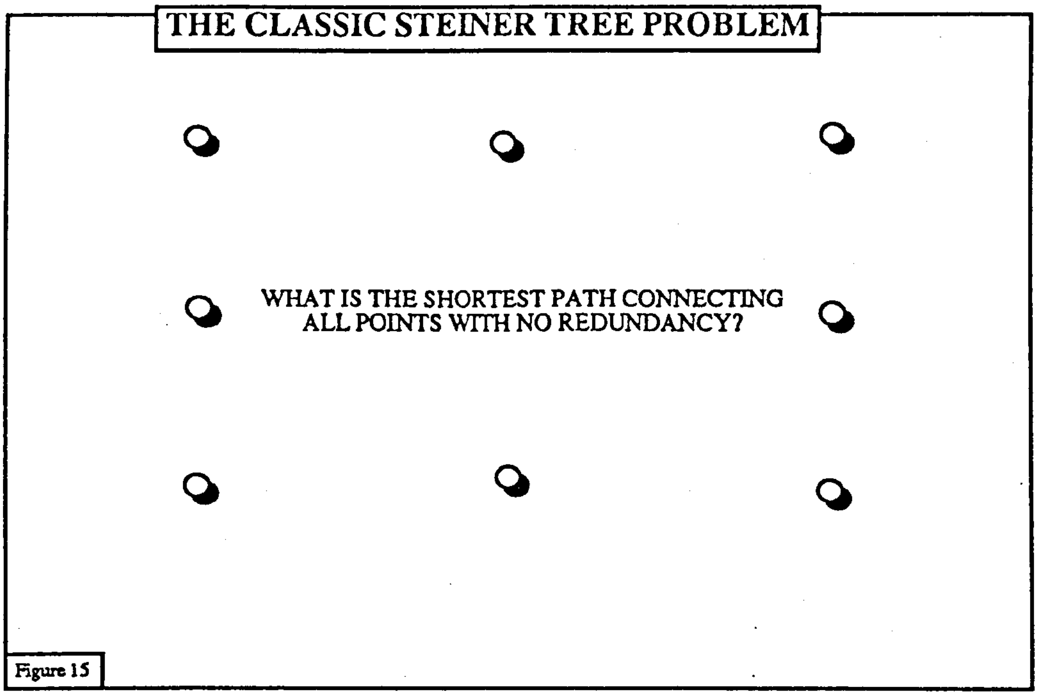


Figure 15

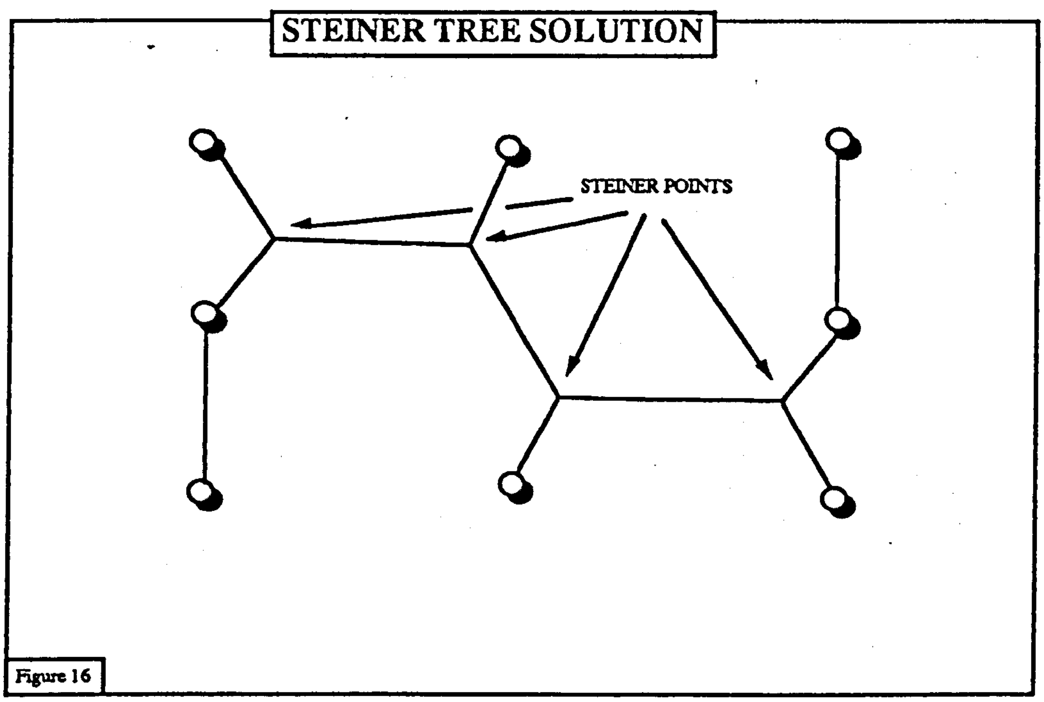


Figure 16

CLASSIC N2 CHART

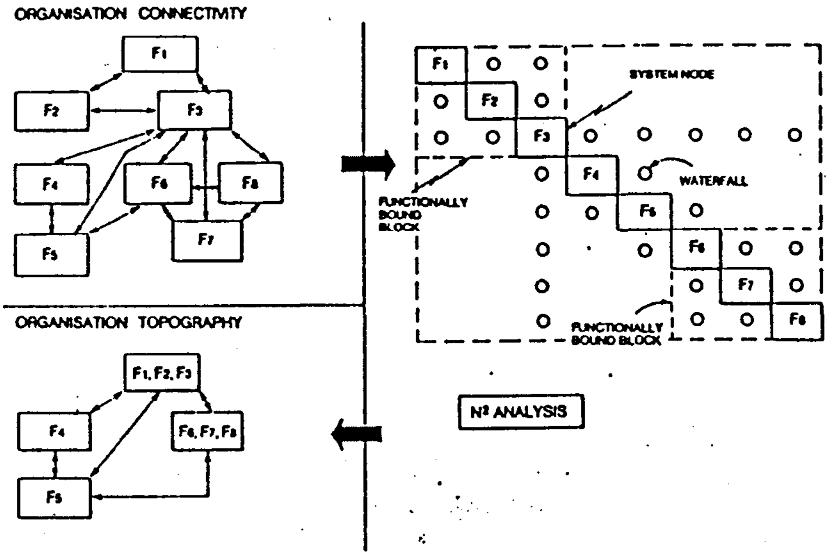


Figure 17

RANDOM ORGANISATION

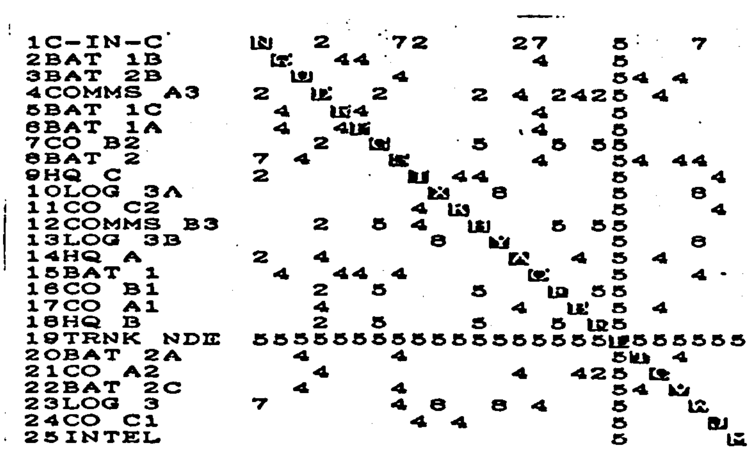


Figure 18

FUNCTIONALLY CLUSTERED ORGANISATION

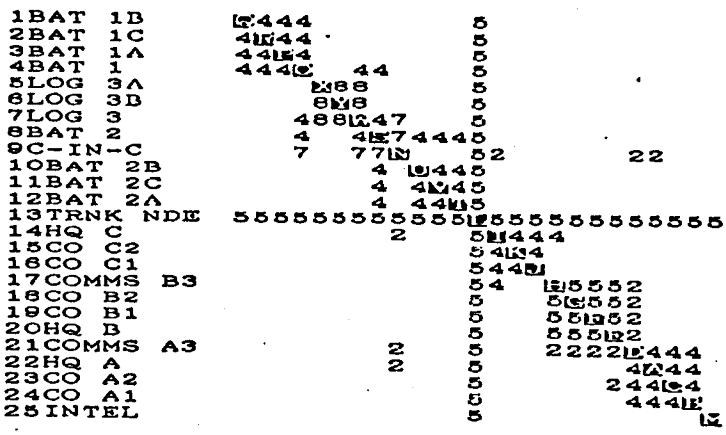
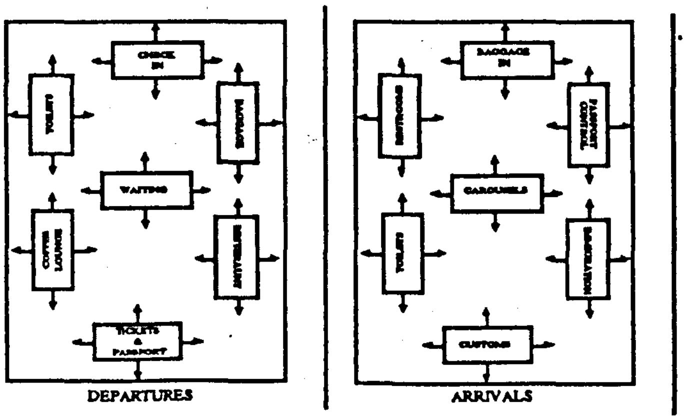


Figure 19

BUILDING LAYOUT



OPTIMISE:- * PASSENGER WALKING DISTANCE
*BAGGAGE DELAY *SECURITY RISK *THROUGHPUT

Figure 20

RESEARCH

COMBINING MEASURES IN DIFFERENT DIMENSIONS

COST EFFECTIVE ANALYSIS

MEASURES OF UTILITY

SHAKE ALGORITHMS

RULE EVOLUTION

Figure 21

EVOLVING RULES

* " IF THEN ELSE" RULES ARE "GENES" THAT CODE FOR SYSTEM BEHAVIOUR
* CREATE RULES WHICH EVOKE A GRADUATED OUTPUT
* PROVIDE INITIAL WORKING "SEED" PLUS BOUNDARIES
* OPTIMISE THE RULES IN A HOSTILE ENVIRONMENT - *IN REAL TIME, ON LINE!*

RESULT?

PERHAPS THE FIRST TRUE ARTIFICIAL INTELLIGENCE - THE ABILITY TO RECOGNISE CHANGES IN THE ENVIRONMENT AND TO ADAPT BEHAVIOUR IN SENSIBLE RESPONSE

Figure 22

A GENERAL THEORY OF COMMAND AND CONTROL

Professor D.K. Hitchins

City University, UK. Written principally when the author was Business Development Director at EASAMS Limited, to whom he is indebted for their continuing support in the preparation of this paper.

COMMAND AND CONTROL THEORY - CONCEPT AND PURPOSE

There is, at the time of writing, no established general theory of Command and Control (C^2). Indeed, there are many C^2 practitioners who would argue that the variety of C^2 applications and the variability inherent in this essentially human activity deny even the possibility of a theory.

That Command and Control System analysts and designers continue to operate in this supposed theoretical desert, however, does at least imply that underlying C^2 mechanisms do exist. A sound basis for appreciating these mechanisms would, if it could be formulated, provide a valuable tool for the optimisation of designs, and could act as a benchmark against which to judge the effectiveness of solutions.

A theory of C^2, like any other scientific or engineering theory, should consist of several parts:

- A corpus of knowledge
- A set of postulates, axioms and laws
- Hypotheses which predict the performance of C^2 solutions, and which can be checked by experiment
- A means of evolving the corpus of knowledge and the rules as the predictions are proved or disproved.

This paper is concerned principally with pursuing the second bullet, but in a particular way, as follows.

Command and Control has become confused in many minds with its supporting technology. Purveyors of contemporary C^2 systems in particular can sometimes overlook the essence of C^2 - Command and Control is "of and by people". The theory which this paper will address concentrates on organisation and relationship, and does not address technology. On the other hand, it will set targets which technology may find useful, challenging or even unattainable.

Where relationships appear to be axiomatic, they will be presented as so-called C^2 Laws; where the relationships are less certain, the term "postulate" will be used.

COMMAND AND CONTROL - WHAT IS IT?

What is C^2? This is at one and the same time an easy and a hard question. Command is the executive function - the decision-taking rôle. Command is a deep responsibility, unlike most civilian counterparts, for the men under command. Command is also the chess player. In war, however, chess is played with most of the opponent's board covered and the chess pieces do not always obey the rules. Control is the management of the Commander's decisions by instructing inaction, or instructing action elements to engage, disengage or move.

APPROACH TO DEVELOPING A C^2 THEORY

Norbert Wiener in his seminal book Cybernetics (1) implied that all systems, living and mechanical, are both information and feedback systems. Taking a leaf out of Wiener's book, the approach adopted in developing a C^2 theory was to develop and explore a variety of different models of C^2, using each to study different aspects of the subject. Types of models employed and their purpose include:

<u>Model Type</u>

- System Dynamics
- Queuing Theory
- Architectural/Structural
- Relationship
- Control Theory
- Timelines
- Information

<u>C^2 Aspect</u>

- System Boundaries and Overviews
- Timeliness of Decision Taking
- Span of Control/Understanding
- Infrastructure Complexity
- Stability
- Timeliness Equations
- Amounts, Granularity and Rates of Information Transfer

While Wiener is undoubtedly correct in his view, neither the Information nor the Control Theory models seems to be productive in generating C^2 theory; their value appears to lie more in the practical aspects of analysing particular systems. Other models are more helpful, however, and enable the following approach to be taken:

- Since technology has potential variety almost without limit, attempting a theory to address technological features is unlikely to be successful
- Instead, follow the line taken by Shannon, Nyquist and many others, and seek the ideal or perfect theoretical solution, devoid of technological context
- Develop theorems, postulates and hypotheses, based on this ideal approach, which can be explored by experiment.

BOUNDING THE CLASSIC C^2 SYSTEM

The technique referred to as System Dynamics by N. Roberts (2) has as many opponents as advocates. It undoubtedly has one distinct merit; it is one of the few ways of bounding a system while at the same time providing an uncluttered overview.

Figure 1 is an Influence Diagram of C^2 which presents Command and Control in action. The diagram is limited in its description of the processes inside the boundary. There are however five vital elements in the total system:

a. Threats and Opportunities

b. Political Influence

c. C^2 Decision - Military Control

d. Action Elements

- Weapon Systems
- War Consumables (Fuel, Weapons, etc.)
- Reserves and Reinforcements.

e. The Enemy, with his equivalent a. to d.

Figure 1 shows that superior control is actually exercised by politicians; they set the priorities and rules of engagement, and they influence the military doctrine which motivates the Commander's decisions.

Figure 1 conceals the delays inherent in the decision and action processes. Such concealment is not fundamental in System Dynamics. Figures 2 and 3 show the decision and action processes in more detail, with sequential and cyclic delays displayed. The interwoven dynamics now start to dominate, and new dimensions can be seen; repair cycles, replenishments, reinforcements and reserves clearly have a major influence on the Commander's options.

These influence diagrams indicate an approach to more quantitative models.

C^2 TIMELINESS

A basic objective of C^2 is to meet the challenge in a timely fashion. This objective forms the substance of the first C^2 equation:

$$\textbf{Opportunity Volatility} \geq \textbf{Perception + Communication + Decision + Planning + Action + Response + Translation + Minimum Engagement Window} \quad ...(1)$$

where:

- Opportunity volatility is the tendency for a situation of interest to evaporate, measured in units of time
- Perception is the delay in noticing the situation (often Intelligence Analysis)
- Communication is the delay in reporting
- Decision is the delay in electing for a particular course of action
- Planning is the delay in detailing the action
- Action Response is the delay in getting action elements to respond
- Translation is the delay in moving to the point of action
- Minimum Engagement Window is the time needed to apply the action effectively.

The equation is illustrated in Fig. 4.

Examples of volatile opportunities include:

- Ballistic Missile attack - Timing and duration of intercept window
- Train in a tunnel - Timing and duration of sabotage window
- Ship changing head - Anticipation of optimum torpedo release point.

Volatility is two measures in one - the period until the opportunity window and the duration of that window. The first reduces with time, the second is notionally constant. The elements on the right of the equation can be estimated from situation, derived from queuing theory and organisational knowledge or - in the case of translation - calculated directly.

$$\frac{\textbf{Translation}}{\textbf{Delay}} = \frac{\textbf{Separation}}{\textbf{Relative Closure Rate}} \quad ... (2)$$

This apparently trivial equation takes on new dimensions when intercepting a bomber before it releases its stand-off weapons, sending a naval surface action group through a stretch of the South Atlantic or traversing a minefield to reach a vital sabotage target.

QUEUING AND THE C^2 LAWS OF RATE EQUALITY

Queuing Theory is a well-established technique used by, amongst others, E. Ruiz-Pata (3) and J. Martin (4), which stands apart from any C^2 Theory in its own right. However, Queuing Theory may lend insight as follows. Figure 5 shows a typical C^2/Action Element organisation, with more-or-less random arrivals, at service stations, of decisions to be taken, plans to be implemented, action elements to be replenished, etc. Simple Queuing Theory is used for steady state analysis; one of the characteristics of steady state for a serial chain of service stations is that the mean arrival rate must be identical at each service station. In the figure, this unique arrival rate is called 'λx'. The figure shows 'λx' plans from "Initiate Action" pairing (meshing) with 'λx' Action Elements from "Replenish/Rearm". 'λx' reports of the action result in 'λx' option adjustments per period - and so on. Both the figure and the concept are reminiscent of a heart which beats at a rate of 'λx', sending actions (arterial blood) pulsing around one loop and plans (venous blood) pulsing around the other, in strict synchronisation. Mean planning and action rates must similarly be identical in the steady state. The figure, which omits overt enemy effects, inspires the following Law of Order/Action Rate Equality.

C^2 Law of Order/Action Rate Equality		
The sustained rate of generating C^2 Action Orders	**=**	**The sustained rate of regenerating Action Elements**

A little thought suggests that the law is axiomatic, but it does not appear widely as a factor in designing C^2 systems.

Figure 5 contains other useful concepts. It is, for example, straightforward in a practical case to calculate the number/proportion of action elements permanently "lost" in the recover - repair - replenish cycle, in the sense that they are never available for action.

Care must be taken in the application of the above Law, in terms of the granularity of "action element". The simplest definition is that one Action Element is the smallest unit for which the relevant C^2 HQ provides action plans. That unit might be a pair of fighters (if they work in pairs), a Surface Action Group comprising several ships or a complete Corps, according to level of viewpoint and command.

The Queuing Model suggests one other important Law, the Law of Action/Reporting Rate Equality.

C^2 Law of Action/Reporting Rate Equality		
The sustained rate of regenerating Action Elements	=	**The mean C^2 Reporting Rate**

This Law may also seem self evident, but is rarely if ever observed in design. Instead, reports to higher authority are sent either at set and frequent intervals to ensure "fresh" information (even though higher authority may not be able to use the information at the time received), or are responses to unpredicted events. Note that reporting in the context of this Law has to provide Requisite Variety (see below); in addition to mission results, recovery rates, damage, repair prediction, replenishment levels of POL and War Consumables all vary as action elements go through their cycle, and are all candidates for reporting at 'λx' (mean) reports/period.

The Law of Action/Reporting Rate Equality is particularly valuable because it provides a basis for calculating the theoretical communications capacity which must exist between elements in a system. If we can define the volatile information which must be transferred, then this Law gives the mean rate of transfer of that information.

DECISIONS AND REQUISITE VARIETY

Command and Control is based upon structured decision formulation. Information is gathered, analysed and used to make decisions which result in actions. This decision process has been characterised as the SHOR paradigm in J.G. Wohl's paper Force Management Decision Requirements (5) - Stimulus - Hypothesis - Option - Response. Decision-taking can be expanded into a cyclic process as shown in Fig. 6, where the decision cycle is shown at the top. The lower part of the figure shows the tabulated decision cycle elements at the left hand column and the conventional organisational C^2 partitions at the top row. The cells in the table indicate some of the contributions these partitions make to various elements in the decision cycle. Even at this superficial level it becomes apparent that structured decision-taking must account for a wide variety of factors and influences. There is, in Control Theory, the Law of Requisite Variety; the Law, from W.R. Ashby in Introduction to Cybernetics (London; Chapman Hall, 1956), states:

> "Given a system with a regulatory process R, intended to maintain goal state G, but affected by disturbance D: The goal state G can only be maintained if the regulator R has sufficient variety and channel capacity to counter the variety in D."

For C^2, this means that the Commander must be able to match and react to anything that could affect efficient operations. The value of this law to an appreciation of C^2 cannot be overstated.

Requisite Variety for Control must include:

- Operations
- Intelligence (by definition, information about the enemy or opposition)
- Logistics, especially war-consumable stocks, reserves and reinforcements
- Transportation, particularly for operations and logistics support
- Engineering, both in support of operations and to maintain C^2 facilities and communications.

THE COMMANDER'S CONTROL MECHANISM

There is a natural inclination to consult the general field of control theory for an answer to the control part of C^2. The promise of a ready-made C^2 theory exceeds the practice, however, for a variety of reasons:

- The transfer functions within C^2 are not time-invariant; they vary because they represent human behaviour and also due to enemy attack. As a result the system is essentially non-linear.
- Methods of solving such complex systems as in Shinners' Modern Control System Theory and Application (6) (e.g. State Variable Analysis), while ideal for computers, obscure the observable physical entities in a C^2 system. As Fig. 7 shows, even a simple representation of C^2 becomes an eighth-order (or higher) differential equation and the routine reduction of such equations to sets of first order equations - while easing solution - detracts from the ability to formulate a C^2 theory.
- The "Russian Dolls" concept is difficult to introduce. In Fig. 7, for example, operations control would tend to contain the C^2 element of Fig. 7 complete, replicated at a smaller scale. Further replications and scale changes should appear in even greater detail and, conversely, Fig. 7 is only part of a greater whole.

Nevertheless, there are lessons to be learnt. Figure 7 raises the spectre of stability of the control system. Simplistically, the system is potentially unstable when:

$$G_2 H_2 + G_3 G_4 H_3 + G_2 G_3 G_4 (H_2 H_3 + H_4) + G_1 G_2 G_3 G_4 H_1 = -1$$

(In practice, non-linear systems display a much more complex stability pattern.) (6). Simplistically, again, we can assess the general form of the various transfer functions and - for small inputs at least - a piece-wise linear approximation might be found to particular solutions. These are a standard analytical processes which will not be pursued further here.

C^2 STRUCTURE, BALANCE AND FLOW

Relationships, Links and Interfaces

Using the decision cycle and contribution tables of Fig. 6 leads to the idea of relationships between the partitions in a C^2 unit. Figure 8 shows an N^2 relationship diagram for a notional C^2 unit somewhere in the middle of a C^2 hierarchy, i.e. the unit has superiors and subordinates. These other features are alluded to by the dotted arrows outside the box; lines out of the right and in at the bottom have passed through subordinate formations, while lines out at the left and in at the top have passed through superior formations.

The N^2 chart is an invaluable device for representing relationships or interfaces. These interfaces represent infrastructure, and the analysis of N^2 charts will prove below to be a rich hunting ground for C^2 theory of infrastructure.

SPAN - THE STRUCTURAL BUILDING BLOCK

Span of control/understanding denotes the number of directly reporting subordinates; it will be symbolised as 's'. s is usually indicative of individual human relationships, but can also be used to indicate, for instance, the number of subordinate formations reporting to, and controlled by, a C^2 HQ.

For humans, research has shown that s is sharply bounded. Where one person has to control a number of others, experiments show that the subordinates split into groups if their number is seven or greater. There also seems to be some deep-rooted physiological difficulty in accommodating seven or more objects simultaneously.

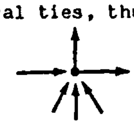
The maximum value of s can be affected by several factors. s_{max} will be less than seven where the subordinates cover such a wide range of activities, skills and knowledge that the Commander would be stretched beyond his ability to understand. Alternatively, if some or all of the subordinates were very similar in their relationship, then the Commander - given sufficient time - might service each subordinate adequately. A Commander might, for instance, control a dozen identical missile batteries. He might be unable, on the other hand, to effectively control three subordinates where one was concerned with Intelligence one with Logistics and one with Engineering, due to their diversity.

THE PYRAMID

Using only the concept of span, it is possible to construct a pyramidal organisation as follows:

s

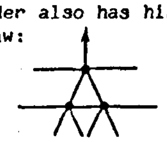
The figure represents a node, with 's' subordinates (span of control/understanding = s), reporting upwards to one superior. Many C^2 organisations also include lateral ties, thus:



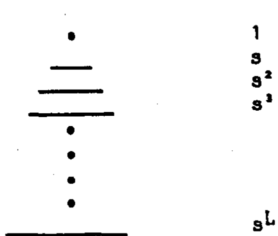
s

The purpose of these cross-ties, mainly to speed C^2 reaction time, will be examined more closely later.

Each Commander also has his Commander. For s = 2, we could draw:



At each node, the span s is two and there is one superior. This structure represents a military organisation archetype, which proves to be rich in providing insight into C^2 theory. Drawing links and nodes can become tedious, however, especially when s is large. The organisation archetype can be represented thus:



A Commander has 's' subordinates, each of whom has 's' in turn giving s^2, and so on. At the bottom of the pyramid there are s^L subordinates, where L is the number of levels. For an organisation of individuals, we have:

$$\text{Total Personnel} = \sum_{n=0}^{L} s^n \qquad \ldots (3)$$

L corresponds simply to the number of effective ranks in a military organisation. The equation is graphed at Fig. 9, from where it is possible to compare actual force numbers, rank levels and $\bar{s}$, the mean span of control/understanding.

The pyramid organisation is in many ways an ideal for military organisations. Most organisations can be represented as pyramids, but the need to identify clear lines of command and to be able to transfer and reform groups swiftly into new organisations leads to standardised pyramidal structures of a more formal kind for military purposes.

THE C^2 LAW OF EQUIPARTITION

When the span 's' would otherwise exceed an acceptable value, a Commander has to divide his subordinates into groups, each with their own leaders who report to the Commander. Is there any particular rule about the way in which the division process should be undertaken? There is a basis for division into equal numbers so that twelve subordinates would be divided into 4 + 4 + 4 rather than 2 + 4 + 6. The reason is to minimise unnecessary infrastructure.

Figure 8 showed an N^2 chart with N(N-1) interfaces; these interfaces represent potential relationships or infrastructure. For N=12, the number of interfaces:

$$\begin{aligned} I &= N(N-1) \\ &= 12 \times 11 \\ &= 132 \end{aligned}$$

Divide the leading diagonal elements of an N^2 chart into two groups, each self contained, with all its internal interfaces, but with only two connections between the groups to represent commander-to-commander interfacing. Let one of the groups have n elements: then the other has N-n.

The total number of interfaces is now given by.

$$I = (N-n)(N-n-1) + n(n-1) + 2$$

Simplifying:

$$I = N^2 - 2Nn - N + 2n^2 + 2$$

$$\frac{dI}{dn} = -2N + 4n = 0 \text{ for a minimum value of I}$$

$$\therefore n = \frac{N}{2}$$

This result indicates that I is a minimum, when the subordinate elements are split into two equal groups:

- For N = 12, I = 132 when the span is also 12, as above
- For N = 12, I = 6(5) + 6(5) + 2 = 62 when the group of elements is split into two equal groups; this is less than half of the potential interface complexity, and is indeed the minimum value.
- If the partitioning is asymmetric, we have for example:

$N = 12, I = 4(3) + 8(7) + 2 = 70.$

It is a simple extension to show that the partitioning rule is true for three or more groupings too. Using N = 12 again, and splitting into three groups, gives:

Equal Groups: $I = 4(3) + 4(3) + 4(3) + 6 = 42$

Unequal Groups: $I = 3(2) + 4(3) + 5(4) + 6 = 44$
$I = 2(1) + 4(3) + 6(5) + 6 = 48$

Clearly, greater asymmetry leads to greater infrastructure. At this point we may formulate a new C^2 Law.

C^2 Law of Equipartition

"Partitioning subordinates into equal groups minimises infrastructure."

C^2 INFRASTRUCTURE

Figure 10 shows an N^2 representation of a three-level pyramid organisation with one superior Commander and three subordinates, each with their three subordinates. Because each Commander occupies one square in each block of squares, we can dispense with 'N' and use the more direct 's' where $N = s + 1$

- Each block has a side of length $(s+1)$
- The interfaces, I, per block $= s(s+1)$
- Each Commander has 2s direct interfaces with his immediate subordinates
- Each Commander has $2(s-1)$ direct interfaces with his immediate peers.

Since every Commander within a hierarchy finds himself with subordinates and superiors, we can add up the "local" interfaces with which he interacts directly.

Subordinates $= 2s$

Peers $= 2(s-1)$

Superior $= 2$

Commander's Interfaces = 4s ... (4)

So, for example, when $s = 3$, a Commander has 12 interfaces, or 6 direct two-way channels.

This identification of interfaces can be extended to consider indirect interfaces too, such as those between a Commander's immediate subordinates. Indirect interfaces are important because they provide information through intermediate nodes, and allow organisation flexibility. In a well designed organisation, all the indirect interfaces will be active to some degree; this is characteristic of tight functional binding. The sumation of direct and indirect links is defined here as infrastructure.

Infrastructure can be viewed in (at least) two lights, as Fig. 11 shows. On the left is a locally connected pyramid, and as the section though A-A' shows there is complete connectivity between subordinates of their immediate Commander. From Fig. 10 it can be seen that this degree of local connectivity, represented by the central block with Force 2 and Subs 2a-c, gives S(S+1) interfaces. At any level, there will be several such clusters. The locally-connected infrastructure amounts to:

$$I_{LC} = \sum_{n=1}^{N} s^n (s+1) \qquad \text{... (5A)}$$

where N is the number of subordinate command levels

I_{LC} is the locally connected infrasture.

The right hand of Fig. 11 shows the same pyramid organisation, but totally connected at each level as shown by the section through B-B'. Clearly there are many more connections, as might be expected in an "all-informed" communications network. In this case the infrastructure amounts to:

$$I_{TC} = \sum_{n=1}^{N} s^n(s^n+1) \qquad \text{... (5B)}$$

where I_{TC} is the totally connected infrastructure.

The two equations, 5A and 5B, tabulate as follows:

Subordinate Command Levels	I_{LC}			I_{TC}		
	Span			Span		
	3	5	7	3	5	7
1	12	30	56	12	30	56
2	48	180	448	102	680	2 506
3	156	930	3 192	858	16 430	120 498
4	480	4 680	22 400	7 500	407 680	5 887 700

A high correlation (1.00) exists between these sets of figures for command levels and infrastructure, using the relationship I = B exp (M*L), where I is the infrastructure, B and M are constants. From the table, we may therefore formulate the C^2 Law of Infrastructure Expansion.

C^2 Law of Infrastructure Expansion

At any point in a command hierarchy, infrastructure complexity increases exponentially with the number of subordinate command levels controlled from that point.

INFORMATION STRUCTURE

The decision cycle may be stylised as a circle. See Fig. 12, in which the decision cycle can be thought of as being rotated by incoming information. The centre figure shows the decision cycle as part of a layered organisation, and diagram C reintroduces the pyramid organisation, with $s = 2$. The structure presents many overlapping, interlocking control loops, but some steady-state characteristics can be perceived.

For an information-decision-action organisation to be effective, there must be a vertical (upwards) flow of information. The Commander at the top of the pyramid cannot be interested in low level detail, which would overload him with irrelevancy. It follows that low level data must be discarded, and aggregated as it passes up the chain of command. The degree of compression must enable intermediate commanders to cope with the data from their subordinates, and to pass up the chain that which the next level needs. Too much compression will isolate the senior Commanders from essential data: two little will overload them. This appreciation leads to the following axiom:

C² Law of Vertical Data Compression

The degree of vertical data compression in a C² hierarchy	=	**The Span of Control/Understanding**

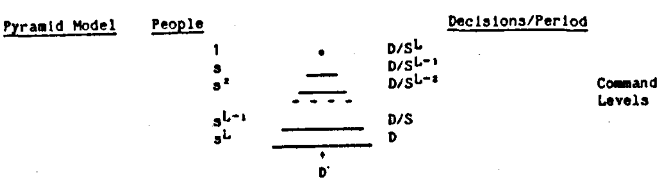
This law forms the basis for some interesting concepts, which follow below.

MEAN PYRAMID DECISION DELAY

The Law of Vertical Data Compression can be seen at work in establishing the mean decision delay through a pyramid organisation.

Mean Delay/Decision = Mean levels Traversed/Decision x Mean Delay/Level

A simple model for this decision process is shown below; 'D' external decisions enter at the bottom of the pyramid.



At each level, Commanders refer 1/s of their decisions upwards.

Consider an intermediate level; s^n people make $\frac{D}{s^{L-n}}$ decisions/period.

Decisions/person/period $= \frac{D}{s^{L-n}} * \frac{1}{s^n} = \frac{D}{s^L}$ decisions/period, i.e. invariant with the level considered.

Mean Levels/Decision

$$\text{Mean Levels Traversed/Decision} = \frac{\text{Total Steps Taken}}{\text{Total Decisions Referred}}$$

- D decisions take no steps
- D/S decisions take one step up and down again
- D/S² decisions take two steps up and down again
- Total steps $= 2 * \sum_{0}^{L} D * \frac{n}{s^n} = 2D \sum_{0}^{L} \frac{n}{s^n}$
- Total Decisions Referred $= D \sum_{0}^{L} \frac{1}{s^n}$

$$\therefore \text{Means Levels traversed/decision} = \frac{2 \sum_{0}^{L} \frac{n}{s^n}}{\sum_{0}^{L} \frac{1}{s^n}}$$

Mean Delay/Level

From queuing theory, References C and D, assuming exponential and independent inter-arrival and service times

$$w = \frac{1}{\mu - \lambda}$$

Where: w is the queuing and servicing delay, the waiting time

μ is the mean service rate

λ is the mean arrival rate

μ is a physiological constant for most people - their maximum sustainable decision rate.

λ is also a constant; $\lambda = \frac{D}{s^L}$: D is an externally driven need for decision and, is assumed constant regardless of C² hierarchy.

s^L = AE, the number of action elements, also assumed constant regardless of C² hierarchy.

$\therefore$ w is invariant with changes in span of control. This result is particularly surprising and is worth confirming, as follows:

Span = 2		Span = 4	
People	Decision Rate	People	Decision Rate
1	4	1	4
2	8		
4	16	4	16
8	32		
16	64	16	64
D = 64 +		D = 64 +	

The left hand of the panel shows s = 2, the right hand s = 4.

For the left, command levels, L = 5, for the right L = 3; D and AE (Decision Rate and Action Elements) are the same, 64 and 16 respectively. The boxed numbers show the picture. In each case at intermediate level, 16 decisions/period coincide with 4 persons giving 4 decisions/person/period. Hence the general conclusion that w is invariant with span is exemplified, and forms the basis for the C² Law of Decision Rate Invariance.

C² Law of Decision Rate Invariance

C² Decision Rate is invariant with changes in mean Span of Control/ Understanding, all factors external to the C² structure remaining constant.

The full expression for Mean Delay/Decision can now be formed.

$$\text{Mean Delay/Decision} = \frac{2 \sum_{0}^{L} \frac{n}{s^n}}{\sum_{0}^{L} \frac{1}{s^n}} * \frac{1}{\mu - \frac{D}{AE}} \quad \ldots (6)$$

This expression is somewhat intractable, but responds to numerical analysis as follows:

For $\mu = 7$

Span	AE = 100						AE = 100 000					
	D/AE = 3.5			D/AE = 5			D/AE = 3.5			D/AE = 5		
	L	N	T	L	N	T	L	N	T	L	N	T
2	7	1.89	.54	7	1.89	.95	17	2	.57	17	2	1
3	4	.96	.27	4	.96	.48	10	1	.28	10	1	.5
4	3	.64	.18	3	.64	.32	8	.67	.19	8	.67	.33
5	3	.45	.13	3	.45	.23	7	.5	.14	7	.5	.25
6	3	.37	.11	3	.37	.19	6	.4	.11	6	.4	.2

Where: L is the number of Command Levels

N is the mean number of nodes traversed per decision

T is the mean time per decision (arbitrary units)

AE is the number of action elements

D is the externally driven rate of decisions required

μ is the sustainable decision rate per node.

As the table shows:

- Increasing the span reduces the decision delay
- Increasing the decision rate increases the delay time (D/AE = μ results in infinite delay)
- The mean number of nodes traversed varies inversely with span
- The mean number of nodes traversed is virtually independent of the number of action elements (and hence of C² organisation size).

C² Postulate of Decision Scale Invariance

For a given mean span of control, the mean number of nodes traversed by a decision is independent of the C² organisation size.

CO-OPERATION

Within a hierarchy, tasks are achieved largely as a direct result of co-operation between peers, that is personnel at the same level within the hierarchy. So, while decisions may be said to invoke vertical information flow, co-operation - it might seem, at least - moves laterally.

Within the pyramidal organisation, peers are separated by organisational groupings. Peers who report to the same superior may be said to be separated by a factor of one. Peers whose superiors themselves report to the same superior in turn may be said to be separated by a factor of two. And so on. Separation, then, is simply the number of senior levels of command which must be traversed before one superior, common to both peers, is found. For co-operation between peers to be effective, authorisation will become necessary as plans and propositions are developed and, generally, that superior who is common to both peers will be that authority. Clearly, the greater the separation between any two peers, the greater will be the number of steps in the authorisation chain before the sought after authority can be forthcoming.

At this point we may invoke either contingent probabilities or, more simply, Occam's Razor. William of Occam (1300-1349) expounded the principle that assumptions introduced to explain something must not be unnecessarily multiplied. The longer a decision/authority chain, the greater the number of steps in reaching agreement and the greater must be the number of multiplied assumptions of agreement. From this point, we may formulate the following law:

C² Law of Lateral Co-operation

In any C² Hierarchy, the effectiveness of co-operation between peers is inverse to their lateral separation within the hierarchy.

This law, all too evident in action in peacetime, does not refer to willingness on the part of peers. Nor should it be taken as suggesting that across-boundary co-operation is not viable. The so-called law does, however, bring into focus the effectiveness of their co-operation, resulting from the implicit infrastructure that their would-be co-operation involves. The law also goes some way to explaining why co-operation seems more effective between senior people - their lateral separation cannot be as great as that of their juniors, since the pyramid is narrower for more senior ranks.

The Law of Lateral Separation is also the one we apply, perhaps instinctively, when we design new organisations and try to group together under one 'umbrella' all of the key components for achieving the tasks we can anticipate. It is the tasks we cannot anticipate which subsequently require co-operation, and it is the Law of Lateral Separation which shows, perhaps more than any other, the weakness of rigid hierarchical organisation.

INFORMATION MODELS

At this point in the analysis, sufficient groundwork has been established for an Information Model. From Queuing Models we have the Law of Action and Reporting Rates Equality to provide Information Rate data. The Law of Vertical Data Compression provides a vertical flow viewpoint. The C² Law of Lateral Co-operation provides a lateral flow viewpoint. The C² Law of Infrastructure gives a view of the complexity of communications and relationships; and so on.

Consider a General viewing an impending battle from a hilltop overlooking his own and the opposing forces. He has a variety of information needs, as follows:

- Location of his own troops
- Location of opposing troops
- Local terrain ('going') which might affect mobility, screening etc.
- Balance between his and opposing forces
- Relationships between his own force elements - interoperability, complementary capabilities, etc.
- Etc.

The list is long, but it is possible to categorise each item on the list, as

- Stable, e.g. Interoperability
- Dynamic, e.g. Movement of enemy forces
- Deducible, e.g. Balance of forces
- Doctrinal, e.g. Keep cavalry in reserve as Force Multiplier.

Such categorisation is useful, because it provides a basis for calculating amounts of stored information and communication system capacities; only the dynamic or volatile data needs to be communicated.

The General would like to know as much about the opposing forces as he should known about his own. Ideally his storage and communication needs should be doubled to accommodate enemy intelligence, except for some deducible factors which - again ideally - should be the same, given full information.

The General is interested, before and during combat, in both performance and in the continuing availability of that performance. We can therefore start to categorise information as shown in Fig. 13, which is evidently not comprehensive, and in particular we can identify volatile or dynamic data as already seen in Fig. 12.

The dynamic model brings to mind information theory and entropy. From information theory as used by Schwartz (7) we have:

$$\text{Information} = \frac{T}{\tau} \log_2 n \text{ bits in T seconds}$$

Where information is a measure of conveyed content:

- T is the period of conveyance of a message
- τ is the time interval (<T) for passing/measuring parameters within the message
- n is the number of variable states, all assumed to be equally probable which can be distinguished in each period τ.

We can therefore conceive an ideal C² system in which message probabilities are indeed all equal. We already have an expression for τ, from queuing: τ is the inverse of the Action Element Regeneration Rate, λx, which in turn equals both the mean Action-order Rate and the Mean Reporting Rate. Then:

$$H = -T * \lambda x * \log_2 n \qquad \ldots (7)$$

Where H is the maximum information content per message (bits)

n is the number of distinguishable levels within unit time (granularity/quantisation)

λx is the mean Action-order Rate

and the mean Action Element Regeneration Rate

and the mean Reporting Rate.

similarly $$C \geq -\lambda x * \log_2 n \qquad \ldots (8)$$

Where C is the minimum required channel capacity in bits/unit time/message.

Note: Reference to Fig. 5 will show that the use of λx is inappropriate for activities not tied to the Action Element Regeneration Cycle, such as Intelligence Collection.

Using these formulae, there is no obvious shortcut to identifying quantities of stored information and rates of conveyed information; the process is one of adding H and C respectively for appropriate categories of information. The crucial parameter is granularity.

Command Granularity

As we have seen from the section on Infrastructure, there is a potential for information explosion if a Commander controls several subordinate levels directly. So, should a Group Captain control directly to wing level, to squadron level, to flight level or to section level? Should a Field Marshall concern himself directly with Armies, Corps, Brigades or Battalions? And so on.

Figure 14 addresses the problem. Each C² HQ falls into a three-layer structure, which usually relates to rank in some measure. The layers are: executive, management and specialist staff. As the figure shows, there will be an overlap with subordinate C² HQs and the optimum granularity for a Commander therefore seems to be up to and including two command levels below his own. Less will not necessarily be effective; more will risk information and communications overload. These considerations lead to the following:

C² Postulate of Command Granularity

Ideally a Commander operates directly up to two command levels below his own

COROLLARY: Operating directly at more than two command levels below his own will risk a Commander (a) operating beyond his span of understanding (b) neglecting parts of his command.

Granularity and Thresholds

As granularity coarsens, the need to convey information may reduce dramatically because any changes occurring since the previous report have not been sufficient to exceed the granularity threshold.

This phenomenon can affect communication capacity in two complementary ways; first, the reporting rate changes, and second, the information content changes. Consider a requirement to report significant changes in visibility, with 10 000 metres or further being considered "clear". Suppose the Mean Reporting Rate, λx, is once per hour. Then applying Equation 8 gives:

$$C \geq \log_2 10\,000 = 13.3 \text{ bits/hour}$$

However, if the range of visibility is divided into 1000 m bands, and visiblity is observed to change at 200 m/hour, then the rate of reporting can be less than λx; it need only be 1 report per 5 hours. The number of levels to be reported is no longer 10 000; it is only 11. Thus we can calculate:

$$C \geq \frac{1}{5} \log_2 11 = 0.69 \text{ bits/hour}$$ assuming all levels to be equally probable

This is a capacity ratio of nearly 20:1 designers evidently have quite extraordinary scope to influence capacity by controlling granularity and this approach can be applied to most items to be communicated - Fig. 10, second panel.

Generally:

$$C \geq \frac{\dot{V}}{Q} * \log_2 \frac{R}{Q} \text{ bits/period} \qquad \ldots (9)$$

where:

C is the communications capacity in bits/period

$\dot{V}$ is the Rate of Change of the variable to be communicated (assumed noise-free)

Q is the Quantisation or Granularity with which the variable is measured

R is the range of the variable.

(Note: Volatile variables may be candidates for more frequent communication than at the rate λx; generally, there is no benefit in exceeding λx since the action-planning loop cannot use the information. In specific cases, changes in important variables may necessitate special priorities.)

Survivability

Had the models of C^2 in this paper been principally of structural groupings rather than of sequential functions coupled by information flow, it might have led to the conclusion that C^2 Survivability was concerned with physical entities. As the functional models have suggested; however, C^2 Survivability is concerned with the continuance of the complete chain of functions and information exchanges in the face of enemy activity.

It is, of course, possible to map between the functional C^2 organisation and its physical embodiment into Sensors, Sources, Action Elements, HQs, Communication Centres and Links. Clearly, loss of some of these physical elements can disrupt C^2 and it is part of the process of system design at the highest level to minimise the impact of physical damage on C^2 functional performance. At its most basic, this design process identifies functional nodes, defined here as singularities in the C^2 activity flow, and either replicates them to prevent their loss being catastrophic (which stops them from being classified as nodes), or protects them from being damaged. Protection ranges from hardening against attack or making elements mobile, difficult to detect and capable of evasion.

These considerations lead to the simple but vital:

C^2 Law of Survivability

C^2 Survivability is inverse both to the number of sequential, functional nodes in the C^2 process and to the degree of their exposure.

The Law addresses four points. First, if the C^2 chain is long and comprises many unique processes, it is more vulnerable than a shorter chain simply because the chance that damage will break the longer chain is proportionally greater. Second, each functional node is vulnerable by virtue of its uniqueness. Were a function to be replicated, it would be less vulnerable (and indeed it would no longer qualify as a node). Third, exposure indicates that the functional node (or at least its physical shelter) is detectable by an enemy. Fourth, exposure means also that the functional node's self defence is inadequate. It is implicit in the C^2 Law of Survivability that it does not refer directly to physical vulnerability; physical loss is only debilitating where it breaks the functional process. It is also implicit that survivability is not a constant; damage may be contained and repaired, functions may be restored, camouflage may be added, self-defences may be strengthened.

SUMMARY AND CONCLUSION

The paper set out to produce a set of C^2 postulates, axioms and laws. The following proposed C^2 Postulates and Laws have been formulated:

- Law of Order/Action Rate Equality
- Law of Action/Reporting Rate Equality
- Law of Equipartition
- Law of Infrastructure Expansion
- Law of Vertical Data Compression
- Law of Decision Rate Invariance
- Law of Lateral Co-operation
- Law of Survivability
- Postulate of Decision Scale Invariance
- Postulate of Command Granularity.

Additionally, equations have been produced which can be used to predict ideal values for:

- C^2 performance as a control mechanism
- Mean span of control
- Commander's direct communication links
- Commander's infrastructure links
- Mean hierarchy levels traversed per decision
- Mean hierarchy delay per decision
- Information content per C^2 information category
- Communications capacity variation with information granularity.

The equations can, hopefully, be used as:

- A benchmark
- In some cases, a warning to avoid combinatorial explosions
- A means of testing the Laws.

REFERENCES

1 Wiener, N., 1961, Cybernetics or Control and Communications in the Animal and the Machine (2nd edition), The MIT Press and Wiley, New York.

2 Roberts, N., 1983, A System Dynamics Modelling Approach, Addison-Wesley 0-201-06414-6.

3 Ruiz-Pata, E., Waiting Line Models, Reinhold, LCCCN 67-16189.

4 Martin, J., Design of Real Time Computer Systems, Prentice Hall, LCCCN 67-18923.

5 Wohl, J.G., 1981, Force Management Decision Requirements for Air Force Tactical Command and Control, IEEE Trans. SMC-11(9).

6 Shinners, S.M., Modern Control System Theory and Application, Addition Wesley, LCCCN 78-52497.

7 Schwartz, M., Information Transmission, Modulation and Noise, McGraw-Hill Kogakusha, LCCCN 70-98490.

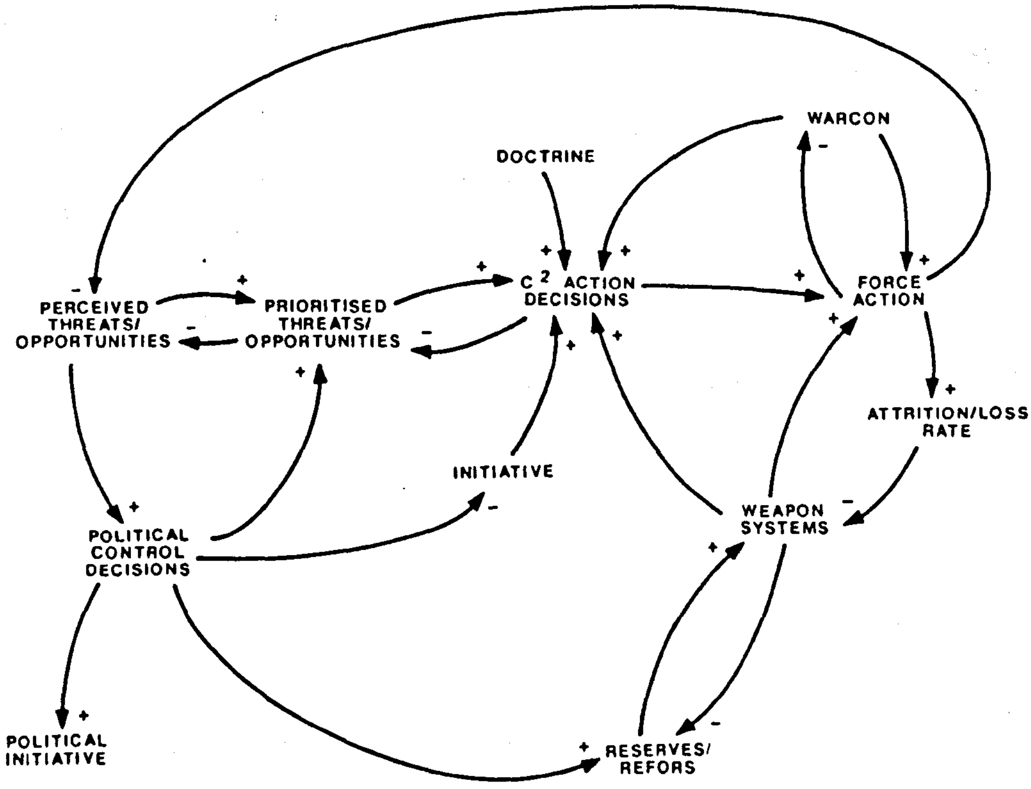


Figure 1 System Dynamics View of Command and Control

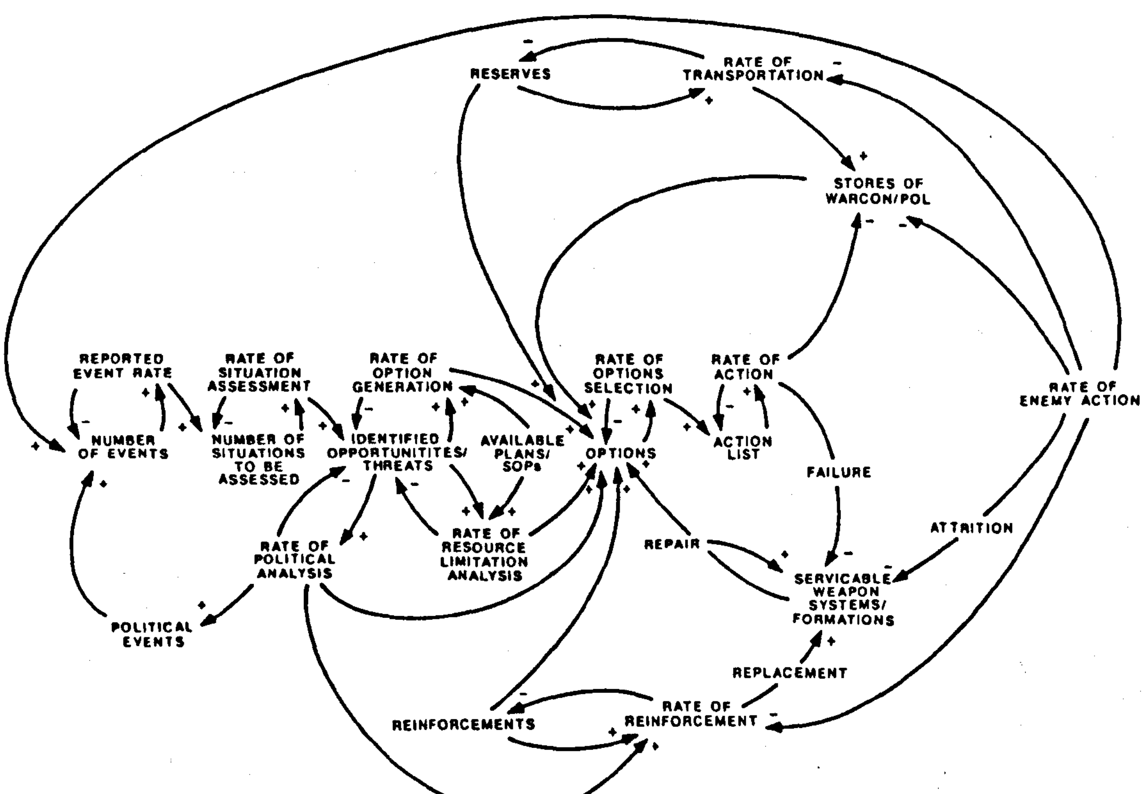


Figure 2 System Dynamics Representation of Delays in Command and Control

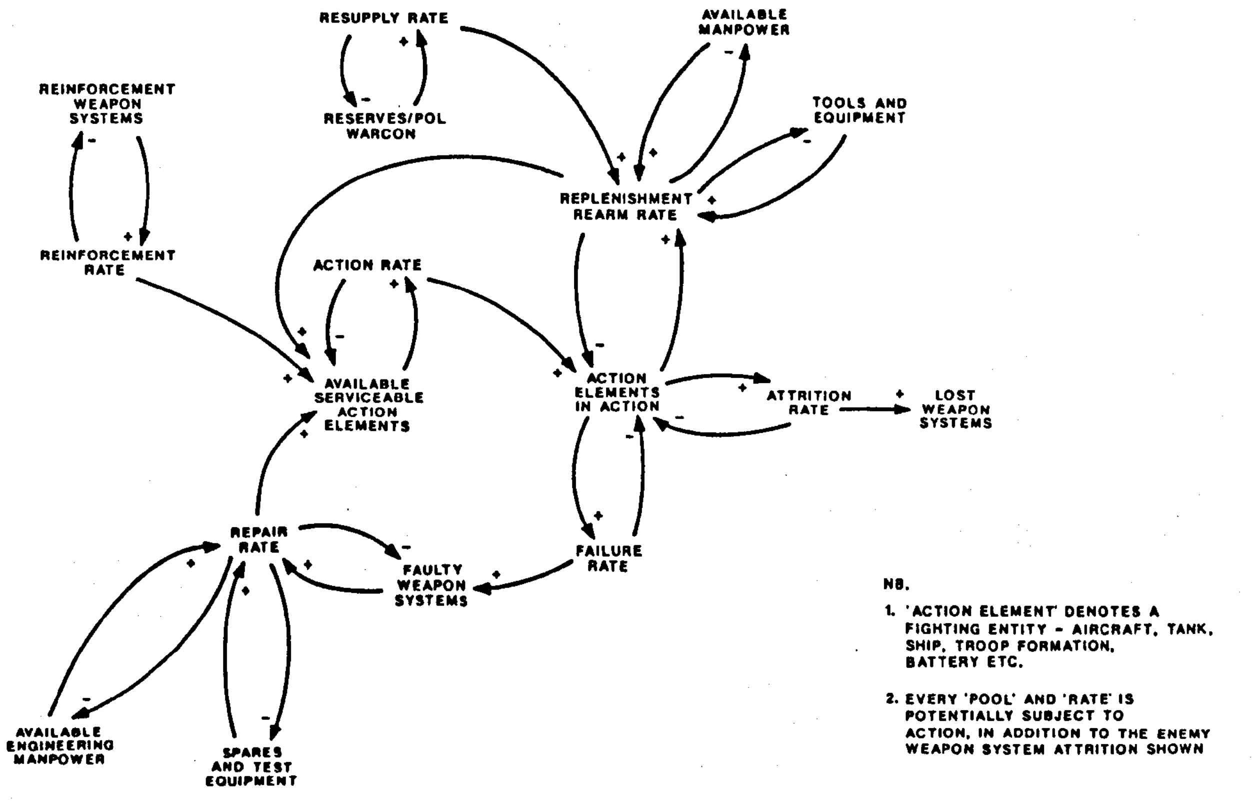


Figure 3 Dynamics of Action Element Availability

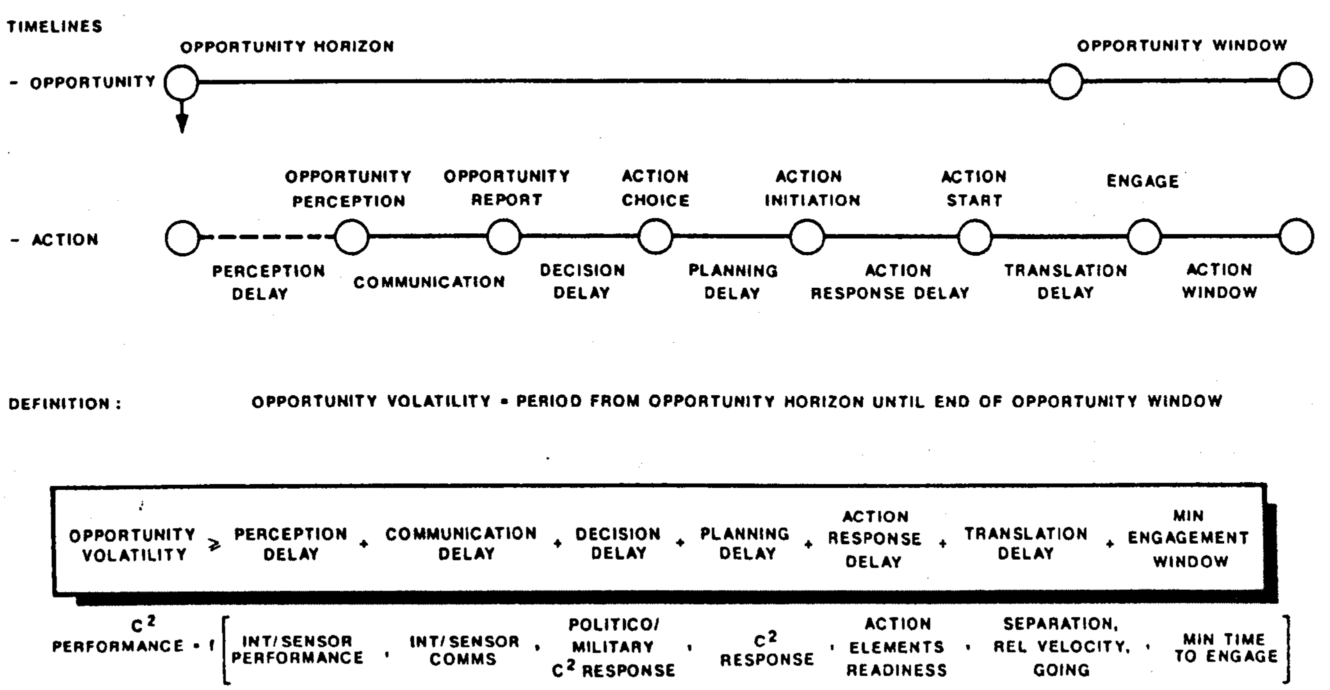


Figure 4 Basic Command and Control Timeliness

NOTES: 1 STEADY-STATE ONLY

2 UTILISATION, $\rho = \lambda/\mu < 1$

3 – λx IS STEADY-STATE ARRIVAL RATE WITHIN ALL LOOPS

– μ IS STEADY-STATE SERVICE RATE PER SERVICE STATION

4 k-ACTION ELEMENT FAILURE PROBABILITY

5 FIGURE ASSUMES SINGLE CHANNELS

6 LOOP DELAYS ARE CALCULABLE

7 ACTION ELEMENTS 'LOST' IN RECOVERY, RESTORE LOOP ARE CALCULABLE

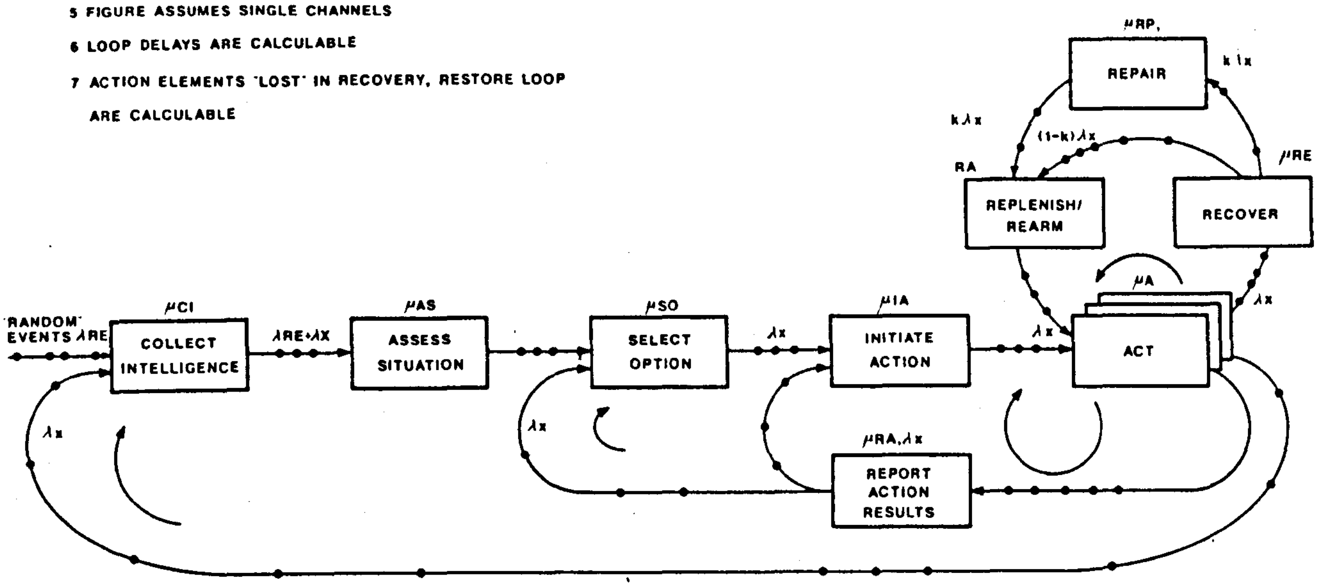
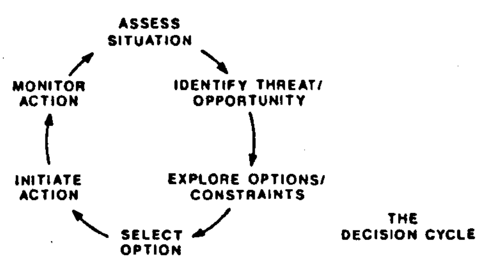


Figure 5 A Queuing Model of Command and Control



CYCLE ELEMENT	COMMANDER	INTELLIGENCE	OPERATIONS	LOGISTICS	ENGINEERING	OPS PLANS
ASSESS SITUATION	ANTICIPATION AND EXPERIENCE	ENEMY SITUATION AND INTENT	OWN FORCE DATA			
IDENTIFY THREAT/ OPPORTUNITY	JUDGEMENT	ANALYSIS OF ENEMY STRENGTH/ WEAKNESS	FORCE BALANCE			
GENERATE OPTIONS	INNOVATION AND EXPERIENCE	VARIETY OF FEASIBLE ACTIONS				OPTIONS IN DEPTH
REVIEW CONSTRAINTS	DATA FOCUS	ENEMY DEFENCES AND TACTICS	AVAILABLE WEAPON SYSTEMS	• STOCKS OF WAR • CONSUMABLES • TRANSPORT	• C^3I INFRASTRUCTURE • PREDICTED AVAILABILITY	IMPACT ON OTHER PLANS AND OBJECTIVES
SELECT OPTION	DECISION					
INITIATE ACTION	COMMANDS	ORDERS	ORDERS	INSTRUCTIONS	INSTRUCTIONS	REPLANS
MONITOR ACTION		COLLECT ENEMY DATA	COLLECT OWN FORCE DATA			ANTICIPATE OUTCOME

Figure 6 Command and Control Decision Formulation

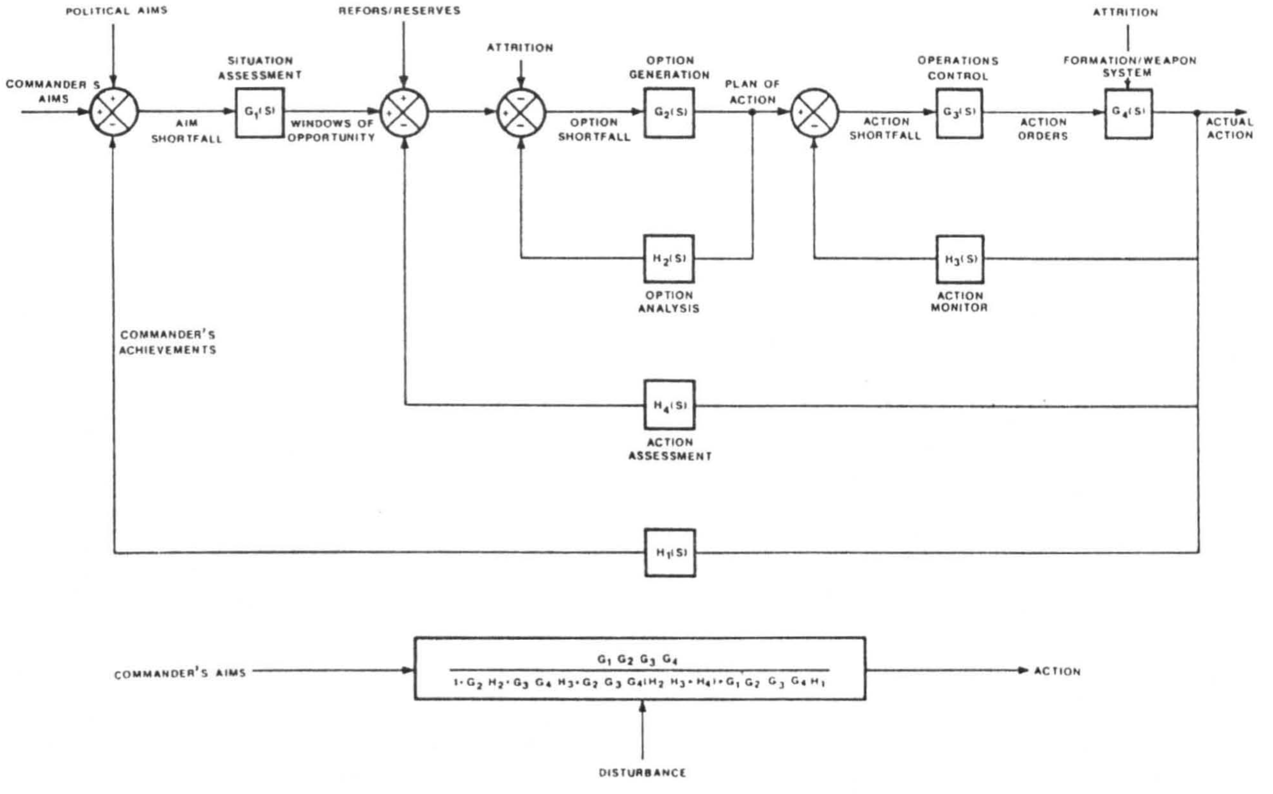


Figure 7 The Commander's Control Mechanism - Simplified

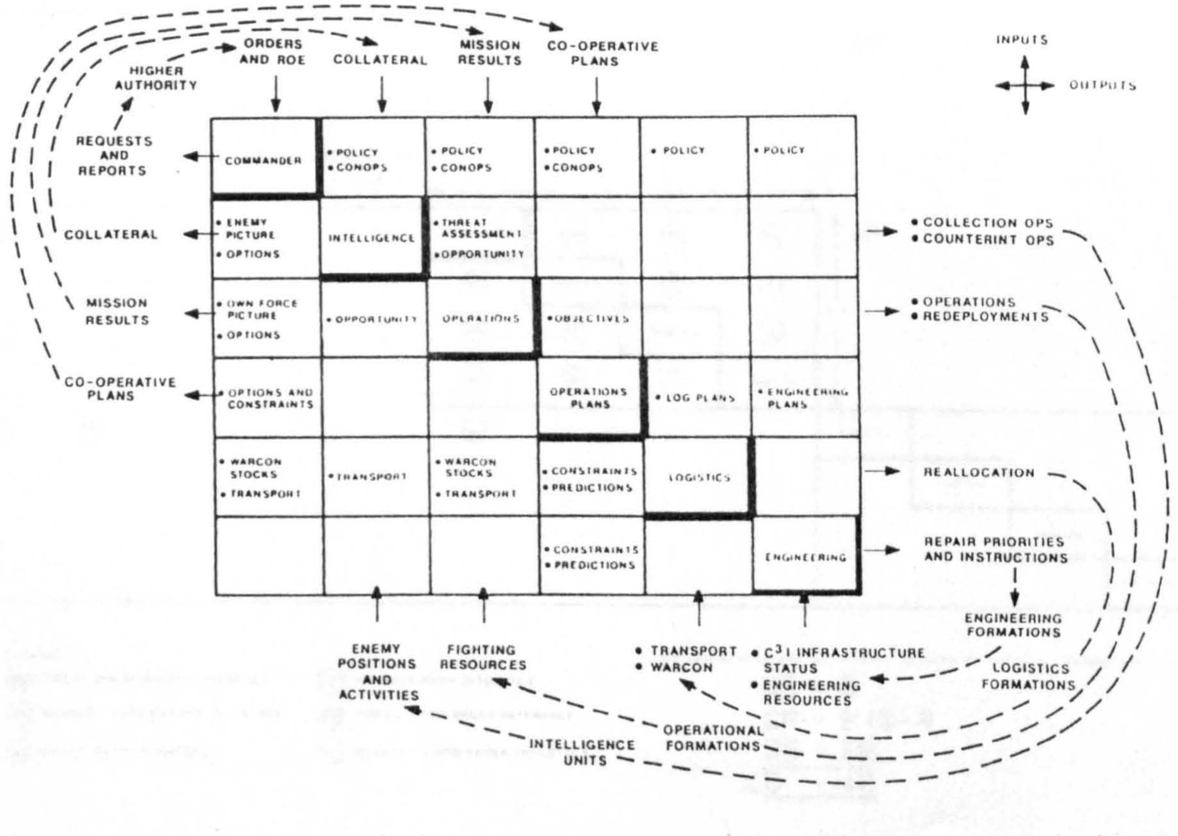


Figure 8 Command and Control Node and Requisite Variety

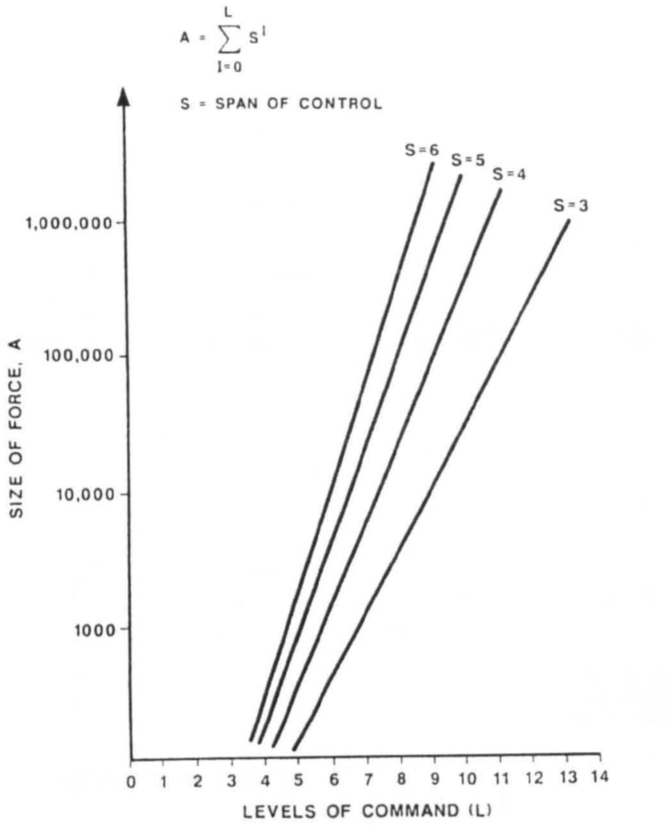


Figure 9 Graph Showing (Simplified) Relationship Between Size of Force and Span of Control

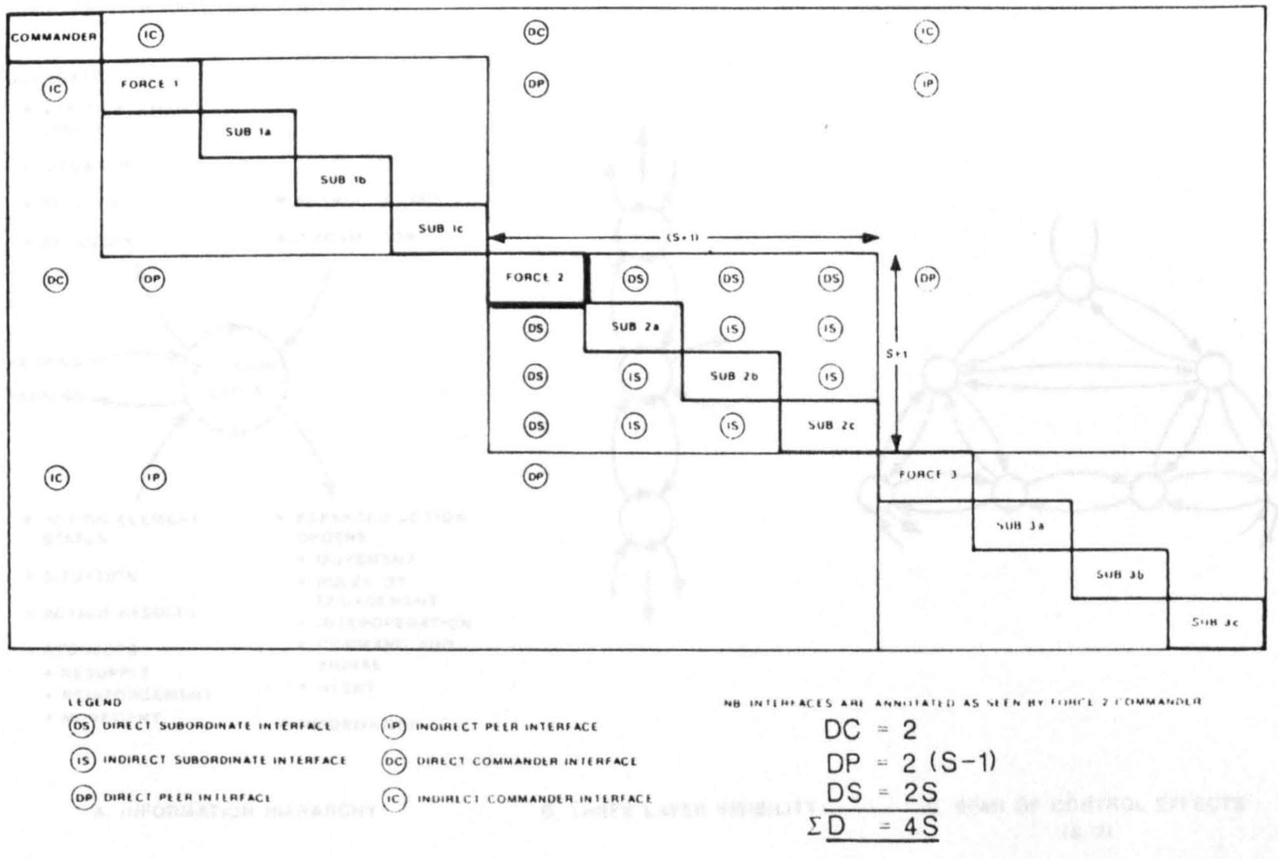


Figure 10 Interfaces

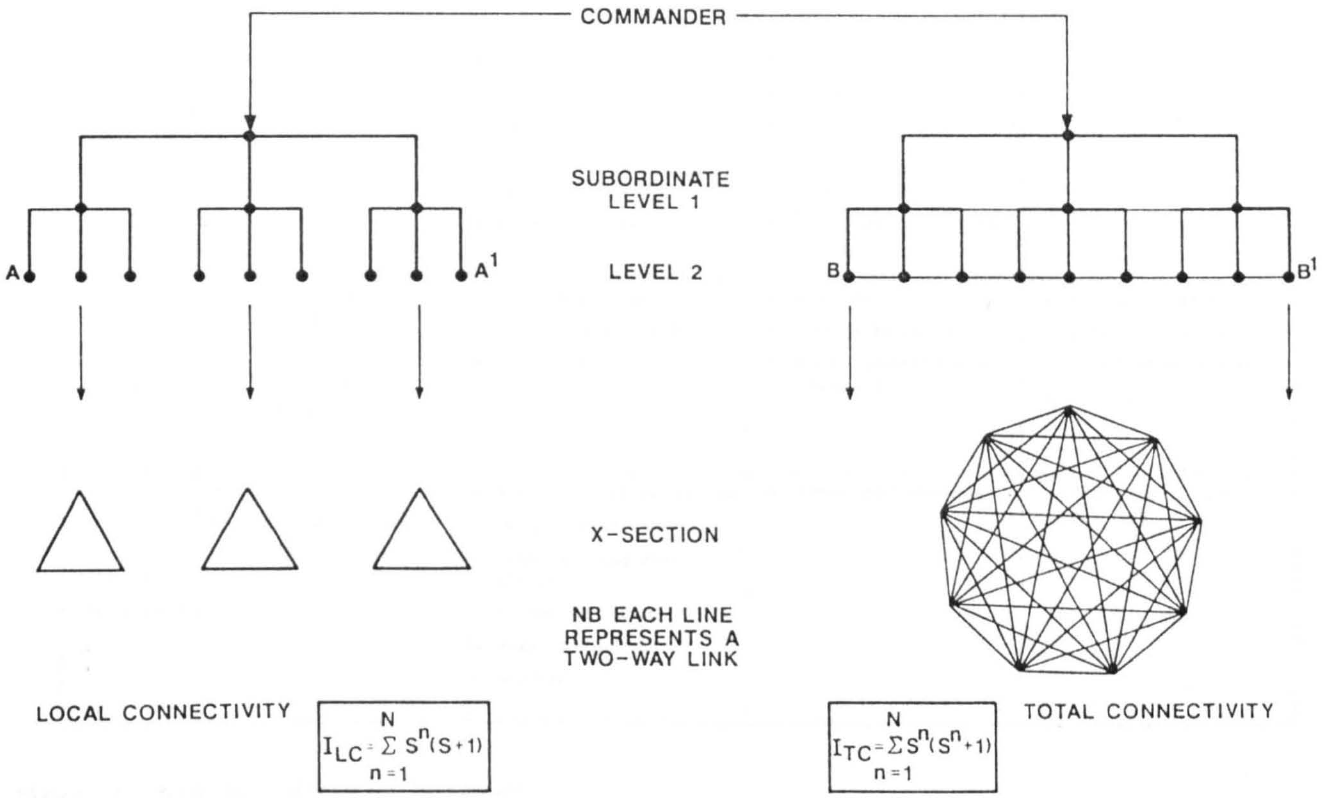


Figure 11 Infrastructure Expansion

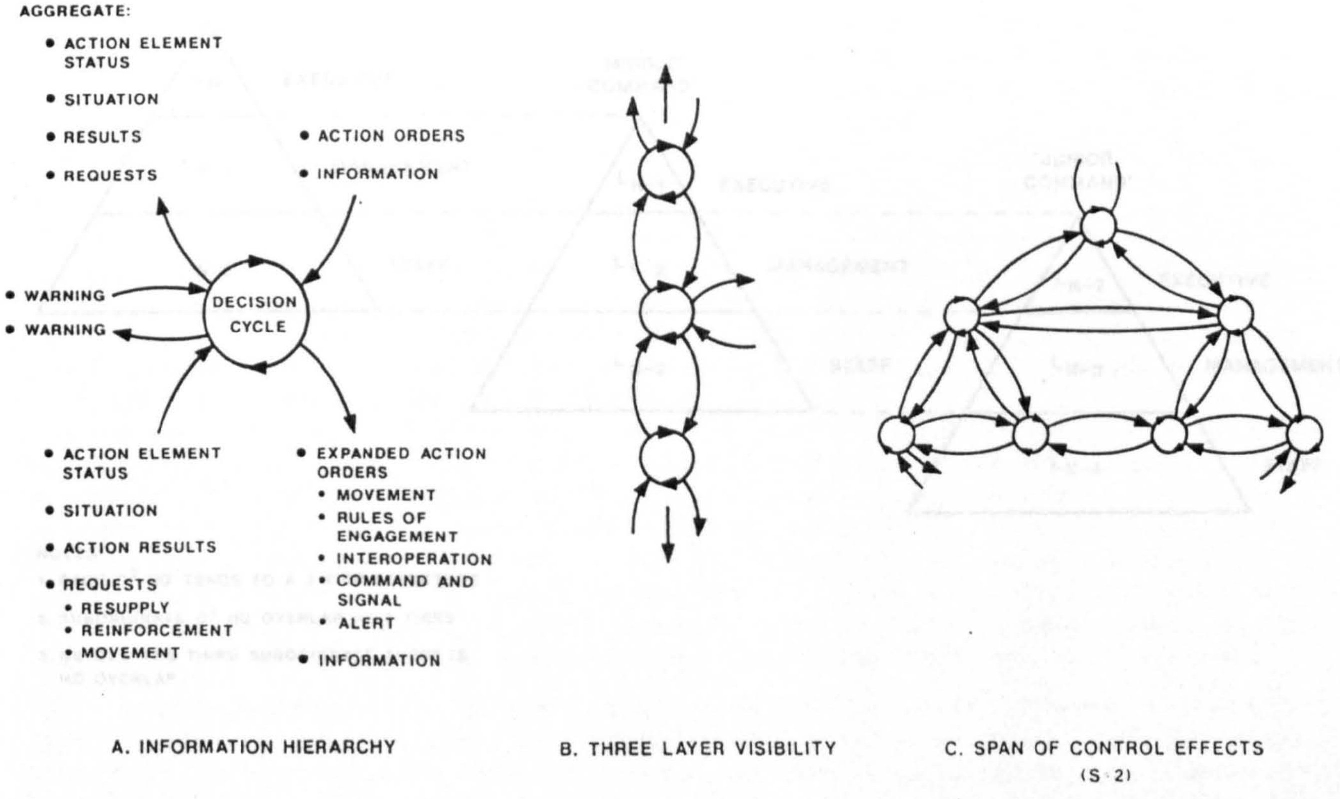
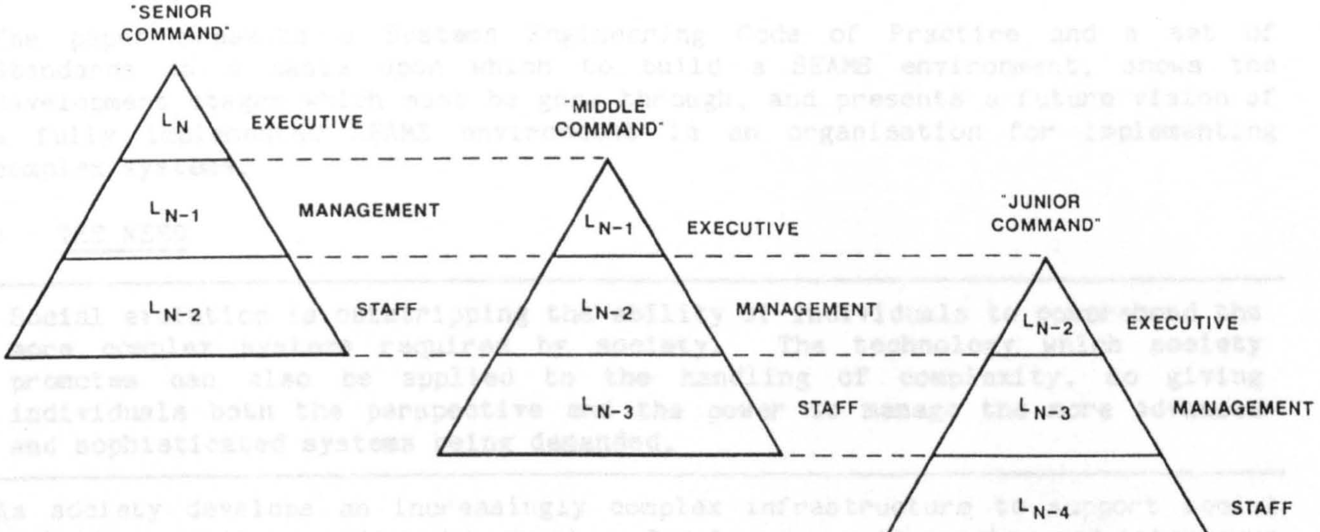


Figure 12 Information Structure Model

	STABLE	DYNAMIC	DEDUCIBLE	DOCTRINAL
PERFORMANCE	• TERRAIN • ACTION ELEMENT CAPABILITIES • INTEROPERABILITY • ORBATS • PLANS	• POSITION (P), $\dot{P}$, $\ddot{P}$ • STRENGTH (S), $\dot{S}$ • GOING (COMBAT) • WEATHER • ROE • PLANS • ENGAGE/DISENGAGE/MOVE	• P, $\dot{P}$ • BALANCE OF FORCES • THREATS,OPPORTUNITIES • RATES OF CLOSURE • WEAKNESSES • ENEMY INTENTIONS • ROUTES OF PENETRATION AND EGRESS	• ADVANTAGEOUS COST EXCHANGE RATIOS • SOPs • ENEMY BEHAVIOUR • TACTICS AND STRATEGY
SURVIVABILITY OF PERFORMANCE	• AVAILABILITY OF ALTERNATES • C^2 AND AE SELF DEFENCE • C^2 AND AE DAMAGE TOLERANCE • EASE OF DETECTION	• ALTERNATE STATUS • SELF DEFENCE STATUS • EMCON STATUS	• ATTRITION • OVERRUN PREDICTIONS • ACTION ELEMENT FUTURE STRENGTHS	• MOBILITY vs ARMOUR • ATTACK/DEFEND RATIOS • STAND/WITHDRAW ETHIC • SUCOC/COLOC
AVAILABILITY OF PERFORMANCE	• TOTAL RESERVES • REPLENISHMENT/REPAIR CYCLES	• AVAILABLE RESERVES/REFORS • WARCON, POL STOCKS • TURNROUND AND REPAIR STATUS • TRANSPORT STATUS • TRANSPORT GOING • READINESS	• COMBAT DAYS REMAINING	• PREDICTED CONFLICT DURATION

Figure 13 Principal Information Categories



NOTES:
1. EACH C^2 HQ TENDS TO A 3 TIER STRUCTURE
2. SUBORDINATE C^2 HQ OVERLAP BY 2 TIERS
3. BEYOND THE THIRD SUBORDINATE THERE IS NO OVERLAP

Figure 14 Command Granularity

GETTING SEAMS STRAIGHT

D. K. Hitchins

INTRODUCTION/SYNOPSIS

Systems Engineering is lagging behind. Software Engineering debacles have fuelled a vigorous response in the form of software tools, fourth generation languages, relational databases and yet more exciting plans; major Systems Engineering fiascos have, in fact, been even more spectacular but have as yet received no corresponding attention. The paper presents SEAMS, a metasystem for developing complex systems; SEAMS is a Systems Engineering Analysis and Management Support environment.

In the context of this document a system is a bounded set of interrelated parts which together form a complex whole. In practice, most systems of interest are "Information, Decision, Action" systems and - while frequently containing processors and software - are wider in context than information handling and processing systems. A typical system might be an aircraft, a ship, a transport system, a power distribution system and so on.

Interest in systems engineering is on the increase, after a period "in the wings". As systems become ever more complex, the need to handle that complexity is being recognised as an essential science and an engineering discipline in its own right. Unfortunately, practitioners are few and the multi-disciplinary origins of potential systems engineers limit the catchment area. There is a need, then, not only to update, revitalise and re-present systems engineering to new generations of engineers, but also to provide practising systems engineers with modern tools and methods to increase their productivity. SEAMS is intended to fulfil that need.

The paper presents a Systems Engineering Code of Practice and a set of Standards as a basis upon which to build a SEAMS environment, shows the development stages which must be gone through, and presents a future vision of a fully implemented SEAMS environment in an organisation for implementing complex systems.

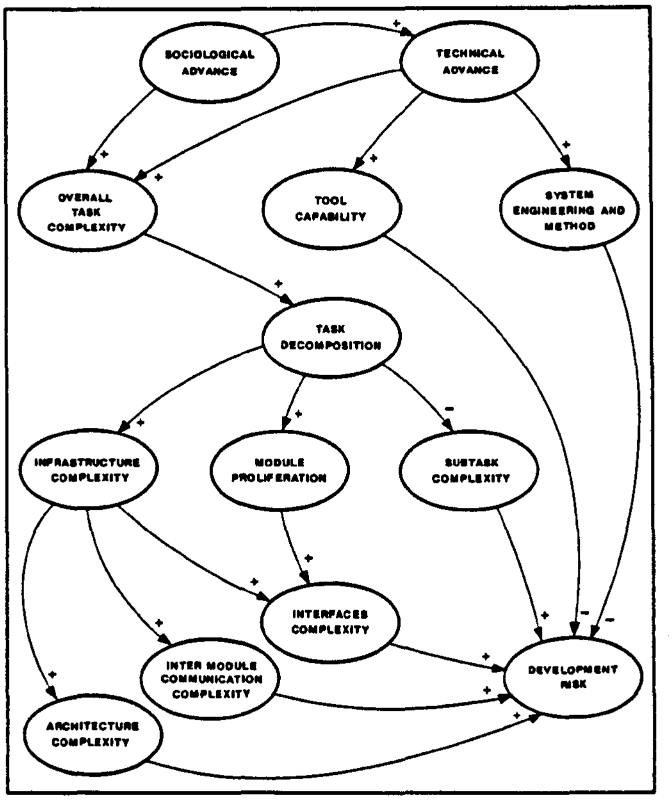
1 THE NEED

Social evolution is outstripping the ability of individuals to comprehend the more complex systems required by society. The technology which society promotes can also be applied to the handling of complexity, so giving individuals both the perspective and the power to manage the more advanced and sophisticated systems being demanded.

As society develops an increasingly complex infrastructure to support social evolution, so the demand and technology for larger, more complex and interwoven systems develops in step. Designers and engineers are not evolving noticeably as individual human beings however, and their need to breakdown (decompose) a system into manageable, man-sized parts has not changed significantly either. In consequence, the number of such parts is continually increasing and, as the figure shows, so too are the interface and intercommunication complexities.

Such a proliferation of modular interrelationships would present a mounting problem were it not for the same technological advance as that which fuels the evolution being available to accommodate the resulting complexity. Computer-supported tools, methods and techniques are either envisaged or are becoming available for requirements capture, configuration management, information handling, modelling, design and development.

Whereas significant advances have been made in the application of modern tools to software engineering, systems engineering has yet to be adequately addressed, except perhaps for the specific example of Information Systems. Wider systems such as transport, energy and military platforms (tank, ship, aircraft) have no such support. This document outlines a remedy for that deficiency.



2 SEAMS - THE METASYSTEM

SEAMS is a systems engineering environment, incorporating a total quality ethic. As a system for developing systems, it is a true metasystem and is designed to learn from mistakes, so providing an improving capability and a "corporate memory".

Systems Engineering, Analysis and Management Support (SEAMS) is a systems engineering environment for the design and development of systems. SEAMS embraces the systems approach to total quality by providing a structured, disciplined approach to the development of complete systems optimised in design for their lifetime.

If, as has been said, the essence of systems engineering is abstraction and the essence of abstraction is the model, then SEAMS should be seen in both veins:

- SEAMS is a high integrity systematic approach to producing systems. SEAMS is, therefore, a METASYSTEM, a system for developing systems.
- SEAMS can also be viewed as a model for developmental control as later topics will show.

The basic tenets of systems engineering are contained in SEAMS. Further, the concept contains the notion of a corporate memory - a near-perfect memory which learns from mistakes and introduces new methods, tools and techniques to avoid repeating old mistakes as well as accommodating new tasks.

Systems - HOLISTIC, INTEGRATED

Engineering - ORGANISED, STRUCTURED, DISCIPLINED

Analysis - INVESTIGATED, ASSESSED, EVALUATED

Management - PLANNED, RESOURCED, CONTROLLED

Support - PROJECT-SPECIFIC, COMPANY-WIDE

3 THE CONCEPT

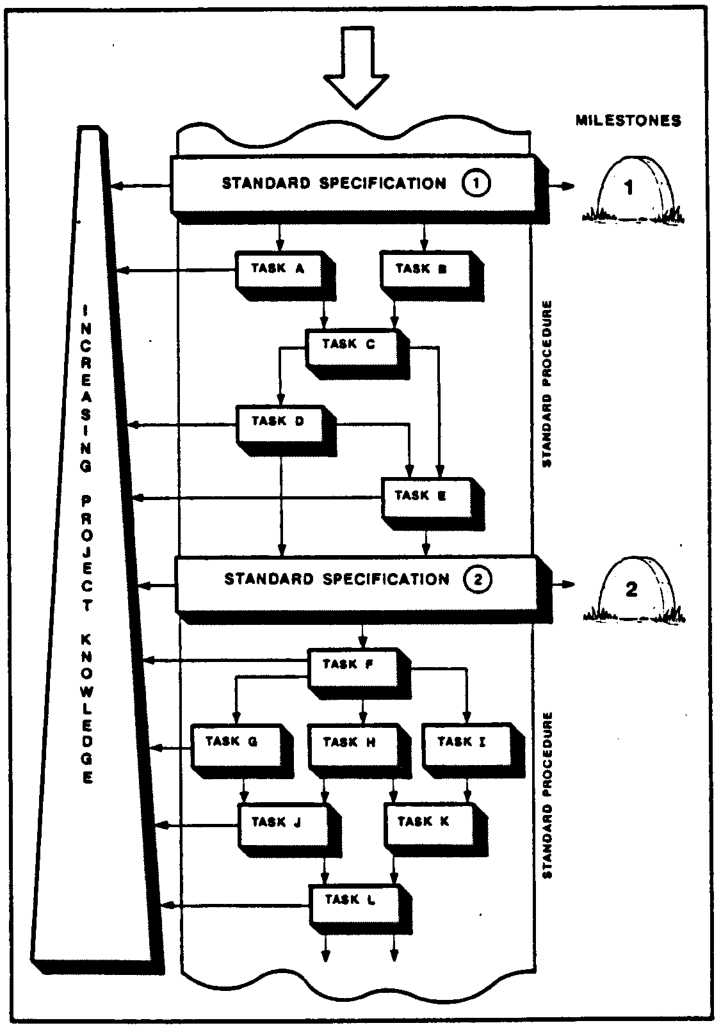
Controlled system developments comprise a phased progression of specifications produced by defined project tasks which interrelate through standard procedures. SEAMS exploits this large degree of project-to-project commonality.

Systems engineering projects are properly controlled by a sequential set of phases:

- Operations Analysis
- Requirements Analysis
- System Design
- Subsystems Procurement
- Test and Integration
- Installation and Commissioning
- In-use Phases.

Phases are punctuated by specifications. Phases comprise a set of activities designed both to produce the end-of-phase specifications and to conduct the task work leading to the implementation.

As the figure shows, a model of the controlled system development can be seen as formalised sets of procedures and specifications, with the procedures providing the relationships between tasks specific to the particular project. It is perhaps, surprising that any system project can be sensibly viewed in this light, and that a very wide variety of projects can be accommodated by a common set of procedures and specification standards. It is this commonality that SEAMS seeks to exploit.

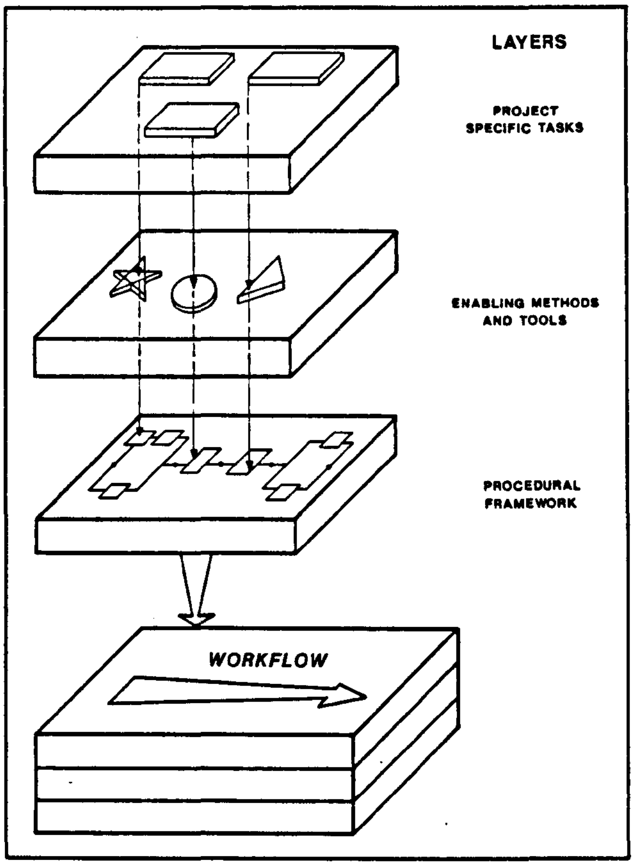


4 THE SEAMS BACKPLANE CONCEPT

Different projects demand different methods. SEAMS must be adaptable to different system needs and it does not therefore contain specific tools. Instead, tools can be selected and employed particular to task.

Systems projects can be large or small, complex or not so complex. It would be quite inappropriate to use tools suitable for large projects on small projects, or vice versa. Projects also have tasks quite specific to their requirement.

The concept emerges of a three layer construction for each phase, as shown in the figure. The lowest layer is the procedural framework which provides sequence, relationship, purpose and flow. The upper layer contains the project specific tasks which must fit into the procedural framework. The binder between the upper and lower layers is the enabling layer which brings appropriate tools, techniques and methods to bear so that tasks can be executed in a comprehensive, economic and timely manner as prescribed by the procedural framework.



SEAMS is not a set of tools. It is an environment into which appropriate tools may be "plugged" as dictated by the job-in-hand. It is analogous in some ways to the hardware backplane familiar to electronics engineers.

5 THE TOOL SET

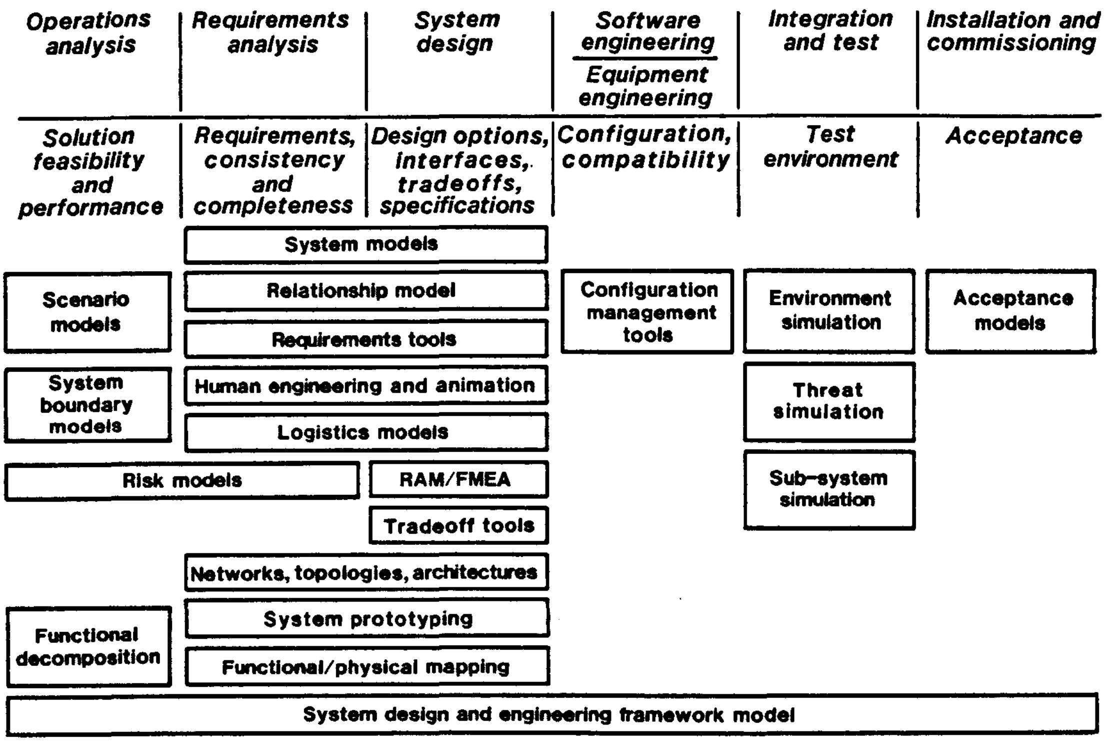
The availability of Systems Engineering models, presently patchy, is being enhanced by the application of software and CAD-based tools for requirements analysis, configuration management, space modelling, display animation, etc. Models much less in evidence presently include risk models, interface models, human engineering tools and models of complete projects in their environment.

The phases of Systems Engineering can be used as a framework within which to generate a range of functions for which tools would be valuable. The figure shows such a taxonomy for Systems Engineering models, simulations and tools. Scenario models and event simulations are well represented in the SE cycle; most major projects have been addressed using models of one kind or another. Risk models, clearly valuable in concept, are much rarer in practice. Systems models - models of the system to be developed - appear to be less in evidence than in previous years, particularly within the defence industry. One possible reason for this trend is the waning role of the Defence Research and Development (R&D) Establishments, and the emergence of the major platform and system contractor. Where R&D Establishments previously promoted the comprehensive analysis of potential system designs by modelling, major platform and system contractors now operate within a fixed price, short-timeframe environment which can little afford models which often explore future operations after the contract period has ended.

Relationship models, which examine interfaces, information exchange, protocols, capacities, configuration and architectures are clearly needed at system level, but are not much in evidence. Requirements Analysis tools abound, on the other hand; Yourdon, DeMarco and Jackson Structured Design contribute in this area, although their heritage is that of Software Engineering, and their applicability to System Engineering is less than complete. Human Engineering models range from display animation (e.g. rapid Man-machine Interface (MMI) prototyping) to physical models; further work in this area is needed to accommodate the most complex system interface, that with its operator/user.

Physical space models of installations are giving way to the more flexible but presently less tangible 2-D and 3-D CAD model. In some severe environments, customers accept systems on the basis of the performance of a system model, rather than by tests of the system itself.

One model which appears to be almost totally missing in software or indeed any other form is that of the project itself, in relation to the company and environment surrounding it. There is a real need to be able to assess the impact of change in the process of a project. Changes in external economic conditions, in availability of external resources and technology, in the customers' requirements, in the future environment can all materially affect the developing project. When two or more of these external factors coincide, the need for a model becomes even more acute. A present example in the public domain is the impact of the Canary Wharf development upon the Docklands Light Railway project which was conceived within a boundary which did not consider that development. A project model could allow such instances to be anticipated and accommodated.



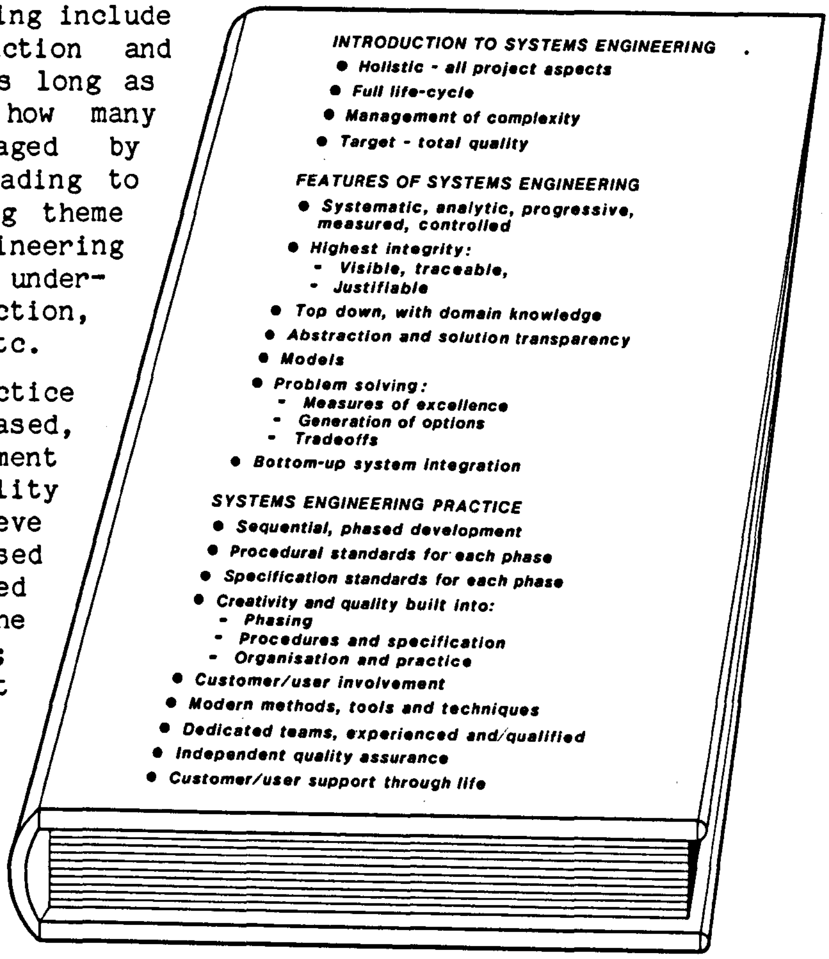
6 THE SYSTEMS ENGINEERING CODE OF PRACTICE

The proposed Systems Engineering Code of Practice contains systems engineering philosophy, principles and tenets, features which are characteristic of good systems engineering and, of course, the elements of good systems engineering practice.

There is presently no recognised code of practice for Systems Engineering, which is surprising in view of the variety of organisations practising as systems companies. A Systems Engineering Code of Practice should include principles, characteristics and practice; shown overleaf is a proposed composition in outline.

Introduction to Systems Engineering is concerned with presenting the scope and objectives of Systems Engineering (SE). As presented, SE is the all-embracing term and therefore incorporates systems design, development, implementation and their management. The target of the Systems Engineering approach has always been a holistic, top-down, life-cycle approach to total quality.

Features of Systems Engineering include its dedication to abstraction and solution transparency for as long as practicable in design - how many projects have been damaged by preconceived assumptions leading to bottom-up design? A strong theme running through Systems Engineering is modelling, (the basis for understanding), analysis, prediction, estimating, visualisation, etc.



Systems Engineering Practice concerns itself with phased, progressive, formal development and with creativity and quality going hand-in-hand to achieve controlled, risk-assessed innovation. The sophisticated does not overshadow the straight forward, however; Systems Engineering is about team work and that requires dedicated, integrated teams and (ideally) continual customer involvement to ensure a smooth development and transition to use.

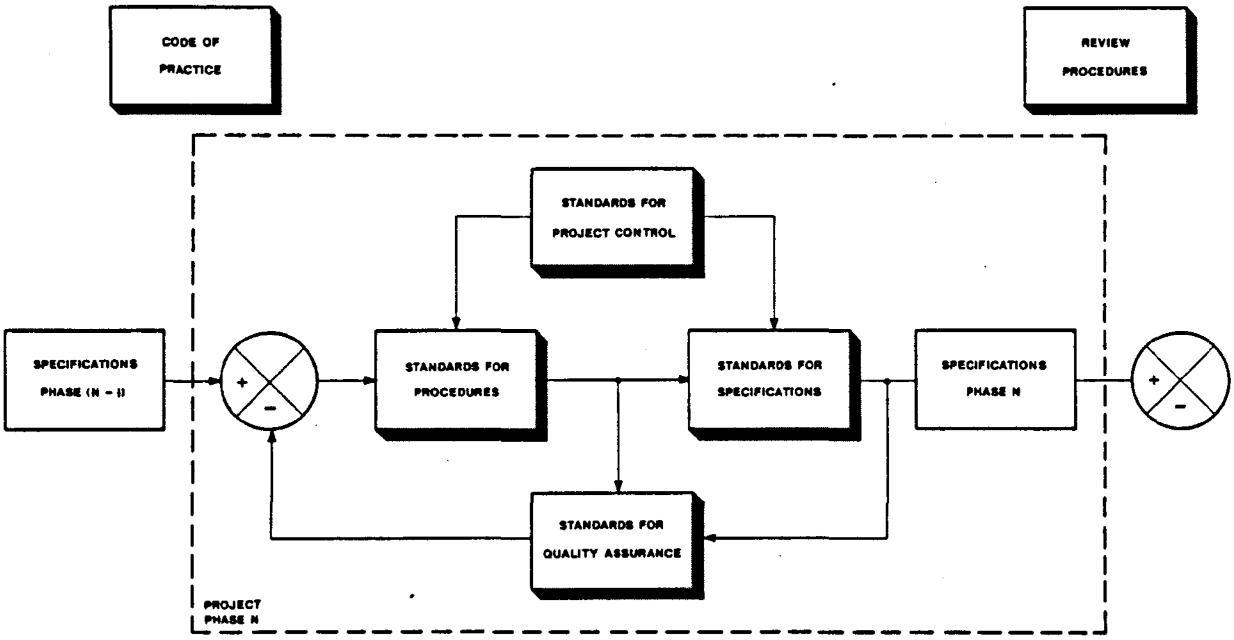
7 RELATIONSHIP BETWEEN STANDARDS

Standards are proposed for Procedures, Specifications, Project Control, Quality Assurance and for review of the standard themselves.

The figure shows the proposed sets of SE standards (in relief) and their interrelationships. A single phase, N, is shown as part of a sequence of system development phases. Phase N uses as its input the specifications from prior phases, and it in turn produces a set of specifications as its major controlling output.

Standards for Procedures show how work should proceed during each phase. Standards for Specifications show how the principal output from each phase should be specified. Standards for Project Control show how the execution of tasks and the achievement of specifications should be regulated. Standards for Quality Assurance show both procedures and specifications for project independent assurance of creativity, design, development, proving and performance. Quality Assurance standards are shown conceptually as a feedback loop to manage quality.

The whole set of standards is seen against the backcloth of the Code of Practice and, since no set of standards can be all-embracing, a procedure is included for reviewing all the standards.



8 SEAMS STANDARDS FOR PROCEDURES AND SPECIFICATIONS

SEAMS standards exist in eight varieties: Specifications and Procedures for Requirements, Development, Project Control and Quality Assurance.

SEAMS Standards are presented in a matrix of activity against type. Activities include:

- System Requirements (Capture and Analysis)
- System Design and Development
- Project Control
- Quality Assurance.

Types of standards are limited to two:

- Standards for Specifications
- Standards for Procedures.

The total system comprises up to five systems, according to customer:

- The user's operating system
- The user's exercise and training system
- The user's maintenance system
- The in-company test and integration system
- The in-company maintenance system.

The basic framework therefore comprises a three-dimensional matrix with 40 cells in it (four activities, two types of standard, five systems). Happily, the majority of standards apply equally to all of the five systems.

	* SYSTEMS REQUIREMENT	* SYSTEMS DESIGN AND DEVELOPMENT	PROJECT CONTROL	QUALITY ASSURANCE
STANDARDS FOR SPECIFICATIONS OF:	SYSTEM BOUNDARIES SYSTEM OBJECTIVES SYSTEM PERFORMANCE USER'S ORGANISATION CONNECTIVITIES CAPACITIES INFORMATION EXCHANGES OPERATIONAL ENVIRONMENTS EXERCISE AND TRAINING RAM EMC / EMP FUNCTIONAL DESCRIPTION THE 'ILITIES'	SYSTEM ARCHITECTURE SYSTEM DESIGN SYSTEM SUPPORT INTERFACES ICDs OPERATING SYSTEMS INFORMATION MANAGEMENT APPLICATIONS DATA DICTIONARY INTEGRATION AND TEST INSTALLATION AND COMMISSIONING ACCEPTANCE	LIFE-CYCLE PLANS PROCUREMENT PLANS INTEGRATION AND TEST PLANS DEMONSTRATION AND ACCEPTANCE PLANS TRANSITION-TO-USE PLANS	SYSTEM DESIGN QUALITY PLANS DEVELOPMENT SYSTEM QUALITY PLANS SYSTEM PROVING QUALITY PLANS
STANDARD PROCEDURES FOR:	REQUIREMENTS ANALYSIS SYSTEM BOUNDING FUNCTIONAL DESCRIPTION FUNCTIONAL DECOMPOSITION FUNCTIONAL-TO-PHYSICAL MAPPING	SYSTEM DESIGN SUB-SYSTEM DESIGN SUB-SYSTEM PROCUREMENT TEST AND INTEGRATION INSTALLATION AND COMMISSIONING CUSTOMER SUPPORT / PDS ESTIMATING OPTIMISING TRADEOFFS	PLANNING SYSTEM DESIGN PLANNING SYSTEM DEVELOPMENT MONITORING PROGRESS COST ESTIMATING COST CONTROL DEVELOPMENT FACILITY MAINTENANCE DATA MANAGEMENT CONFIGURATION MANAGEMENT STANDARDS REVIEWS PROCEDURES REVIEWS	DESIGN QUALITY AUDIT DEVELOPMENT SYSTEM QUALITY AUDIT DEVELOPMENT PROCEDURES QUALITY AUDIT PROJECT CONTROL QUALITY AUDIT QUALITY PLANS REVIEWS QUALITY AUDIT REVIEWS

* THE 'SYSTEM' COMPRISES 5 SUB-SYSTEMS - THE USER'S OPERATIONAL SUB-SYSTEM
- THE USER'S EXERCISE AND TRAINING SUB-SYSTEM
- THE IN-SERVICE SUPPORT SUB-SYSTEM
- THE DEVELOPMENT INTEGRATION AND TEST SUB-SYSTEM
- THE IN-COMPANY SUPPORT SUB-SYSTEM

9 REALISING SEAMS

SEAMS will start life as an evolving set of Systems Engineering Standards. While these standards are maturing, an automated version of SEAMS will be designed, developed and tested, using Manual SEAMS as a template. A final transition phase will move the organisation across to full Auto-SEAMS.

SEAMS is a concept yet to be realised, even though a number of the building blocks may be available. Such an undertaking cannot be approached lightly. The procedure for developing SEAMS is shown opposite and comprises three phases:

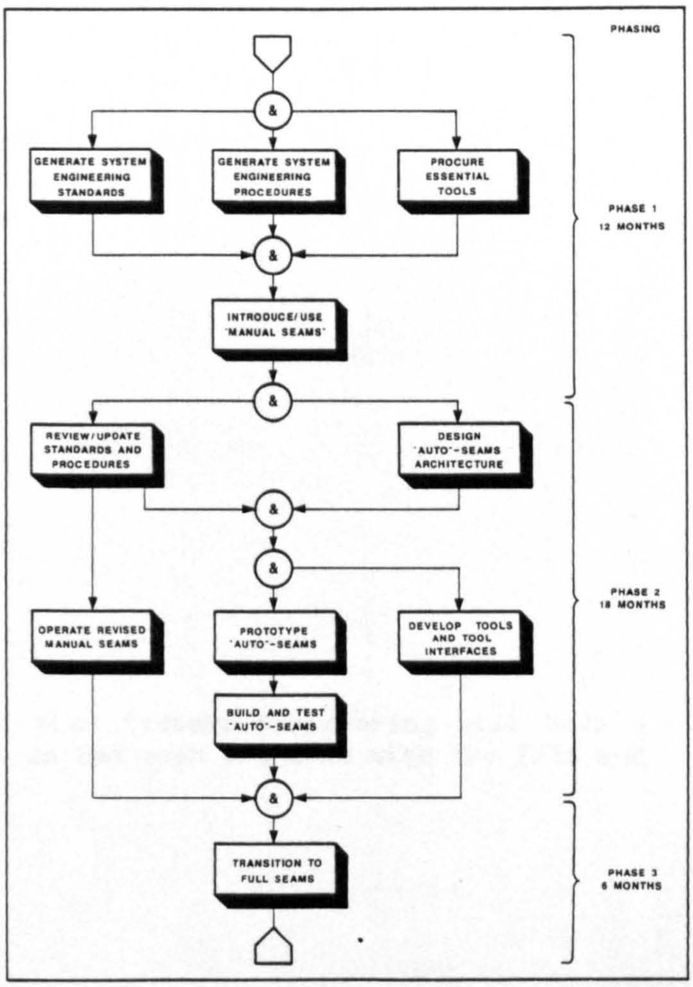
Phase 1 Design and Introduce the Standards - the so-called "Manual SEAMS"

Phase 2 Design, prototype and develop Auto-SEAMS while operating and improving Manual SEAMS

Phase 3 Transition to Auto-SEAMS.

Auto-SEAMS is, of course, a company wide, computer supported SEAMS with integrated procedures, specifications, management controls, information handling, project reporting, financial controls, etc.

The prudent three-phase introduction is necessary not only to develop a system with integrity, but also to manage the cultural change inherent in introducing a major new approach to company activities.



Integrated Project Support Environments (IPSEs) point the way for a higher order Systems Engineering Analysis and Management Support (SEAMS) environment, with an integrated set of tools, techniques, models and formats providing a comprehensive backcloth upon which to develop the future system. IPSE and SEAMS, not dissimilar in concept, will be quite different in realisation, and Software Engineering will then bootstrap Systems Engineering into a formal discipline.

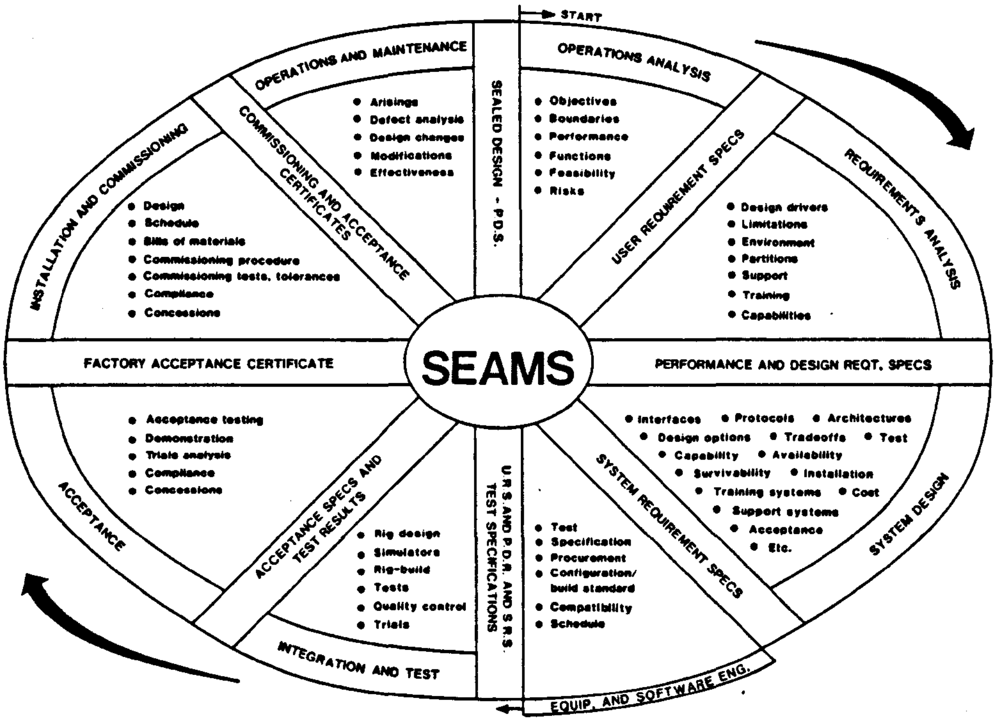
Foregoing topics have shown the need for, and the route for achieving, a sound, high-integrity approach to the engineering of complex systems. The Software Engineering IPSE has led from underneath. The need is apparent for a corresponding, but wider capability at systems level to bring together under one integrated system those various tools, techniques and methods which together will form a SEAMS environment.

SEAMS is an apposite acronym; its purpose is to provide a continuous, supporting blackcloth of tools, procedures, models and methods to work within each project phase and to blend each phase into the next via the appropriate specification using a common project database. (Current jargon favours "seamless" for such concepts, and SEAMS should provide the customer/user with a seamless, complete system balanced and optimised to his needs.) Some of the tools are presently available for each phase, although they are far from being in an integrated set.

SEAMS would comprise a project database, accessed through a series of workstation displays for each phase. There would be in SEAMS:

- Operations Analysis Tools
- Capability and Performance Models
- Requirements Analysis Methodologies
- Interface and Configuration Management
- Life-cycle Cost Models
- Survivability/Durability Models
- Cost Effectiveness Analysis Tools
- Financial Forecasting Models
- Model/Skeleton Specification
- System MRI and Drawings
- Space Models
- HF Models
- Availability Models
- Risk Models
- Architecture Models.

When SEAMS is realised, as it will be, then Systems Engineering will have a valuable focus for technique development as has been the case with the IPSE and Software Engineering.



11 THE VISION

SEAMS will become an integral part of company management, technical organisation and processing support. The cost of SEAMS will be justified by the enormous reduction in risk and corresponding increase in quality its introduction will invoke.

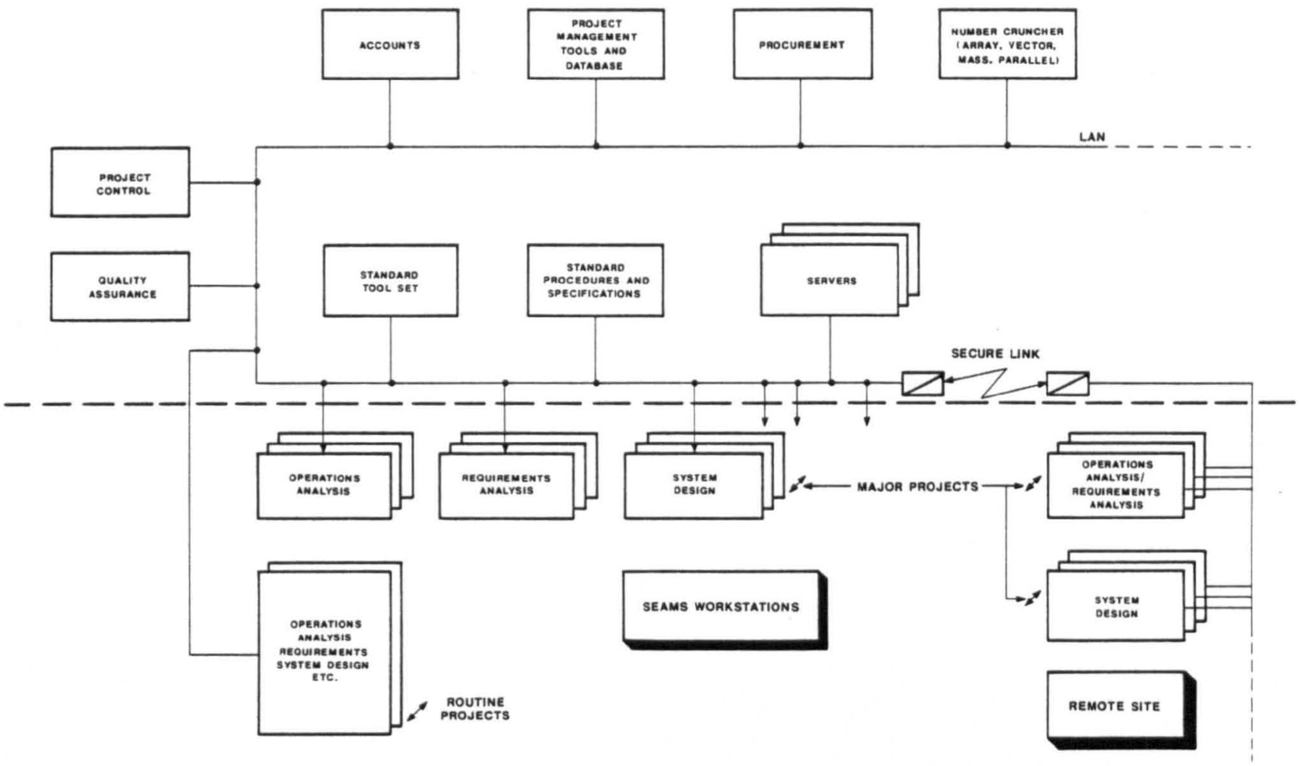
What will a future SEAMS environment look like? The figure shows the general concept; SEAMS is seen as an integral part of the company computing system. The non-SEAMS specific elements are shown above the dotted line and include project control and Quality Assurance workstations; project control operates at a level "above" SEAMS in the sense of authority for project time and budget. Quality Assurance operates properly outside the project on behalf of the company and the customer.

SEAMS itself is seen as sets of workstations, dedicated to individual project phases, and as single workstations able to operate in any or all phases according to need; both workstation concepts will be needed to accommodate large and small projects. Workstations will call down appropriate tools, standard procedures and specifications from read-only sources, all formally quality approved and under Quality Assurance jurisdiction. Current project data, procedures, specifications and "tools-in-action" are held in file servers, and are called down to the intelligent SEAMS workstations for working. Major processing tasks can be delegated to a "number-cruncher" on the Local Area Network which (funds and technology permitting) could be anything from a powerful but conventional serial processor to an exotic parallel processor.

Access to all facilities is envisaged from remote parts of the company, from individuals working as company consultants to external projects, or from home. Work from home, for example, could be executed using a suitable PC and a secure modem, security being a keynote of off-site operations in particular.

Obstacles to achieving SEAMS? Few technically, apart perhaps from a secure, distributed operating system, and that need is not too far from realisation. The principal obstacle probably arises in the business justification of the full expenditure, since it would be difficult to attach the full cost of SEAMS to a single project at project start up. For the faint of heart in this respect, consider the cost of countless recent major system projects which failed (or lost big money) quite simply due to grossly inadequate systems engineering. The cost of SEAMS pales ....

SEAMS is presently only a vision; it is realisable and potentially very valuable as a cogent power for total quality in the engineering of complex systems.



FOREWORD

Automatic Threat Assessment is at the so-called 'leading-edge' of technology. Airborne Automatic Threat Assessment cannot be said to exist. This volume is therefore concerned more with establishing a blueprint for future designs, than in assessing contemporary systems.

The volume is presented as a series of topics, which address particularly those areas believed to represent the more difficult hurdles in the path to successful design. The 1990s threat is prognosticated, so that readers may follow the design rationale from first principles through to a recommended design concept which can be seen at Figs. 5.10 and 6.4.

Finally Topic 6.2 lists the more significant, outstanding research studies, together with a series of tasks which propose a route to successful designs in the future.

CONTENTS

CONTENTS (Cont'd)

ABBREVIATIONS

AMRAAM	Advanced Medium-range Air-to-air Missile
APC	Armoured Personnel Carrier
ASRAAM	Advanced Short-range Air-to-air Missile
CRT	Cathode Ray Tube
CW	Continuous Wave
DSD	Data State Design
EO	Electro-optical
EH	Electro-magnetic
EW	Electronic Warfare
FEBA	Forward Edge of the Battle Area
FLIR	Forward-looking Infra-red
HARM	High-speed Anti-radiation Missile
HFI	Hostile Fire Indicator
HMS	Helmet Mounted Sight
IFF	Identification Friend or Foe
IR	Infra-red
JTIDS	Joint Tactical Information Distribution System
LWR	Laser Warning Receiver
MSP	Multi-sensor Processor
MTI	Moving Target Indication
NIS	NATO Identification System
POST	Passive Optical Seeker Technology
RPV	Remotely Piloted Vehicle
RTB	Return to Base
SAM	Surface-to-air Missile
TWS	Track-while-scan
UV	Ultra-violet
WP	Warsaw Pact

1 THREAT SOURCES

1.1 THREAT SCOPE

The threat comprises weapon and environmental hazards external to the aircraft, which are expected to exist in a Central European Theatre in the 1990s.

This study is concerned with threats and it is necessary to define the scop of that term. *'A threat is that which may hazard the aircraft, the pilot o. his mission, and which is external to the aircraft and its systems.'*

An engine fire, therefore, would not be a threat in this context although it may have been caused by, say, a cannon shell which was a threat. Aircraft system failures are accommodated by the aircraft's primary and auxiliary warning systems.

Threats may originate from:

- Enemy action
- Misdirected action by own forces
- Natural hazards

This study is concerned principally with the Central European Theatre in the 1990s and beyond. Thus the threat characteristics are not limited to the parameters of present-day arsenals; weapons being introduced now will be obsolescent in ten to fifteen years, and new weapons and sensors will be in use. It is necessary therefore both to forecast the environment and to recognise the limitations of forecasting by designing a reconfigurab system for threat assessment.

For convenience, threats may be grouped into:

- Groundfire - projectiles
- Surface-to-air guided missiles
- Nuclear radiation
- Biological and chemical agents
- Air-to-air guns and missiles
- Beam damage weapons
- Hazards to flying

Each of these will be addressed in more detail in succeeding topics to enable consideration of the characteristics of the threat as seen in the future scenario by extrapolation from current (especially US) technology.

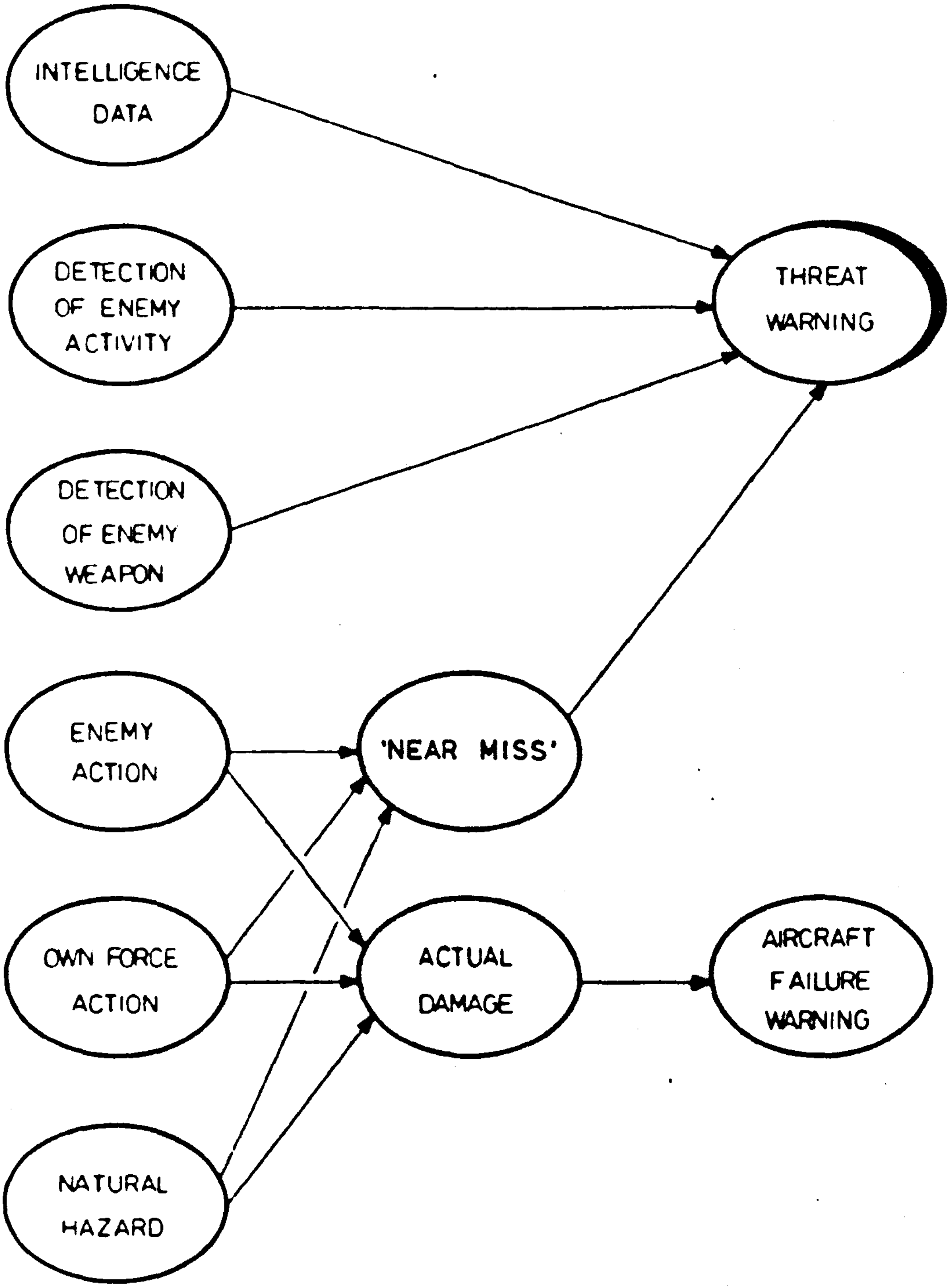


Fig. 1.1 RELATIONSHIP BETWEEN THREAT WARNING AND FAILURE WARNING

1 THREAT SOURCES

1.2 GROUNDFIRE THREAT

The ZSU-23-4 and its successors will constitute a significant point defence threat against which avoidance is more likely to afford protection than armour. Small arms fire will be a constant threat against which some protection may be practicable, while 'smart' rockets are a possibility.

The principal anti-aircraft gun at present in the Warsaw Pact (WP) Armour in the ZSU-23-4 SP, which as the title indicates is a tracked vehicle carrying four guns each of 23 mm calibre, purpose-built for point air defence. The guns fire mixed-belt ammunition, three high explosive rounds to one armour-piercing incendiary, with a muzzle velocity of 970 m/s (Mach 3.0) at a rate of 200 rounds/minute/barrel. The ZSU radar works in J-Band for target acquisition and fire control, with a range out to 20 km line-of-sight; Moving Target Indication (MTI) is fitted and there is an optical sight. ZSU-23-4 proved very effective in the Middle-East Wars, with a practical range of 2000/2500 m: four vehicles are allocated per Motor Rifle Regiment and eight per Tank Regiment so the ZSU-23-4 is likely to be encountered when attacking WP armour.

The ZSU-23-4 has limited range which could be increased up to a point by increasing the calibre to 30 mm or more. It seems likely that future point defence guns of this type will also employ thermal or millimetre wave passive imagers/trackers to obviate the need for self-revealing transmissions in clear weather.

The size and composition of a 23 mm (or larger) shell are such that a high-performance aircraft is unlikely to sustain a hit without being seriously damaged, and there seems little purpose in attempting to provide external armour against such heavy calibre shells; effective external armour would be very heavy and would reduce the aircraft weapon load. Internal armour and the use of, say, avionics, to shield vital flight critical components, may offer some measure of protection however. The ZSU-23-4 and its successors are considered to be a significant threat therefore, particularly at ranges of less than 2500 m. Figure 1.2 shows the parameters for a notional successor.

Small arms fire will be an ever present threat; 7.62 mm shells will be typical. The distribution of this hazard is likely to be both wide and unpredictable, so that warning with a view to avoiding the threat may be profitable only if the fire rate has exceeded some threshold and there is some prospect of reducing the risk by evasive- or counter-action. It may be practicable to protect some parts of the aircraft against small arms fire, so as to absorb direct hits without incurring short-term danger.

One other possible development which may be read across from the contemporary US scene is that of laser-guided projectiles or rockets. 'Smart' rockets are being examined and have considerable potential for point defence out to 6000 m or more - once a cheap reliable laser guidance head can be employed to reduce dispersion. Their advantage is their potential for saturating a volume of space or an area of ground.

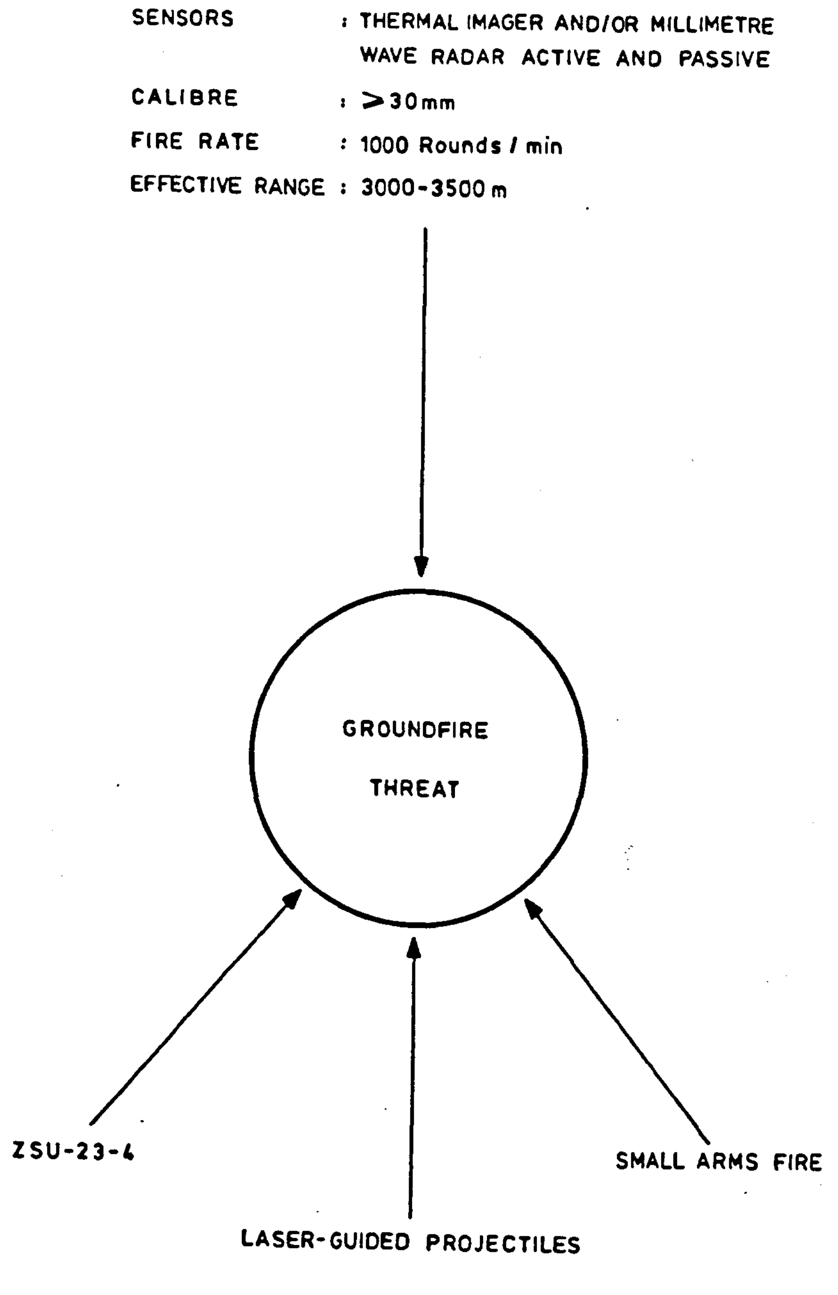


Fig.1·2 POSSIBLE WARSAW PACT ANTI-AIR GUN - 1900's

1 THREAT SOURCES

1.3 SURFACE-TO-AIR MISSILE THREAT

Battlefield Surface-to-air Missiles (SAMs) such as SA7, SA9 and their successors will probably constitute the single biggest threat to anti-armour/ground attack aircraft in the future battle scenario.

Warsaw Pact forces have a number of proven second-generation SAMs, including the SA6 GAINFUL, SA7 GRAIL, SA8 GECKO, SA9 GASKIN and the (as-yet) unknown SA-X-10. The SA7 is of particular interest in the battlefield scenario, being man-portable. This missile is a tail-pursuit infra-red (IR) tracker with an impact fuze only; the range is said to be up to 10 km, and improvements have been made to withstand the distraction of IR flares. The SA9 comprises four GRAIL mounted on a wheeled amphibious vehicle. SA6 and SA8 are longer range SAMs; SA6 for example has a minimum range of 4 km and a maximum of 60 km, and is intended primarily for area air defence.

Likely improvements to WP SAMs may be forecast by comparison with the US STINGER and Alternate STINGER programmes. The STINGER employs Passive Optical Seeker Technology (POST) which is a 'two colour' image scan using both IR and ultra-violet (UV) to improve performance against countermeasures while Alternate STINGER is a laser beamrider. The successors to SA7 and SA9 are likely also to employ more elegant guidance, control and fuzing techniques which will be aimed to nullify countermeasures and to conceal the launcher. SA7B is believed, *at present* to have a kill probability against helicopters of about 90%. Similar performance can be reasonably anticipated against low-level, fixed-wing aircraft in ten to fifteen years' time. The deployment of SA7 and SA9 is such that a future battle is likely to see more WP kills accomplished by SAMs than by guns.

Figure 1.3 indicates the operational characteristics of a notional WP missile for the 1990s, based on current US developments.

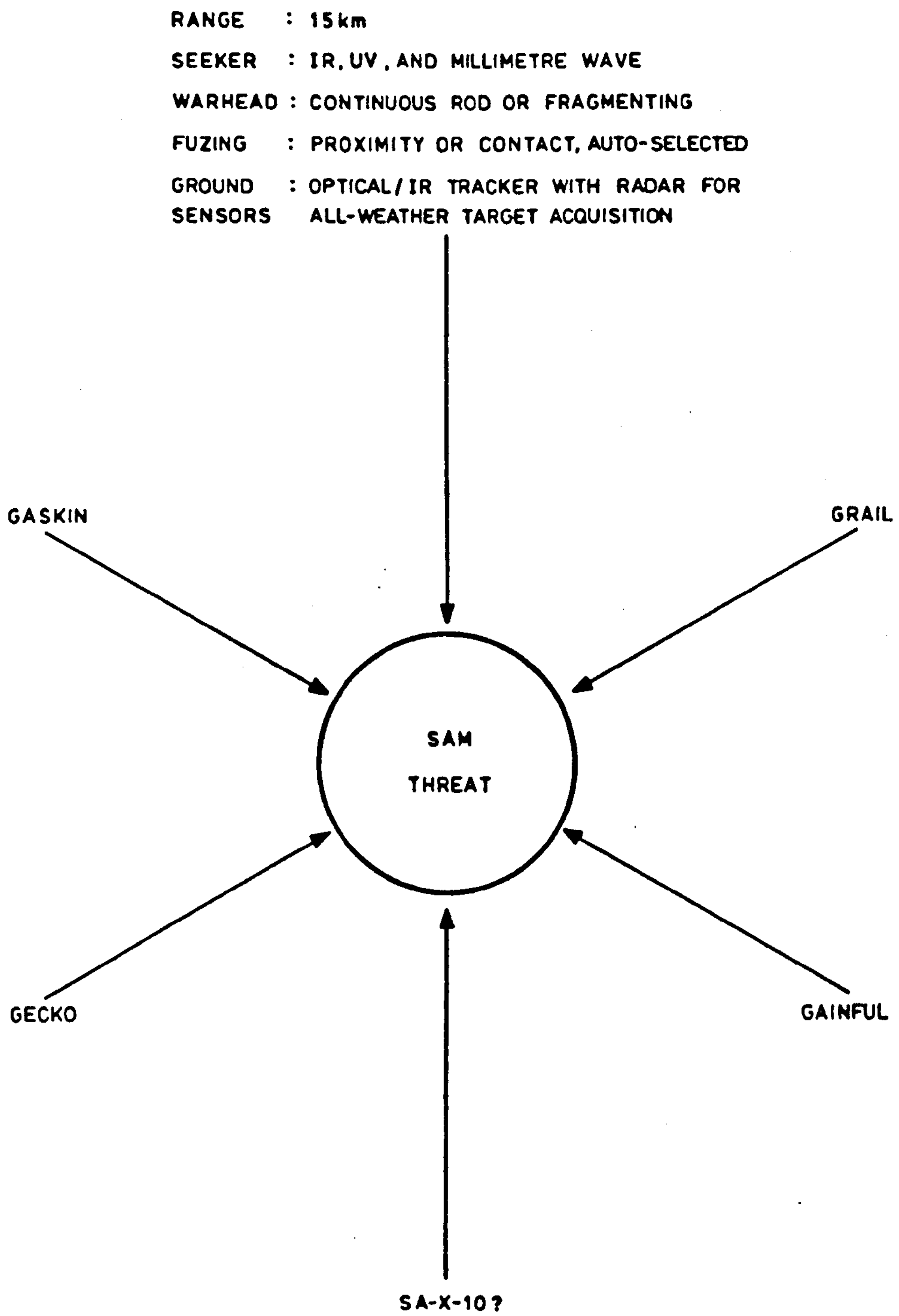


Fig. 1.3 POSSIBLE 1990's WARSAW PACT SAM

1 THREAT SOURCES

1.4 NUCLEAR, BIOLOGICAL AND CHEMICAL THREAT

Biological and chemical agents may be filtered out of supply air. Nuclear radiation may present a more significant threat.

Warsaw Pact nations are known to employ chemical weapons in large quantities and the stockpiling of biological agents is suspected. Neither of these is expected to constitute an immediate threat to an aircraft or its pilot provided that possible effects are anticipated in design by filtering cockpit air, selected use of 100% oxygen, appropriate aircrew clothing etc.

Nuclear radiation could present a more significant hazard however, since there is no means in current aircraft designs to detect radiation directly and hence avoid (or limit duration of stay in) an area of dangerously-high radiation levels. The use of tactical nuclear weapons and of enhanced radiation (neutron) weapons can be anticipated, and the pilot's efficiency could be permanently impaired within minutes of flying through dense nuclear radiation. Avionics solid state devices are also sensitive to radiation and 'hardening' incurs both weight and cost penalties; the probability of serious avionics failure is also related therefore to the total radiation dosage received, and *total* hardening against this threat cannot be realised in small aircraft. (Hardening against neutron damage appears to be impractical in any event.) Even modest degrees of protection may be expensive in both money terms and in reduction of payload/radius of action.

The detection of neutrons is more difficult than that of charged particles. Direct detection of neutrons is difficult because their lack of charge allows them to penetrate the open structure of atoms. However, occasional direct hits on atomic nuclei may eject charged particles which can be detected by their ionisation trails within gases.

e.g. $^{10}_{5}B\left(^{1}_{0}n,\ ^{4}_{2}\alpha\right)\ ^{7}_{3}Li$

Neutrons have a 932 s half life outside the atomic nucleus, after which they decay to one proton and one electron. Being charged particles, these are easily detected, but both travel distances of only a few centimetres in air.

$$^{1}_{0}n \rightarrow {}^{1}_{1}H + {}^{0}_{-1}e$$

1 THREAT SOURCES

1.5(a) AIR-TO-AIR THREAT - MISSILES

Reading across from contemporary US developments suggests that significant improvements can be forecast in Warsaw Pact missile performance.

Warsaw Pact nations employ a variety of air-to-air missiles:

- ACRID AA-6 has a range of 40 to 50 km in its semi-active, radar-guided version, and of 20 km using IR-homing. It is currently fitted to FOXBAT.
- APEX AA-7 has ranges of 35 and 15 km respectively for its radar and IR-homing variants. APEX is fitted to MIG-21 FISHBED and MIG-23S/FLOGGER-B fighters.
- APHID AA-8 has ranges of 8 km or more for its radar and IR models. It too is fitted to the MIG-23S.

Current WP philosophy is to fit suitable aircraft with either a mix of missile types or a mix of missile guidance. MIG-25 carries four AA-6 missiles - two with IR homing and two with semi-active radar heads. Alternatively, MIG-23S carries two AA-7 and two AA-8.

Developments in the US air-to-air missile scene, include advanced medium- and short-range missiles and the introduction of Anti-radiation Missiles (ARMs), which will be discussed in Topic 1.5(b). The Advanced Medium-range Air-to-air Missile (AMRAAM) will employ inertial mid-course guidance, active radar terminal guidance (possibly with a passive back-up system) and proximity fuzing, and will, of course, be fire-and-foreget, with the carrying fighter therefore able to engage several targets simultaneously. The Advanced Short-range Air-to-air Missile (ASRAAM) is less well-defined, but active millimetre wave radar is being examined too (94 GHz) and the missile is also expected to be fire-and-forget and highly agile, with a wide firing cone, good short-range performance and all target-aspect launch success performance.

The use of millimetre waves encourages the idea that the seeker head will also have a passive capability, using thermal radiation from the target in that band. This can be very effective against a clear-sky background, but would be less effective against a thermally-patterned background or through cloud, where the active radar performance would probably be more successful. Warsaw Pact air-to-air missiles in the 1990s can therefore be expected to have:

- Active and passive seekers, possibly multi-spectral
- High agility and good short-range performance

- Fire-and-forget performance
- Wide cone of fire
- All target aspect performance
- Good look-down capability and anti-clutter features
- Resistance to distractors such as flares, chaff and electronic jamming.

1 THREAT SOURCES

1.5(b) AIR-TO-AIR THREATS - ANTI-RADIATION MISSILES

Anti-radiation missiles could reduce NATO dependence on active target sensors and increase the threat of attack by passive seeker-head missiles.

The introduction of ARMs could alter the tactical choice of sensors by denying or limiting the use of active sensors to either or both sides. The US High-speed Anti-radiation Missile (HARM) AGM-88A is an example; based on SPARROW it has an 'open-front' crystal receiver designed to seek radiation from ground radars over a wide frequency range. AGM 76 STANDARD ARM and AGM 45A STRIKE are also air-to-ground ARMs. Other US developments include the use of long endurance drones or Remotely Piloted Vehicles (RPVs) which may patrol an area for five hours or more, seeking specific radars or radiation in general. There have been indications of surface-to-air and air-to-air ARM developments (e.g. BRAZO - or PAVEARM as it is known by the USAF) and interest has been noted in the radiation emitted by aircraft power generation and digital computer systems, as well as in their communications, local oscillators, and radar emissions. The widespread use by NATO countries of JTIDS/ Link 16 could reasonably be expected to increase WP interest in JTIDS/ periodic transmissions as a mid-course guidance signal for an air-to-air missile for example. Similarly, any on-board identification transponder which can be induced to transmit may serve as a suitable source of radiation, even although its emissions may be quite unintelligible.

Nations of the WP have never been slow to exploit such obvious techniques as ARMs, and it is reasonable to anticipate, therefore, that ARMs will be introduced into the WP arsenal of air-to-air missiles (and possibly SAM too), if indeed they are not already covertly present.

1 THREAT SOURCES

1.5(c) AIR-TO-AIR THREATS - GUNS

Warsaw Pact aircraft will continue to carry large calibre cannons in the 1990s because of their inherent flexibility.

Most WP fighters carry a gun, usually 23 mm calibre. The twin-barrel GSh-23 cannon is fitted to FISHBED-N and FLOGGER-B, for example, while FLOGGER-D has a six-barrel rotary 23 mm cannon. HIND-D, has a chin-mounted four-barrel GATLING-type gun, probably 12.7 mm calibre, and can carry a pylon-mounted 23 mm cannon, and four rocket pods, each 32 x 55 mm, which could combine to present a formidable barrage in the face of a low-flying aircraft, even although such aircraft might not be the HIND-D's principal target. The current air-to-air gun threat is therefore not insignificant, but generally, it is probably a relatively short-range problem only. Future WP aircraft are still likely to carry guns as these have the advantage of flexibility. For instance, they can fill the short-range performance gap left by most air-to-air missiles, can be employed against soft skinned vehicles and personnel and - vital in many circumstances - can warn without actual damage, unlike missiles.

1 THREAT SOURCES

1.6 BEAM DAMAGE OR DIRECTED ENERGY WEAPONS

Laser damage weapons may present an essentially short-range ground-based threat which is unlikely to be met on a widespread scale judging by the present state of development in the US. Soviet research and development could be more advanced than US developments however.

Current developments in the US and USSR are apparently concentrating on laser damage weapons for military applications. Charged particle beams such as proton beams defocus owing to mutual repulsion between the hydrogen ions, so their range is short (Soviet scientists have put charged particle accelerators into space for experiments, however, apparently with some success). Neutrons and gamma rays are difficult to collimate into tight beams and their delivery energies are presently below those needed to damage aircraft. Laser damage weapon research has concentrated on two spectral lines; 3.8 μm, for the pulsed deuterium-fluoride chemical laser wavelength, and 10.6 μm for the CO_2/helium/nitrogen continuous wave (CW) gas laser.

A US deuterium-fluoride laser scored a notable success in 1978 by destroying a small, supersonic, anti-tank missile. (The speed of the beam obviated the need for any lead angle computations.) Further successes have been recorded since. Soviet scientists are known to be interested in both wavelengths, and have patented techniques for boring holes through fog using CO_2 lasers. Current opinions suggest that laser damage weapons could damage structure, avionics or warheads at ranges up to 5.5 km, but only at the expense of carrying bulky equipment required to generate the power needed to support the low power efficiency of lasers.

Such weapons would therefore most probably be ground-based, although a US KC135 transport is currently being fitted with an airborne laser beam weapon. Perhaps a more widespread threat could arise from the action of lasers on thermal imager and direct view optics systems, including the human eye. Lasers in the visual and near-visual spectrum 0.4 to 1.0 μm, with modest powers, could damage a pilot's eyes, especially if he employed direct view optics; current laser designators operate in this band. Similarly, the increasing use of 10.6 μm CO_2 lasers for ranging, coupled with 8 to 13 μm band thermal imagers, suggests that the detector elements in the imagers could be attacked with some prospect of success, unless specific filters were interposed. The development of fast-reaction photochromic glass for visors and other optical devices would act both as a threat counter and as a threat warning, but could be used by the enemy for 'spoofing'. The 1990's threat from beam damage weapons is therefore seen as likely to be sparse, and based on laser beams at one of several predictable wavelengths, with the weapon's radius of danger limited to some 5 to 6 km. However, speculation is rife that Soviet beam weapon research and development is far more advanced than US work. The use of lasers as airborne weapons is possible, not least because of their psychological impact on opposing crews, and in this role they are seen as air-to-air anti-personnel weapons, as anti-missile weapons, and possibly as an alternative to short-range gun and missile fits.

Note: A comprehensive unclassified assessment of directed energy and particle beam weapons can be seen in the July 28th issue of Aviation Week and Space Technology.

1 THREAT SOURCES

1.7 FLYING HAZARDS

Natural hazards threaten mission success in a similar way to enemy action, and so form a logical part of the threat spectrum.

The anti-armour/ground-attack aircraft of the 1990s will require the ability to fly very low and to use terrain shielding in order to reduce risk of attack. There will be a need to detect and avoid terrain and other hazards such as masts, cables, bridges, etc., in clear weather, by day and night, and probably in poor weather also. The pilot will be able to avoid terrain by day, in clear weather, unaided, but he may not be able to detect cables strung across a valley for example, until it is too late to avoid them. By day, therefore, threats arise from natural hazards according to the speed and altitude of the aircraft, the degree of concentration of the pilot on terrain and hazard avoidance, visibility and so on. The pilot's problems are compounded at night and by poor weather.

These natural hazards constitute threats to the aircraft just as much as enemy action, and are therefore considered as part of the threat assessment task.

1 THREAT SOURCES

1.8 SUMMARY

Some passive sensors will be able to detect a variety of threats. Other threats will necessitate special or active sensors, use of the latter putting the Ground Attack aircraft at risk.

The threats to a ground attack fighter in the 1990s battlefield will tend to share various characteristics, as the table shows. There will be shared themes of: radar emissions for target acquisition and active guidance/ homing, especially in bad weather; IR flash emissions during the firing of large calibre guns and missile motor burns; detectable air over-pressures caused by projectile near-misses (or hits) at the aircraft and detectably high rates of change of local magnetic field as ferrous warheads pass nearby; ionised gas trials left by high temperatures of missile motors; and, generally, detectable levels of thermal radiation, where this term covers the band from 1 μm to about 4 mm.

Notable exceptions to the themes of recurrence are means of detecting residual nuclear radiation, biological and chemical agents, groundfire and cables. Similarly, the ability to detect some threats in clear weather will be poor where the enemy uses passive sensors and camouflage.

Not only may an attacking aircraft need to employ active sensors, therefore, but it may need to activate them in clear weather in order to highlight non-radiating targets/threats. This puts the 'hunter' at risk.

	SIGNIFICANT EMITTED CHARACTERISTICS							
WEAPON TYPE	ELECTRO MAGNETIC RADIATION (INTENTIONAL)	I R FLASH	THERMAL RADIATION	DETECTABLE BY EYE	IONISED GAS TRIAL	LARGE FERROUS CONTENT	AIR OVER PRESSURE	COMMENTS
ANTI-AIR GUNS	POOR WEATHER	√	√	CLEAR WEATHER		√	√	TARGET ACQUISITION RADAR USED IN BAD WEATHER OR AT LONG RANGE
GROUNDFIRE (SMALL ARMS)		POSSIBLE		POSSIBLE			?	DIFFICULT TO DETECT UNLESS HIT
SAM	POOR WEATHER	√	√	CLEAR WEATHER	DURING MOTOR BURN	√	√	ACQUISITION RADAR WILL RADIATE IF USED
NUCLEAR RADIATION	SHORT LIVED	SHORT-LIVED	POSSIBLE		POSSIBLE		SHORT-LIVED	NUCLEAR RADIATION IS DETECTABLE USUALLY BY ITS IONISATION EFFECTS
B & C AGENTS								
AIR-TO-AIR MISSILES	POOR WEATHER	√	√	CLEAR WEATHER	DURING MOTOR BURN	MAY NOT BE FERROUS	√	INTERCEPTOR RADAR DETECTABLE IN BAD WEATHER
AIR-TO-AIR GUNS	√	SMALL	√			√	√	USUALLY ASSOCIATED WITH RADAR
ARM		√	√	CLEAR WEATHER	DURING MOTOR BURN		√	
LASERS	√		√	UNLIKELY	POSSIBLE		POSSIBLE	
TERRAIN			√	√				
CABLES ETC	SOME	-	√	SOME				

B & C BIOLOGICAL AND CHEMICAL

Fig. 1.4 THREAT TABLE

2 THREAT DATA SOURCES

2.1 RADAR WARNING

Radar warning receivers may find less data content to analyse in signals received from the next generation Track-While-Scan (TWS) radars, while jammer emission analysis may expand.

Radar warning is a valuable source of data especially when the sensed-signals can be analysed to give indications of the enemy's mode of operation. For example, it would be valuable to know when a radar was in search mode, and when it changed to lock-on and fire control modes. Receivers for analysing current Warsaw Pact radars are available and provide useful, timely information. ALR 56 is a contemporary US electronic warfare (EW) management system which provides signal analysis, direction finding, establishment of threat priorities, jammer blanking, look-through capability, etc., in one centralised control package.

The advent of TWS radar will reduce the quality of information available from radar warning receivers. A TWS radar need not lock-on to a target in order to provide its associated fire control system with data. The aircraft radar warning receiver may not, therefore, observe any change of mode or scan-rate when being tracked by a TWS radar. This will significantly reduce the timeliness of the information available from that source alone, although it would still be of value to be aware of the radar scan and its type, direction range, and likely source, in order to facilitate correlation with other sensor data. Parallel developments in radar technology include the use of higher frequency, lower power, coherent, frequency-agile transmitters and more sensitive receivers, with narrower aerial beamwidths and single scan transmissions. These factors will combine to make the task facing the radar warning receiver designers increasingly more difficult.

Radar warning receivers can also analyse jamming signals designed to disrupt the aircraft radar, radio/telephony (R/T), data links, navigation aids, radio altimeter and transponders. It may prove possible to locate the source of jamming and to assess its objectives automatically to a much greater extent than at present.

2 THREAT DATA SOURCES

2.2 LASER WARNING

Laser warning is at an early stage of development and progress does not appear to be fast, owing to serious difficulties.

Laser warning systems are currently being developed but with some difficulty. Contemporary CO_2 lasers have typical parameters as follows:

Beamwidth	:	0.1 to 0.2 mrad
Mean Power	:	500 mJ
Pulse Repetition Frequency	:	200 pps
Pulse width	:	20 ns
Efficiency	:	8 to 9%
Sidelobe Energy	:	10^{-6} x Main Beam.

Thus energy per pulse is in excess of 1 MJ, and the main beam would produce a circle of diameter 50 cm on an aircraft target at 5 km.

To be sure of observing the laser energy, an aircraft would require either a large number of sensors spaced over the aircraft surfaces to pick up main beam energy or a detector to sense the (lower) sidelobe energy. A sidelobe energy detector would need a wide dynamic range in order to withstand full beam energy. At the same time, its low threshold would make it subject to nuisance trips, especially at low altitude where laser energy could be reflected off terrain.

The design of a suitable laser warning system is clearly scenario dependent as well as aircraft dependent, and at present an entirely satisfactory system is not known to exist.

Detector head designs are known to include the etalon, which is a type of Fabry-Pérot interferometer, giving sharp fringes and high resolving power, but little data is available on support processing in extracting useful and reliable information from receiver laser signals. This paucity of data could well reflect a low success rate.

Perkin-Elmer are under contract to the US Army however for the development of the AN/AVR-2 Laser Warning Receiver (LWR) for helicopters, which is integrated with the AN/APR-39 Radar Warning System. The LWR uses an etalon detector system covering visible, near and far infra-red. The system is required to detect, identify and categorise:

- Ranger - Coherent
- Designator - Coherent
- Beam rider - Coherent
- Radar - Coherent
- Weapon - Coherent
- Searchlight - Non-coherent
- Flare - Non-coherent

The detector should discriminate between coherent and non-coherent radiation to avoid false triggers due to natural or man-made non-laser illumination. Characterisation factors include:

- Wavelength
- Intensity
- Pulse rate
- Pulse width
- Position.

Four sensors are used per quadrant to locate the emitter.

2 THREAT DATA SOURCES

2.3 PRIMARY RADAR

Many developments may be expected in radar technology, which should refine the information content available from radar.

Primary radar may be required for target search, recognition and acquisition. Radar developments indicate that future systems will share certain characteristics:

- Higher frequency, moving up to the millimetre wavebands and beyond - to IR scanning laser systems
- Narrow beamwidths
- Frequency agility
- Lower transmitted power
- More sensitive receivers (high quantum efficiency for optical wavelengths)
- Coherent radiation
- Doppler beam sharpening
- Moving target indication
- Smaller mass and volume
- Extensive signal processing to give TWS, automatic target recognition, synthetic aperture performance, etc.
- Passive modes of operation to reduce aircraft total electro-magnetic (EM) signature.

The increase in frequency, leading to finer beams, coupled with high resolution in range, will provide small equivalent radiated energy packets, so that target characteristics may become apparent to the eye, and the progressive introduction of computing power should allow systems to emulate, and in some cases exceed, human eye pattern recognition performance.

Thus future airborne radar systems should be able to present far less raw data to the other aircraft systems and to the pilot; instead, more useful information should be presented.

A move to 10.6 μm CO_2 scanning laser radar seems a logical progression. Relatively high lasing efficiencies and powers may be achieved, and weather penetration is good up to tens of kilometres. 'Aerials' will be small, and a suitably designed receiver might double as a thermal imager.

2 THREAT DATA SOURCES

2.4(a) ELECTRO-OPTICAL SENSORS

Electro-optical (EO) sensors fall into two groups according to their wavelength passbands; visible and IR detectors employ photon detectors, while millimetre wave imagers will use more conventional thermal detectors and radar aerials. The imaging range of this latter waveband would be severely resolution-limited by the aperture size available in an aircraft.

The need to reduce aircraft emissions will increase dependence on EO sensors. These could sensibly operate in one of four or five passbands as determined by atmospheric transmittance (see Fig. 2.1(a)).

- 0.4 to 1 μm - visible and near IR
- 3 to 5 μm - medium IR
- 8 to 14 μm - medium IR
- Millimetre wave passbands, especially 1 and 3 mm.

Receivers operating in the last-mentioned passband would operate in a significantly different manner from the first three. Visible and IR radiation detectors are of the quantum or photo-electric type, and can use relatively small apertures to 'capture' sufficient photons for successful operation. However, millimetre wave detectors must employ larger EM field detectors (aerials) since the energy in those passbands is of the random thermal radiation type. Figure 2.1(b) shows the photon ($h\nu$) energy increasing with frequency, and the random thermal emission (kT) which is invariant with frequency but proportional to absolute temperature. Imaging in the optical and IR passband can employ photographic cameras, TV, optical scanners or (active) laser radar. Millimetre wave imagers would probably employ a scanning radiometer technique (passive) or more conventional radar (active).

Other differences between the passbands include their theoretical information content (see Section 5.1) and the contrast in mean levels of received energy from targets and typical backgrounds. Less is known in this last respect about millimetre waves than about the visible and IR bands. Additionally, the use of conventional image presentation of targets recognisable to the human eye would require very large apertures in the millimetre passband. To recognise a tank image at 3 to 4 km might require a circular aperture of more than 25 m diameter. Alternatively, a more sensible diameter for an aircraft of, say, 0.5 m, would limit target recognition range to less than 100 m. Evidently millimetre wave systems will not compete with thermal/visible spectrum imagers in presenting images to a pilot for visual target recognition, but their use for radar and passive detection of thermal radiation seems likely to expand rapidly. The ability to recognise targets by other characteristics of their reflected radiation pattern is being developed at present and will be available in new radars.

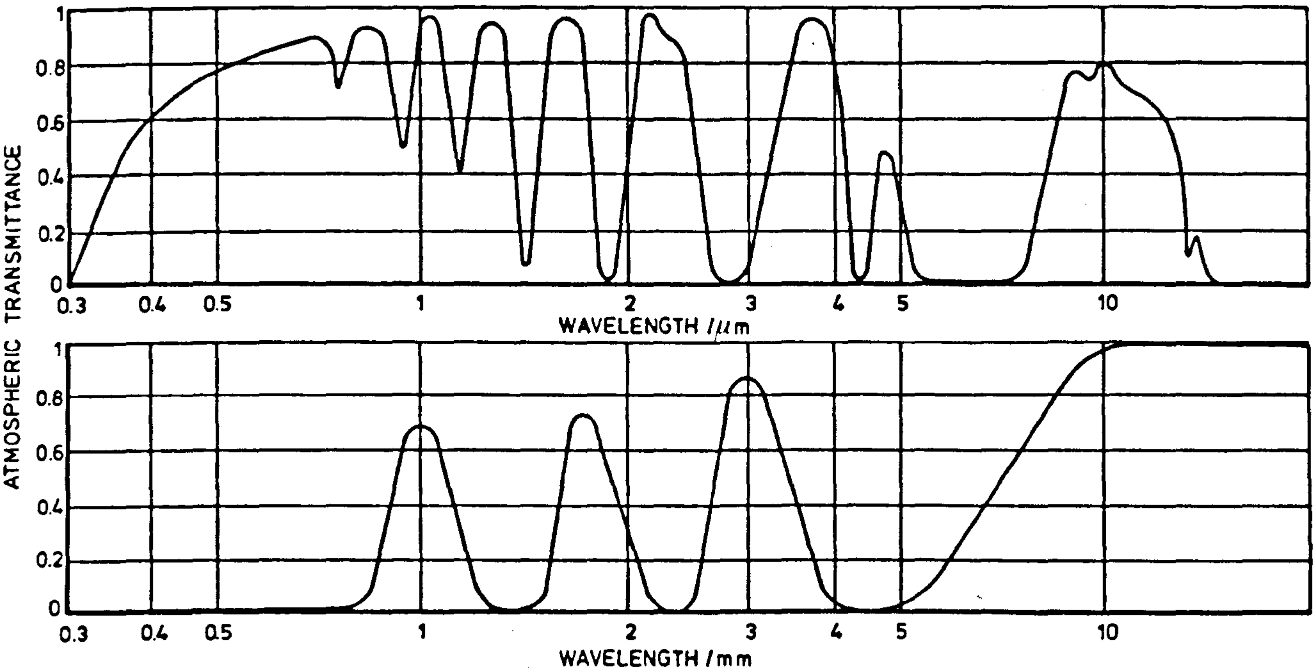


Fig. 2.1(a) SPECTRAL TRANSMITTANCE OF THE ATMOSPHERE (TREND)

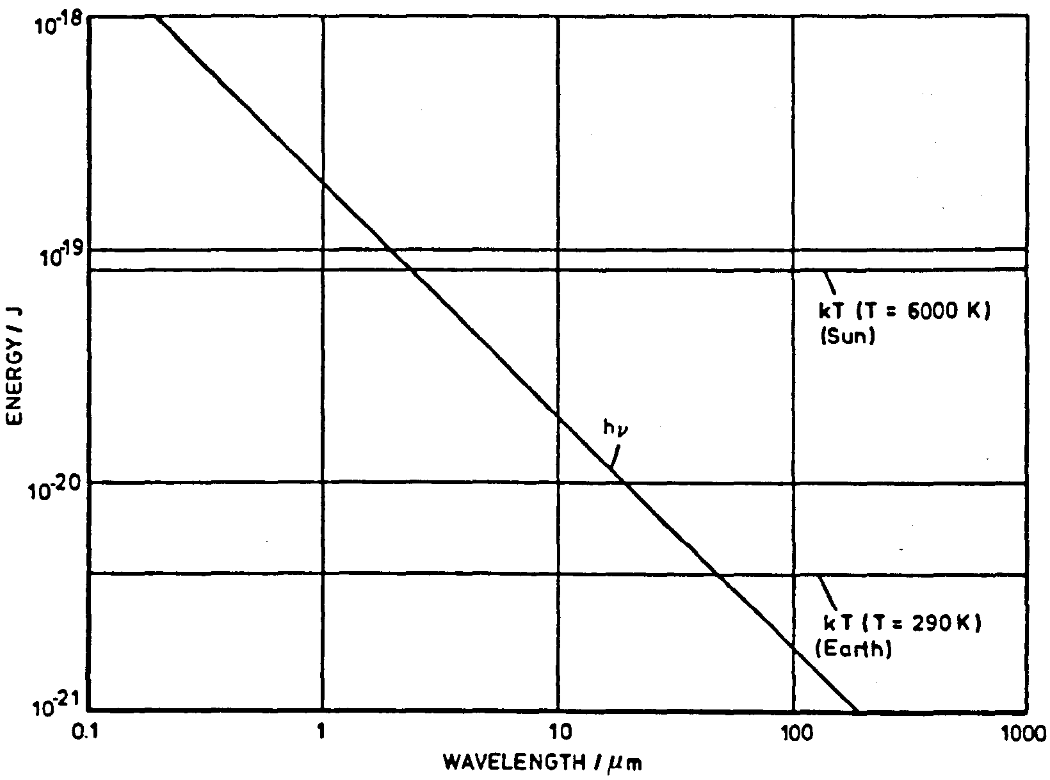


Fig. 2.1(b) PHOTON AND RANDOM THERMAL RADIATION ENERGIES OF A BLACK BODY EMITTER

2 THREAT DATA SOURCES

2.4(b) ELECTRO-OPTICAL IMAGING SENSOR CHARACTERISTICS

The three principal imaging wavebands have such different characteristics that none is likely to be self-sufficient. The combination of visual spectrum with 8 to 13 μm Forward-looking Infra-red (FLIR) imaging seems the most likely future solution.

The various passive imaging sensors have differing operational characteristics:

a) Optical

Optical sensors (TV) in the visible spectrum will work in light levels down to starlight, but not in total darkness. Images may be fully coloured, even in starlight, so offering to the pilot an image most compatible with his eyes. The characteristic radiation from hot targets is not discernible at visible wavelengths. Apertures for low-light TV may be quite large (>10 cm) to achieve sufficient photon capture.

b) 3 to 5 μm

Black bodies at 750 K emit their maximum radiation intensity in the 3 to 5 μm band, which is therefore particularly valuable for detecting hot targets. Atmospheric absorption in the 3 to 5 μm band is due mainly to CO_2, not water vapour, and is therefore the most suitable band for imaging hot targets at longer ranges (>10 km) in humid conditions. Detectors can be tuned for terrain imaging, and apertures may be quite small. A hot tank at 3 to 4 km could be resolved by an aperture of some 5 to 6 cm, although cold targets might require a larger aperture to increase photon capture area. Imagers will work in total darkness.

c) 8 to 13 μm

Black bodies at about 273 K (0 ^{o}C) emit their peak energy in the 8 to 13 μm band (see Fig. 2.2) which is ideally suited therefore to terrain imaging. This band also detects significant energy levels emitted by hot objects, which can be consequently highlighted. Imagers will work in total darkness. Atmospheric absorption is increased by humidity so that only shorter ranges are practicable in high humidity conditions. However, this band offers improved imaging compared with the 3 to 5 μm band through some (battlefield) haze and smoke conditions. Resolution may be aperture limited: to recognise the image of a tank at 3 to 4 km would require an aperture of 20 cm or more.

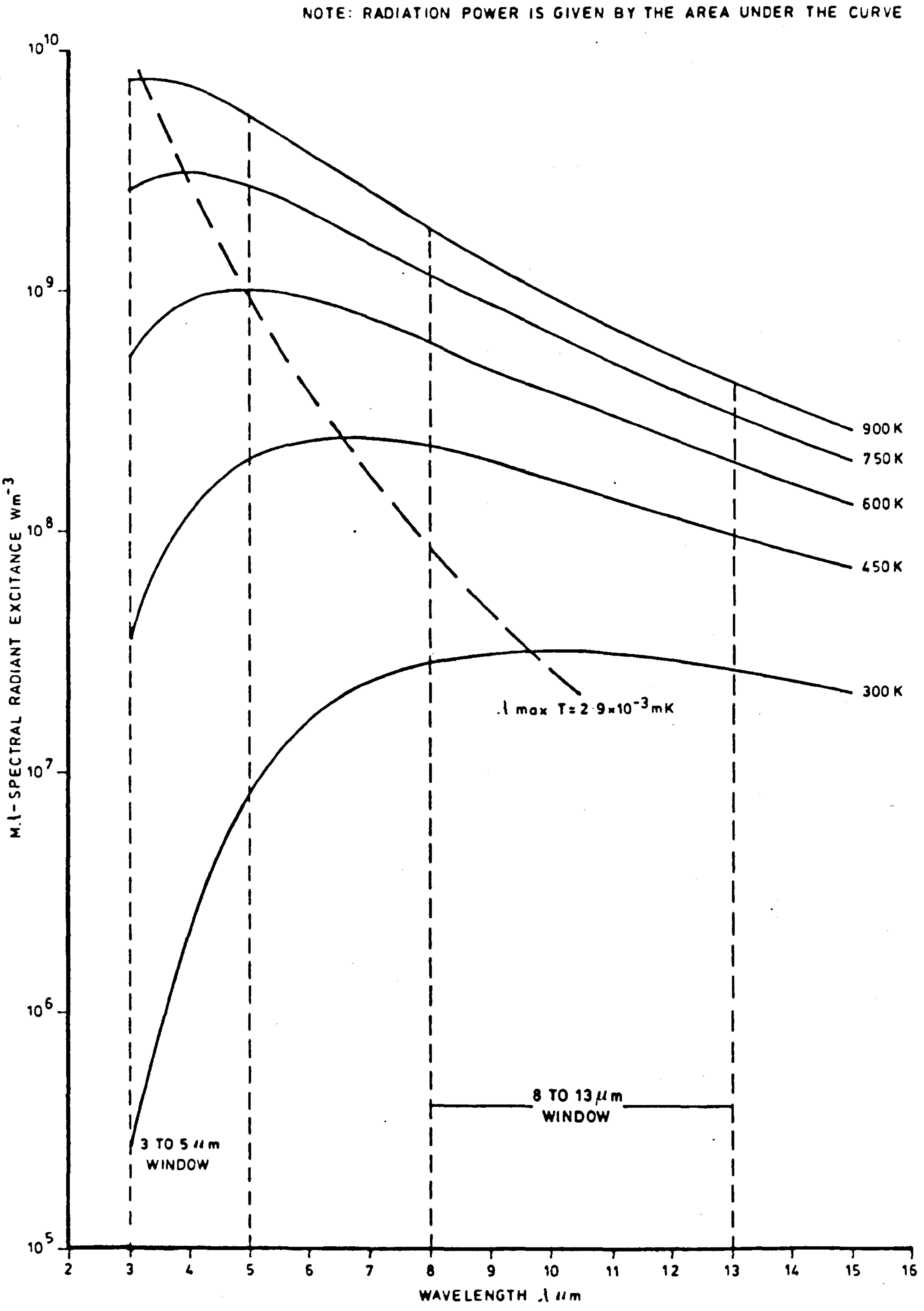


Fig. 2.2 BLACK BODY RADIATION CURVES FOR 300 K (27 °C) TO 900 K (627 °C) IN THE 2-15 μm WAVEBAND

2 THREAT DATA SOURCES

2.5 MISSILE WARNING

The most reliable means of missile warning will be active doppler radar, which, in future, may use scanning IR laser radar as a basis.

Missile warning can be provided in several ways:

- By detecting changing radar emission patterns (see Topic 2.1)
- By detecting the IR or visible emissions caused by rocket motor burn, and by atmospheric heating of the missile
- By using an active doppler radar to detect the approach of a missile directly.

Future radars may not change their emission patterns, and a future battlefield scenario in Central Europe generally suggests widespread use of missiles so that 'burn' emissions will be widespread; thus selecting the appropriate emission may not prove practicable. Therefore, the most reliable means of detecting an approaching missile is likely to be active doppler radar, which unfortunately involves self-revealing transmissions. The US has several such systems, including the AN/ALQ 156 and the F-15 EAGLE'S ALQ 153/4 tail warning, both of which are pulse doppler radars designed to respond to missile approach by initiating countermeasures. (There will be little purpose in providing warning against future missiles unless some effective countermeasures can be taken.)

Developments seem likely to include a move to higher frequencies, possibly employing scanning CO_2 laser heterodyning radar which offers good doppler performance, narrow beams and high quantum efficiency receivers, so enabling lower power transmissions to be both effective and less detectable.

2 THREAT DATA SOURCES

2.6 HOSTILE FIRE INDICATORS

The merits of a Hostile Fire Indicator (HFI) in a high-speed aircraft are questionable since countermeasures are at present doubtful.

Ballistic warning systems or HFIs usually sense pulses of atmospheric pressure associated with the nearby passage of a projectile; their use in the air has generally been restricted to helicopters, and the installation of reliable systems in these vehicles has not proved entirely successful. 'Nuisance-trips' have resulted in thresholds being raised to such levels that relatively near misses might not be detected.

The future of HFIs in high-speed aircraft would seem likely to depend more on the use which might be made of the warning than on development problems, however. The only countermeasure likely to be effective in the face of a large number of projectiles will be to evade them; this requires knowledge of the source of projectiles and of 'safe' directions for evasion. A future scenario is unlikely to be so neatly ordered that projectiles will come from only one direction and that safe routes inside enemy territory will be known. Moreover, the hostile fire is quite likely to emanate from point defence of the target(s) to be attacked (see Topic 1.2).

A CO_2 doppler radar for missile warning (see Topic 2.5) might be able to provide projectile detection too, for shells of 23 mm calibre and over.

2 THREAT DATA SOURCES

2.7 NUCLEAR, BIOLOGICAL AND CHEMICAL WARNING

There seems to be a case for the development of nuclear, biological and chemical detection systems for aircraft.

Techniques for detecting some organic and hydrocarbon gases are already established and warning devices can be purchased for use in the home. However, the range of detectable chemical and biological toxins is not extensive, and little work has been publicised on militarised airborne detectors. Such detectors would provide a valuable source of battle area intelligence. Nuclear radiation detection techniques are well established but not in a form suitable for battlefield aircraft use, where simple Geiger Muller tubes and dosimeters will be of limited value. Instead, a radiation gradiometer would seem to be essential which will indicate the direction(s) of the radiation source(s), record exposure, compute danger levels, and provide warning, intelligence, and recording facilities.

2 THREAT DATA SOURCES

2.8 CABLE WARNING

A suitable choice of primary radar would permit active detection of non-radiating cables. The radiation from some power cables might permit useful passive detection.

Power cables may carry sufficient current to produce a detectable magnetic field, enabling their radiation to be located using field detectors within 1 to 2 km, which may offer barely enough time for a slow-flying aircraft to take evasive-action. Current detectors are tuned to 50 to 60 Hz mains power. There is no prospect of detecting non-radiating cables reliably without emission. A number of active systems are being developed, principally for helicopter use. CO_2 scanning laser systems lead the field at present, but millimetre wave radars seem practicable too. It should prove possible to design the primary radar with the means to detect this particular hazard - which has grown in importance recently with the advent of wire-guided missiles which may leave festoons of wire across valleys.

2 THREAT DATA SOURCES

2.9 PILOT

Threat assessment systems should not expect inputs from the pilot, but should provide warnings to him so that he may correlate against observation where that is possible.

The pilot may sense a wide variety of threats. However, he is unlikely to be able to inject comprehensive information into the threat assessment system for intelligence gathering and correlation with other sensor data, owing to his potentially high workload. Such devices as Helmet Mounted Sights (HMSs) may enable him to inject a threat direction simply. However, categorisation of threats may prove more cumbersome unless, for example, the avionics contain a voice processor able to distinguish spoken threat categories such as 'Five T62 tanks' or 'ZSU-30-gun - 3 batteries'; this may be practicable in the future, but would push the current state-of-the-art in its airborne application. Ideally, the situation should work in the reverse sense, with the threat assessment system detecting threats more ably than a busy pilot and drawing his attention to those threats which are considered significant.

Thus although the pilot is a source of threat data, it would be imprudent to depend upon useful inputs from him for correlation. Rather, he should correlate external to the threat assessment system by relating warnings to observations.

2 THREAT DATA SOURCES

2.10 INTELLIGENCE

Intelligence data will cover a wide variety of topics, but the mean rate of data transfer is unlikely to exceed greatly the rate at which the pilot may use the resulting information.

Intelligence data will be pre-briefed or received in flight, probably via digital data link. Data will include:

- No-go zones
- Targets, ground and air
- Defended areas
- Weapon details
- Safe corridors
- Communications channels
- Own ground force details
- Early warning (EW) data
- Weather, etc.

The amount of information which could be received in flight will be phenomenal: the capacity of one Joint Tactical Information Distribution System (JTIDS) net for example is 30 Kbits/s and there could be some 20 or 30 nets. However, scenario studies will limit the scope to that which is usable by the avionics processors and the pilot - who has a very low information bandwidth of some 2 Hz. Intelligence data could also be transmitted in flight for use on the ground or by other aircraft, to augment their target recognition capability for example.

3 THREAT CATEGORIES

3.1 THE RATIONALE FOR CATEGORISATION

The threat assessment system will have to satisfy so many constraints that flexibility will be the design keynote.

Threats may be categorised in several ways:

- Immediacy of effect
- Operational implications
- Risk and probability
- Avoidable/unavoidable
- Respondable/non-respondable.

It is vital that the design concepts for a threat assessment system should reflect the interest of both the aircrew under direct threat and the need to execute the mission in hand. This raises some interesting questions:

- Should 'minor' threats (whatever that term means) be notified to the pilot? He is going to face some risks in any event.
- Supposing a 'minor threat' threshold is set prior to flight- could the pilot be inundated with irrelevant warnings?
- Which is most important - the pilot, his aircraft, completion of attack or successful return to base?
- Should the pilot be notified of a threat against which he has no countermeasure?
- Under what conditions should countermeasures be initiated automatically?
- Should the pilot be told that countermeasures have been initiated - does he always need to know, when he may have other problems on his hands?

3 THREAT CATEGORIES

3.2 IMMEDIACY OF EFFECT

The employment of time or immediacy as a sole threat index is not practicable, although the data is relevant to relative threat weightings.

One parameter for comparing threats might be their relative timescales. A SAM radar, for example, may change mode indicating that a missile is about to be launched which will have an estimated time of t_m seconds to impact. The attack aircraft might also detect that the tank squadrón which it is about to engage is defended by a ZSU-30 anti-aircraft gun and that the aircraft will enter effective gun range in t_g seconds.

This example raises several points:

- The missile time, t_m, will count down inexorably once launched, but the time to gun engagement, t_g, can be changed by the aircraft pilot as he manoeuvres.
- The threat from the missile, once launched, will vary as a complex function of aspect angle, altitude, speed, etc., but could be considered as essentially constant when the missile is fired from within its launch success zone.
- The threat from the gun will also be a complex function, but essentially it will increase as the separation from the aircraft decreases, being initially (probably) lower than that from the missile.

Clearly, then, while time is a very relevant factor, there is little value in its use as the sole selector of rank-order, or importance. (See Topic 3.4(a) for further analysis of this problem.)

It is evident that the logic supporting a threat assessment system will need to be comprehensive, adaptable and responsive to changing situations even within the timescale of a mission. This flexibility will enable us in 'beg the questions', in some senses, since many of the choices will be left to the eventual operators of the system. Some queries remain, and will be discussed in following topics.

3 THREAT CATEGORIES

3.3 OPERATIONAL EFFECT

The use of operational effect to categorise threats may prove conditionally useful as part of a set of threat parameters.

Threats might be categorised according to their likely relative effects on the success of an operation. A successful operation could be defined as one in which the weapon load is effectively launched against a valid target, even if the aircraft is lost in the process. This definition will not suffice for a photo reconnaissance mission however, where return to base is essential at present.

In either case threat categorisation could be problematical. Small arms fire, for example, could start an aircraft fuel fire, which would burn for some indeterminate time before the aircraft was affected operationally. This indeterminate time could traverse the time boundary marking a successful mission; in other words, the threat posed to a mission by ground fire is indeterminate, since the effects of ground-fire may take some time to accrue.

Generally, however, the criterion of threat to mission completion is felt to be useful. In unemotional terms, a SAM which will arrive after aircraft free-fall weapon release is no threat to that part of a mission. It is of course a threat to crew and aircraft, and the problem of warning still has to be addressed; EW might permit successful countermeasures but 'spoil' the aircraft's attack routine, which may or may not be repeatable. By itself, therefore, operational effect is an inadequate parameter, but it could well contribute to a more comprehensive threat index.

3 THREAT CATEGORIES

3.4 RISK PROBABILITY AND CONFIDENCE

The establishment of orders of threat will necessitate the use of statistics and models, which will require validation to establish any degree of credibility.

Inevitably, figures will have to be invoked to describe and specify threat risk probability and confidence. In order to compare the risk from attack by a SAM and an anti-aircraft gun, the following could be considered for each threat:

- Probability of being in range
- Probability of detection
- Probability of engagement by the enemy
- Probability of receiving a hit
- Probability that one hit will defeat the mission object
- Variation of probabilities with range and time, etc.

With so many probabilities involved however, each likely to be poorly established, there is little point in performing comprehensive calculations, since the quality of the results will only reflect the summed uncertainty of each input parameter. Instead, algorithms may well serve to provide a model sufficiently accurate for comparison. (A simplified example of that approach is given in the next topic.)

Such models could assess alternative courses of action to reduce the threat, and could offer assistance to the pilot.

The degree of confidence which can be placed in such assessments will be of paramount interest, and a compromise will be necessary between algorithimic complexity and computing brevity. Computations will be valid only at the instant of data sampling and will require real-time updating, so that time-consuming optimisation routines would require careful design. Essentially, some confidence will be generated when simulation has been validated against realistic scenarios, but at best any such system would depend upon intelligence assessment of threat parameters. At worst however, it should prove preferable to leaving the problem for the hard-pressed pilot.

One major uncertainty will be that of recognising the potential target. The approach most likely to succeed, and which emulates human behaviour, is to correlate between different sensors and to build up confidence by integrating the repeated (snapshot) recognition process. This will be discussed in Topics 5.5 and 5.8.

3 THREAT CATEGORIES

3.4(a) COMPARATIVE THREAT ASSESSMENT - EXAMPLE

Comparisons between threats may present conflicting information; an intelligent threat assessment system could evaluate alternative tactics to reduce the threat, and present options to the pilot.

Intelligence and aircraft sensors supply the following data (example only):

	Surface-to-air Missile	Anti-aircraft Gun
a) Effective range	6000 m	3000 m, but see (b)
b) Probability of kill/round	0.25	1000 ÷ (Slant range in m)^2 (Both arbitrary choices)
c) Fire rate	1 per 15 s	800 rounds/min
d) Lethality	Single-hit kill	Single-hit kill

Aircraft data and the geometry are shown on the diagram.

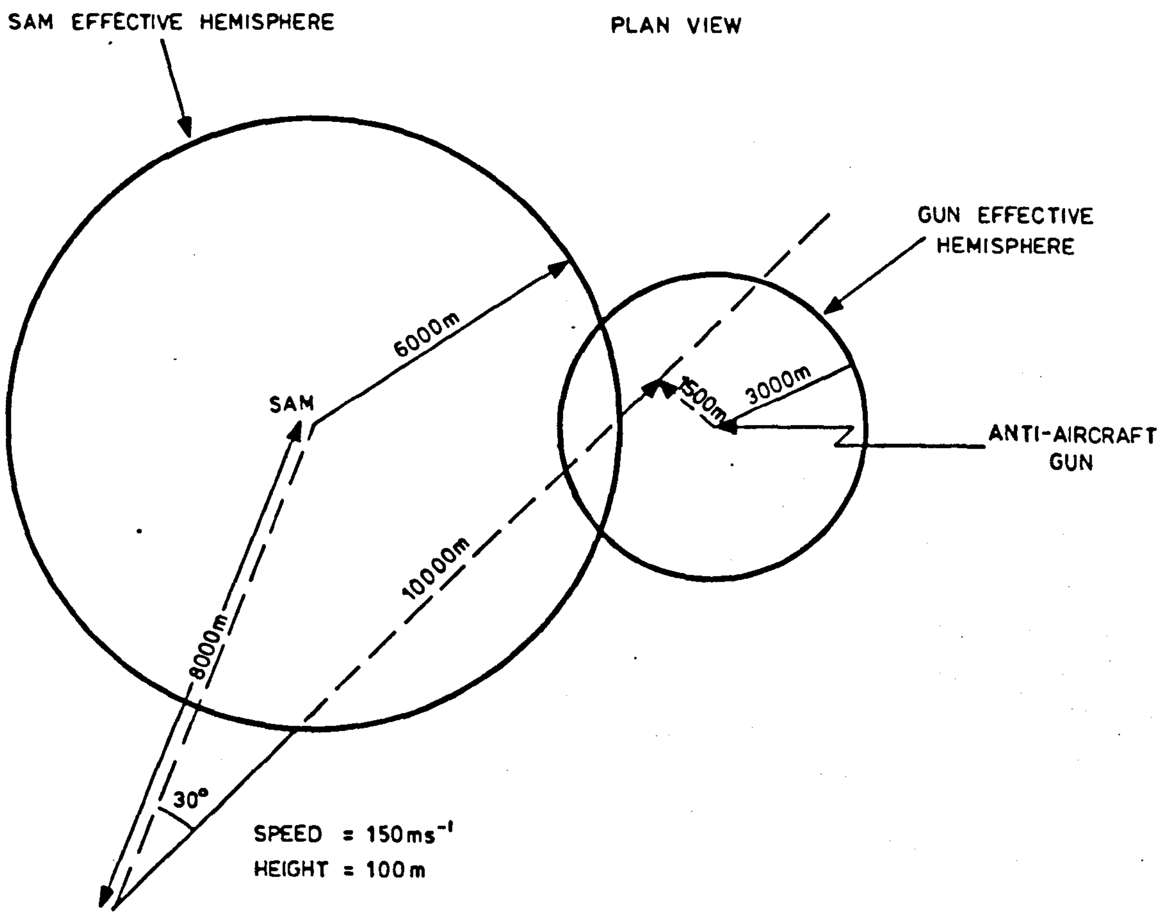
The two apparently similar sets of data present quite different problems; the missile problem is principally geometric in nature, with the calculations revolving around determining time before entering the lethality hemisphere, time in that hemisphere and hence how many missiles can be launched, each with the same kill probability. The anti-aircraft gun calculations are concerned with calculating the probability of kill which varies continuously with slant range to the target, throughout the overpass.

The results given by the two simple models are interesting in themselves:

	Surface-to-air Missile	Anti-aircraft Gun
Time to enter danger zone (s)	16	47
Time in zone (s)	60	40
Number of missiles launched/fired	4	533
Total kill probability estimate (assuming each launch is an independent event)	0.68	0.773

The sample calculations show that there is a 68% probability of being killed by the missile battery, starting in +16 s and ending in +76 s, whereas the anti-aircraft gun threat affords a 77.3% kill probability starting in +47 s and ending in +87 s. So, the gun presents a potentially greater, but less imminent, threat.

The approach employed suggests that the aircraft should carry real time simulation models of known threats, should use its sensors to locate and recognise those threats, and should analyse the threat implications by some form of simulation. The results of the analysis would not necessarily mean much to a pilot acting under pressure. However, he would have to make a decision based on two conflicting factors - lethality and immediacy. The simulation approach could help here, too. The computation could assess alternative tactics for presentation to the pilot. For example, he could accelerate to 300 ms^{-1}, on the same track. The model would show the anti-aircraft gun risk reduced from 0.773 to 0.528 and the missile risk reduced from 0.68 to 0.44. The missile risk being less than the gun risk, the simulation could further determine a path of minimum risk between them, by steering to port, or of reduced overall risk by steering around the two zones. So, instead of offering confusing information to the pilot, he could be offered a choice of tactics to reduce or avoid the threats. If the system were aware of the pilot's objectives, these could also be taken into account.



Simplified Gun Effects Model

For gun, kill probability = $\frac{k}{d^2}$/round

where:

d is instantaneous slant range from gun to aircraft

and:

k is an empirical constant.

Then:

$$d = (R^2 - 2\,vt\sqrt{R^2 - h^2 - s^2} + v^2t^2)^{\frac{1}{2}}$$

and overall probability of a hit:

$$P \Omega 1 - \left[\pi\left(1 - \frac{k}{R^2 - 2\,vt\sqrt{R^2 h^2 - s^2} + v^2t^2}\right)\right]$$

as t varies from 0 to $\frac{2}{v}\sqrt{R^2 - h^2 - s^2}$ in intervals of $1/f$,

where:

f is the gun fire rate in rounds/s
R is the effective gun range
v is the aircraft velocity
h is aircraft height

and:

s is across-track miss distance to the gun.

3 THREAT CATEGORIES

3.5 AVOIDABLE/UNAVOIDABLE THREATS

Unavoidable severe threats may be considered as candidates for display on a central warning system.

Once it is established within the aircraft system that a threat is imminent, it may be categorised according to whether it need be experienced or not. A missile which is within a few seconds of impacting the aircraft, and against which countermeasures either do not exist or have failed, represents a high, unavoidable risk.

Attacking a tank squadron, point defended by SAM and anti-aircraft gun batteries, represents a risk which the pilot's mission may decree should be taken. The risks differ in that the second is strictly avoidable, and a 'logical' system design could therefore sensibly recommend that the attack be broken-off if the risk exceeds some threshold. A logical system, in the first case, might recommend that the pilot take some drastic action. This type of unavoidable risk perhaps enters the area of immediate danger warning, and is on a par with an engine fire warning; it is therefore in a separate category and could be considered for special display priority.

3 THREAT CATEGORIES

3.6 COUNTERMEASURES

Threats can be categorised by the availability of countermeasures, and by the need to initiate such countermeasures manually or automatically.

Threats may be divided into those which can be countered and those which cannot. To some extent this division may be set by the pilot, since he may elect to 'run silent' and not use jammers, so that an approaching IR missile might be in either group according to pilot selection.

Countermeasures may also be considered for automatic initiation. The principle criterion for judgement could be response time; a threat for which a countermeasure is available and permitted by pilot selection may be detected with such short notice that the pilot could not reasonably be expected to assess a display and respond in time. A severe threat which is imminent and to which the pilot has not responded (although notified) might be considered for automatic response in some circumstances. Clearly, there is a sensitive divide between allowing the system to respond 'at the last instant' in order to preserve the pilot and his aircraft, and usurping the pilot's control by anticipating his response. There need be no difficulty if the system is flexibly designed and developed with co-operation from operationally experienced crews.

4 THREAT ASSESSMENT SYSTEM CHARACTERISTICS

4.1 THE PILOT INTERFACE

The introduction of an automated threat assessment facility must be accompanied by a comprehensive, yet simply operated, system for configuring the facility to suit the individual pilot, the particular phase of flight and the contemporary threat environment.

The severe operating conditions forecast for low-level battle zone flying in the 1990s will impose a high workload on the pilot of a single seat aircraft. Survival may necessitate high speeds at altitudes lower than those currently achieved with terrain following systems. Pilot-use of terrain screening may involve manual terrain avoidance in two axes. In addition he will be faced with a burgeoning variety of threat sensors and weapons designed to detect and impede his progress towards his target.

The pilot will therefore be faced with a widening range of tasks, and some assessments indicate that his workload will be unacceptably high. The obvious solution is to provide some levels of automated assistance in those areas where such assistance is both practicable and acceptable to the pilot. These two factors are not necessarily compatible, and it seems reasonable to suppose that pilots will wish to employ much of their time flying head-up, when in the battle zone, in order to maintain the safety of their aircraft and an awareness of their situation *vis-à-vis* other aircraft targets and threats. Ideally, therefore, automated assistance should be directed to assisting them with these principal tasks, head-up. During other phases of operations, and during exercises to build confidence in automated systems, head-level or head-down displays may prove valuable and practicable.

A concept emerges which favours two types of display:

- Head-up presentation of essential data relating to aircraft situation, targets and threats only
- Head-level and head-down displays which may be referred to, where supportive information is available, to confirm or expand the head-up data.

Moreover, the data presented head-up must be filtered to present only that 'of interest to the pilot'; herein may lie difficulties, since the term is essentially subjective; the pilot may prescribe parameters before, or during early phases of, a sortie which will define his limits of interest. Threats, for example, could be prescribed by significance level, imminence, potential lethality, closeness of encounter etc. Thereafter threats as such, need not appear head-up; instead, threat reduction measures could be presented to the pilot in the form of advised headings or steer signals, or as an advice to initiate countermeasures. Even the last category, initiation of countermeasures, need not appear if the pilot had confidence in their automatic initiation and had so configured the system.

The threat environment is dynamic; not only are new weapons being continually introduced, but new data is constantly being discovered about existing weapon performance. Moreover, the siting of threats in a battle zone may be known to the pilot through a combination of sensor data and intelligence. Intelligence may be stale, so that a mission planned on such information may need to be replanned in real-time, using either observed data or information received in flight. The threat assessment system, if it is to be effective, must therefore be highly flexible, since it must:

- Accept new target/threat parameters easily, in its operational role, as well as during development
- Correlate poorly registered intelligence, sensor, and pilot data in 'real-time'
- Adapt to 'rules of the day' or even 'rules of the instant' under pilot control which may be imposed before and during a sortie
- Provide high integrity advice to the pilot, and/or initiate countermeasures with similarly high levels of confidence.

The next topic examines some control and display implications in more detail.

4 THREAT ASSESSMENT SYSTEM CHARACTERISTICS

4.2 DISPLAYS AND CONTROLS

Integrating the threat assessment system equipments should reduce the number of separate equipment controls and displays, but some will be needed for the integrated system; in particular, a Helmet Mounted Sight (HMS), Inject/interrogate, marker control, a head-level/-down threat display format and 'attention-getters' seem likely requirements.

The threat assessment system will bring together a variety of facilities, including possibly:

- Forward-looking Infra-red (FLIR)
- Radar
- Radar warning
- Missile warning
- Laser warning
- Identification Friend or Foe (IFF)/NATO Identification System (NIS)
- Cable Warning
- Manoeuvre Monitor, etc.
- Countermeasures - Jammers, flares, emission control, etc.

Some of these have other uses; radar could have a mapping/navigation role, for example. Generally, however, their combined purpose can be described as 'Target and Threat Detection, Location, Recognition, Identification and Response'. From the pilot's point of view, combined functions may be grouped as:

a) Broad scan for threats and targets of opportunity.

b) Search for planned targets, with evasion/reduction of threats.

c) Attack following a reduced-risk flight profile *vis-a-vis* terrain and threats.

Category (a) might be applied during transit to, from and over the Forward Edge of the Battle Area (FEBA), and into enemy territory, and implies no detailed pre-knowledge of enemy dispositions. Sensors would be making wide searches, and the pilot may wish to direct them towards visual contacts or detail of interest on radar or FLIR displays; there is an implied need for an HMS with an inject facility to mark spatial angles of interest, and pilot controlled markers on radar, FLIR and possibly head-up display (HUD). Recognition of a threat implies the need for counter-measures; the pilot may wish to assess the threat himself and store/transmit threat data for intelligence purposes. The need thus emerges for some form of threat assessment display, for use when workload permits. The display need not be complex, but should encourage simple recognition of the threat situation, and enable the pilot to mark and store threats of interest. Countermeasures will be related to the threat; the display may conveniently indicate countermeasure plans formulated within the system, for pilot approval. These would include: planned threat reduction tracks; emission control advice; flare/chaff ejection advice; and IR/radar jamming advice. Such a display is deemed necessary for peacetime training and to foster confidence. An automatic mode is also suggested which would initiate countermeasures other than evasion. In either case, 'attention-getters' may be required, plus manual initiate facilities.

Categories (b) and (c) imply planned operations in an area of high threat density. The pilot may have little time for head-down displays and would concentrate on finding a relatively safe and effective run-in to the target. He may require to know when the target has been acquired (if it is sensor-discernible), and to make best use of local terrain and knowledge of threats to improve his prospects for survival. While the threat assessment system might be able to plan a reduced risk run-in to the target, it seems unlikely that such a path could usefully incorporate knowledge of the terrain or that the pilot would wish to trust such a plan even if it were available. A compromise would seem necessary, in which the pilot would fly to take screening advantage from the terrain while the threat assessment system would advise which flight-path options offered least risk.

Thus while any steer display must sensibly operate head-up, a head-level or head-down display may still prove necessary to allow occasional pilot assessment of his overall threat situation. Moveover the threat reduction plan must operate in real-time, to give the optimum route to the target from any present position. Depending on the degree of sophistication required, this route could accommodate sequential targets, and could re-order them in real-time.

The warning sensors may provide large quantities of data in areas of high threat density. It would be unreasonable to expect the pilot to assess all such data, and some filtering would seem appropriate, particularly since their range is such that detected threats could be many miles distant. Independent threat sensor displays are not therefore considered vital; instead, the threat assessment system should act as the correlator/filter and should present data of immediate or imminent interest. The suggested head-level head-down display will satisfy some of this aim, but the need is foreseen for 'attention-getters' to warn of approaching missiles or hazards. These could be specific, for example pre-recorded vocal messages, or form part of a central warning system.

Positive identification of targets prior to attack may be required by the Rules of Engagement. A suspected target may be observed visually by FLIR or by radar. Once marked by the pilot, he may need some form of interrogate control, which could be associated with the marker button - on a rocker switch for example.

The system sensors may need to be selective; radar or laser warning may need to be programmed to seek particular types of radiation. Generally, such requirements will either be pre-briefed or well established and seem unlikely candidates for prime cockpit panel area.

As the threat assessment system design concept emerges in the next section, it will be seen that the pilot may also require the ability to specify some of the processor parameters, including:

- Selection of learn and run modes
- Setting of warning time and clearance factors
- Selection of single sensor data in multi-sensor mode etc
- Selection of areas of interest for fine search

There will also be a potential to display tabular ordered lists of targets and threats should they be required.

The indication of targets and threats on FLIRs and radars seems most likely to take the form of simple characters of alphas superimposed over the displays in the correct spatial position. Thus a tank might appear as a letter T and a SAM battery might appear, say, as an S with continually reversing field indicating its dual target/threat role. In the air-to-air role, the radar may present an azimuth/elevation C-scope display (in addition to a plan or B-scope range/azimuth or frequency/azimuth display). It may prove useful to scale radar and FLIR so that their respective displays can be readily compared, or to produce a composite FLIR/radar C-scope display at long-range, in order to reduce the number of independent displays. Such displays could also have recognition alphas overlayed on targets/threats.

4 THREAT ASSESSMENT SYSTEM CHARACTERISTICS

4.3 CORRELATION BETWEEN SENSORS

Sensor correlation is essential for effective threat, location and recognition.

Automatic threat recognition is an area of concentrated research and development effort at present. Few working systems have emerged from the laboratories, but those that are operating tend to employ more than one sensor. Correlation between sensors may not be simple, but it offers greater confidence. Single sensor data from, say, a FLIR might detect hotspots at long-range, but it is more difficult to analyse that hotspot with sufficient confidence to categorise it as emanating from a particular type of vehicle. Efforts are proceeding along the path of single sensor signal processing, and it may well offer greater promise for long-range recognition. At shorter ranges, however, where two or more sensor inputs can be correlated, a more confident recognition will be possible. A threat assessment system design may therefore be expected to contain both single sensor and multiple sensor capabilities.

Threat sensors will include a variety of specialised devices in addition to imagers and radars. Radar and laser warning seem likely, in the foreseeable future, to give good recognition selectivity but poor spatial resolution. So, a radar warning system might recognise that the aircraft is being scanned by a particular Target Illumination Radar, but might locate that radar with an angular accuracy of no better than, say, 10^{o}. A radar might be able to search within that cone of uncertainty to locate a target to within 1^{o} or so, and provide range data. Thus two types of sensor correlation emerge:

a) Parallel correlation between sensors providing similar data but perhaps in different spectral bands

b) Sequential correlation in which the sensors provide complementary data, with one sensor effectively alerting the others.

The threat system design must accommodate both types of correlation if it is to be effective at location and recognition. (Note the parallel to human behaviour, where a) might correspond to optical and aural data, while b) might correspond to peripheral flicker trigger followed by head-turn and detailed foveal examination.)

4 THREAT ASSESSMENT SYSTEM CHARACTERISTICS

4.4 RESPONSE TO FAILURE

Target/threat recognition introduces a new category of failure, failure to recognise with sufficient confidence, which requires special consideration owing to its safety implications.

A threat assessment system should 'degrade gracefully', provide an acceptable low level of incorrect outputs, and should make apparent any failures. This last point is particularly important since failure to display a threat may encourage the pilot to follow a dangerous route which he could easily avoid.

The system should therefore contain redundancy, should select 'best' answers, and should provide 'best' advice. These requirements are consistent with a multi-sensor configuration which can operate in single-sensor modes. Superimposed on such a system will be means for:

- Computing 'best' answers
- Suppressing secondary outputs referring to the same threat, to reduce traffic and display loads
- Detecting partial failure and suppressing or replacing the associated threat data
- Detecting inability to provide data of adequate confidence
- Advising the pilot of such inability
- Detecting and indicating failure.

This list includes the unusual category, which is neither success nor failure, of 'lack of confidence'. For example, the system could be designed to display only threats which had been recognised with greater than some arbitrary fixed confidence level. In a particular situation the aircraft might be entering a threat zone which had been recognised, but with a degree of confidence just less than that fixed level. The threat would be encountered but not displayed, while the system would not have failed in any conventional sense. Clearly such inflexible rules should give way to more elegant display criteria.

The detection of routine equipment failures would be effected using the now conventional practices of passive and active on-line monitoring, built-in test facilities, etc. In some instances, the high packing density of microprocessor devices may allow built-in redundancy so that apparent reliabilities may be considerably extended.

5 PROCESSING

5.1 SENSOR INFORMATION CONTENT OF ELECTRO-OPTICAL SENSORS

Sensible upper limits may be established for the information and capacity of future communication channels such as colour TV, thermal imagers and other EO displays.

Information content (I) and capacity (C) are defined by Shannon as:

$$I = \frac{T}{\tau} \log_2 n \text{ (bits)}, \quad C = \frac{1}{\tau} \log_2 n \text{ (bits/s)}$$

where:

T is the period of receiving data

τ is the minimum interval over which signals can change

and:

n is the number of distinguishable levels which could occur in time τ.

These simple relationships can be used to define the upper limit of information content which can be produced by various sensors, as follows.

An EO TV display consists of N lines, an aspect ratio of 'a' (i.e. a x N black and white elements per line) and a frame rate of p/s. Thus the number of elements is:

$$N \times (a \times \frac{N}{2}) = \frac{a}{2} N^2, \text{ drawn in } \frac{1}{p} \text{ s, thus:}$$

the information content (I) = $\frac{a}{2} N^2 \log_2 n$ bits, and:

the system capacity (C) = $\frac{a}{2} \cdot p \cdot N^2 \log_2 n$ bit/s

The value of n relates to the distinguishable variable states which can be displayed in the interval τ. For an EO display this could be 'b' brightness levels, and 'c' shades of colour or temperature (radiation polarisation angle is not generally detected). The full expressions for maximum information content and capacity become:

$$I = \frac{a}{2} N^2 \log_2 (b.c), \quad C = \frac{a}{2} p N^2 \log_2 (b.c)$$

The number of raster lines displayed in an aircraft is likely to be higher than the domestic 580 lines (625-line convention) because the pilot will observe the display from as little as 25 cm. At this distance eye resolution is about 0.1 mm, so that a 10 cm high display would require some 1000 lines to prevent individual lines from being resolved. Technical problems are likely to limit the number of lines to about 800, however.

Typical figures for the two formulae would be as follows:

Aspect Ratio $a = \frac{4}{3}$

Number of lines N = 800

Frame rate = 25/s

Brightness levels b = 7

Colour shades S = 30.

For a colour TV, the information content per frame then becomes:

$I = 3.29 \times 10^6$ bits (= 3.29 Mbits)

and the capacity (C) = 82.3 Mbits/s.

For a black and white thermal imager:

S = 1

I = 1.20 Mbits,

C = 30 Mbits/s

These figures should not be confused with system bandwidth equivalent to $\frac{1}{\tau}$; rather, they indicate the maximum storage and processing tasks associated with the sensors; so a single frame-store for an 800-line full colour TV would require a maximum storage capacity of 3.29 Mbits. This upper limit is based on the premise that each pixel in the format contains independent data, which is generally untrue when terrain is imaged or targets are tracked. Some estimates suggest that domestic TV displays are, on average, well over 90% redundant in their content. (The concepts of data compression will be discussed in Topic 5.3.) Current research work by the Hitachi Company has just produced a full colour frame store for 192 lines and 256 elements per line occupying only 18 kbytes, using an 8-bit microprocessor to pack the data. This compares favourably with the usual commercial colour TV frame store of 2 Mbits for about 510 lines and indicates that research is active in this area.

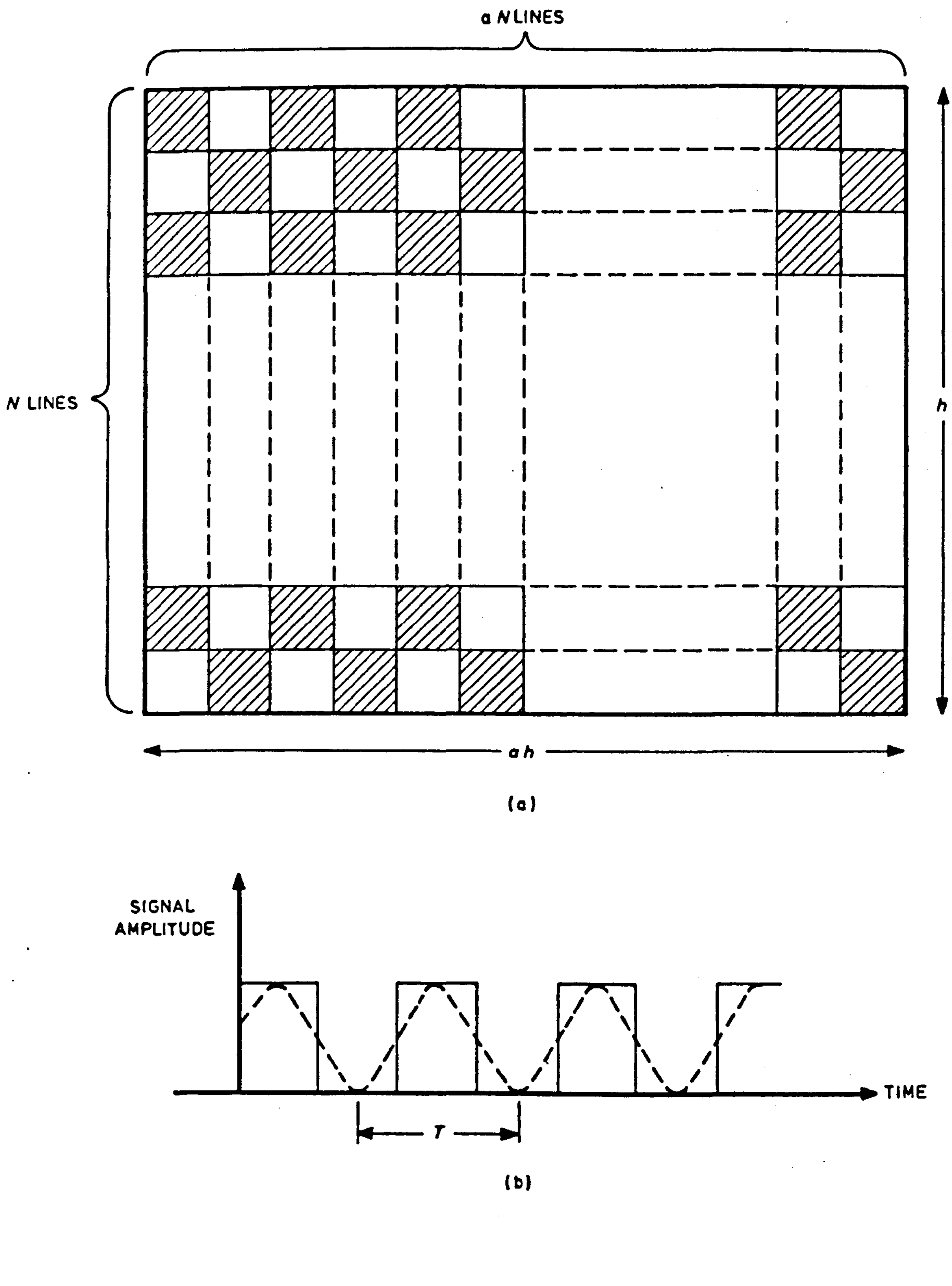


Fig. 5.1 PATTERN USED TO ESTIMATE TELEVISION BANDWIDTH REQUIREMENTS. (b).... WAVEFORM CORRESPONDING TO (a).... ACCEPTABLE APPROXIMATION

5 PROCESSING

5.2 SENSOR INFORMATION CONTENT OF RADAR SENSORS

The maximum information content available from an active radar is related to its particular function/mode/range, but in broad terms is greater than that obtainable from a passive sensor.

For active radar a different approach is required to that outlined in the previous topic. The diagram shows a sector swept out by a radar beam. The number of discrete elements in the swept volume is given simply by:

$$\frac{\theta}{\delta\theta} \times \frac{\phi}{\delta\phi} \times \frac{R}{\delta R}$$

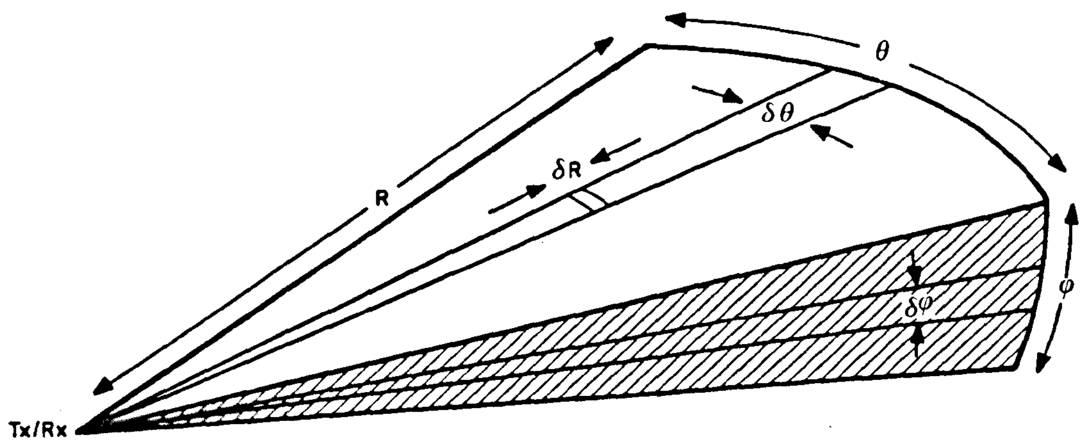
These elements represent discrete echo locations, and the echoes can have a number of distinguishable levels, or cells, within the measurable parameters of amplitude, polarisation and frequency or phase.

At present radar echo polarisation data is usually ignored, although the need to recognise and track targets is likely to change that; observing the response from terrain and targets in differing polarisations can provide useful data. The number of distinguishable polarisation angles (ψ) might have a value of about 18, to give angle measurements to within 10^{o}.

Amplitude data relates to target emissivity, reflectance, orientation and size, and is clearly complex in nature. The number of distinguishable levels may be about seven owing to the relatively high noise associated with such received signals.

Frequency data will prove increasingly valuable for moving target indication (clutter suppression), and doppler beam sharpening, etc. The techniques have different requirements. For air-to-air use, relative velocities of 4000 knots (2078 m/s) can be anticipated and accuracy of measurement is limited by noise, sampling rate and the separation between frequency cells. A typical upper limit might be, say, 2048 frequency cells, giving a relative velocity accuracy of better than 2 m/s (3.75 knots). For air-to-ground use against stationary targets, doppler beam sharpening could be valuable, effectively to reduce azimuth beamwidth. A 6 m-long tank at 3000 m subtends 2 mrad, so that a 1^{o} radar beam (17.5 mrad) needs to be subdivided into about ten parts in order to resolve the target. Thus ten frequency cells might suffice, f = 10.

Typical information content and capacities for two notional radars are shown opposite. Compared with EO and radar, other sensors present small volumes of information, although that information might be uniquely valuable, such as radar or missile warning. The capacity of such information channels is nonetheless minor in terms of impact on processing requirements.



NUMBER OF RESOLVABLE ELEMENTS = $\frac{R}{\delta R} \times \frac{\theta}{\delta\theta} \times \frac{\varphi}{\delta\varphi}$

Fig.5·2 RADAR SWEPT VOLUME

Information = $\frac{\theta}{\delta\theta} \times \frac{\phi}{\delta\phi} \times \frac{R}{\delta R} \log_2 (a\ \psi\ f)$ bits

Capacity = $\frac{I}{T}$ bits/s

TYPICAL VALUES FOR RADARS

	Ground Target Search	Air Target Search
θ	80^o	80^o
$\delta\theta$	1^o	0^o
ϕ	80^o	40^o
$\delta\phi$	80^o	1^o
R	10 km	20 km
δR	2.3 m	6 m
a Distinguishable amplitude levels	7	7
ψ Distinguishable polarisation angles	18	1
f Distinguishable frequency cells	10	2048
I (per single 'frame'	3.58 Mbits	147 Mbits
C (upper limit)	10-100 Mbits/s	50-350 Mbits/s

Note: In principle, the capacity of airborne radar links could be much higher; at present it is limited partly by mechanical scanning constraints, which may be alleviated with the advent of electronic scanning in airborne radars. The capacity required of such systems may then be dictated by the rate at which the data can be processed within the receiver.

5 PROCESSING

5.3 DATA COMPRESSION

Large amounts of data justify the employment of data compression techniques, which can be quite simple in essence. Although storage capacity and transmission bandwidth can then also be reduced, some loss of spatial resolution must be anticipated.

The amount of data potentially available from EO and radar sensors is so large that some form of data compression will undoubtedly be necessary. Data compression can be applied effectively to imaging devices in particular because of their high redundancy. Effective compression can significantly reduce storage capacity and the bandwidth needed for tranmission.

A wide variety of techniques is available; the general public is now familiar with satellite images of the earth which take the form of patchworks of small rectangles; these are produced by a data compression technique.

Most techniques appear to reduce the entropy of the received signal with varying degrees of localisation. Essentially such techniques divide frames into groups of pixels horizontally and groups of lines vertically, and then examine the signals within the groups.

Figure 5.3 shows how five successive groups of five pixels each might be processed along a single line. The top diagram shows the data to be compressed; the centre diagram shows the linear regression line representing each group of five pixels; the bottom diagram shows the reconstitution (by interpolation) in black, against the original in red. The approximation is evidently reasonable. The original data was specified by twenty-six pieces of information - 'y' co-ordinates for twenty-six pixels. The regression line approximation required only ten pieces of information - 'y' for the start and finish of each of five lines. The compression ratio in this simple example is therefore $\frac{26}{10} = 2.6$, which could be highly significant in terms of storage or transmission bandwidth. The reconstituted data erred by 11 squares in 150 = 7.3%. More sophisticated techniques are employed in some applications. The regression analysis need not produce a line: a curve or arc may be represented by polynomial coefficients which may still be less in number than the pixels they represent. Successive regression arcs can be designed to start and finish at one value of x and y, cutting the storage requirement again; (the technique then becomes known as 'spline approximation'). Data reduction may operate across lines as well as along lines and may compress frames together, and so on.

The techniques vary in their compression ratio and the validity of reconstitution. Generally, by reducing signal entropy, some data will be lost and with it some spatial resolution. This may be acceptable where there is strong correlation between adjacent pixels. Decreases in spatial resolution can be offset using edge enhancement techniques.

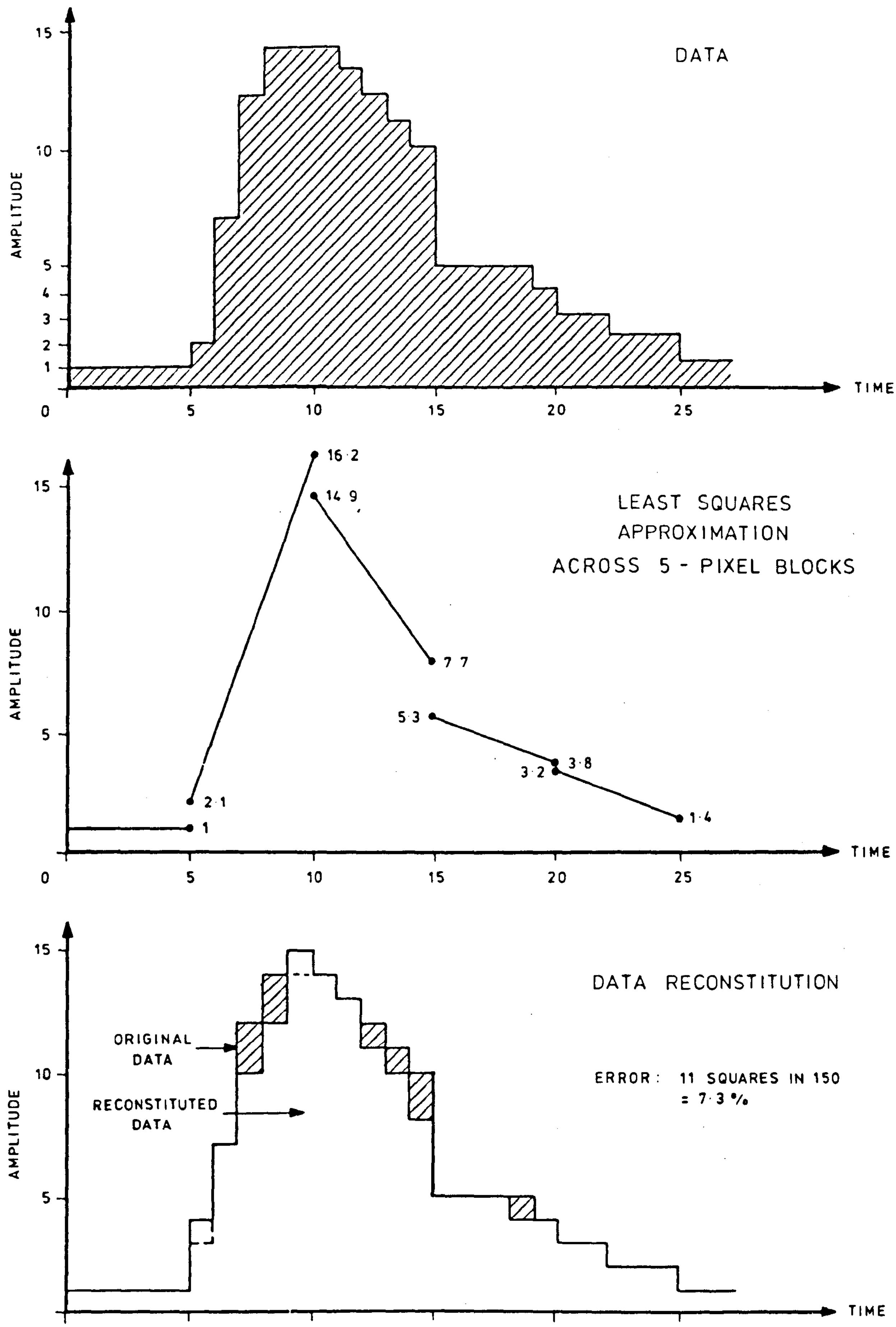


Fig.5.3 DATA COMPRESSION BY LEAST SQUARES

5 PROCESSING

5.4 DATA REDUCTION AND FRESHNESS

Data may be reduced by discarding irrelevant portions, providing relevant details can be specified and detected. Convolution can discard stale data and provides a learning capability.

The previous topic indicated that some forms of data could be compressed without significant loss of information; the advantages of reduced bandwidth and storage capacity could offset minor loss.

However, the amounts of information potentially available from sensors (see Topics 5.1 and 5.2) would still be vast, to the extent that irrelevant data would need to be discarded prior, say, to correlation between sensors for threat confirmation. There is a number of techniques available; their general theme is to identify characteristics of interest and to transfer these in preference to other information. For example, edges could be transmitted while surface detail with no changes, or small-scale variations might not; thermal images might be reduced by selecting detail at particular energy levels or radiation frequencies; radar targets might be selected from within a certain range bracket; etc. Alternatively, data could be transferred periodically; an image pattern might be stored only once in every 5 or 10 frames for example, so reducing mean processing rate requirements. Each technique has its application and incurs a reduction in information entropy which may, or may not, be acceptable in a particular case. The elimination of 'stale' information is also necessary; this can be carried out using some form of the so-called convolution integral:

$$g(t) = \int_{-\infty}^{\infty} f(\tau) * h(t-\tau)\, d\tau \qquad \ldots (1)$$

or its digital equivalent:

$$g(t) \triangleq \sum_{n=-\infty}^{\infty} f(n\Delta\tau)\, \delta(t-n\Delta\tau)\, \Delta\tau \qquad \ldots (2)$$

Equation (1) is generally reduced to: $g(t) = \int_0^t f(\tau) * h(t-\tau)\, d\tau$

where: $h(t-\tau)$ is the system impulse response.

The convolution integral may be thought of as a progressive weighting function, with more weight being given to current than recent samples, 'older' samples being progressively reduced in effect. It has a similar effect to Cathode Ray Tube (CRT) persistence.

The application of a convolution process in a system for compressing successive image frames would gradually erase stale information as it overlayed new signals. This would give the frame store an ability tantamount to learning from what it 'saw'.

5 PROCESSING

5.5 TARGET SIGNAL ACQUISITION

Target signal acquisition is a sequential process which is best achieved using several sensors. The quality of recognition performance improves both with time and practice.

Acquiring a target signal implies detection and recognition. In the context of this topic it is initially valueless to distinguish between a target and a threat, since in many cases targets will be threats and vice versa.

The ability to detect a target automatically will depend on certain target characteristics:

- Movement relative to stationary background
- Movement radial to sensor (radar)
- Characteristic shape and colour (visual spectrum spatial distribution and intensity)
- Thermal shape (thermal spectrum spatial distribution)
- Thermal intensity spectral distribution (temperature and surface characteristics)
- Electromagnetic emissions (other than thermal and visual)
- Etc.

Human 'target' acquisition is highly developed, and possibly therefore worthy of emulation; it comprises a number of steps:

a) Coarse detection of area of interest. This is generally achieved by locating a sound, sensing movement at the periphery of vision or from memory.

b) Close search in area of interest.

c) Tracking moving object to give stationary retinal image against moving background.

d) Target recognition against some *a priori* stored model (memory of similar objects), and correlation between sensors (ears and eyes, rates of sightline movement, range estimates, etc.)

e) Repeated correlation between target parameters and stored knowledge of targets.

For example, a distant 'helicopter' travelling at M1.0 would be rejected as not realisable, and a fixed-wing aircraft would be considered instead.

Similarly, a fixed-wing aircraft in hover would be a Harrier or would be rejected. In both cases, recognition is not based on 'single-shot recognition', but employs an integration process to establish behaviour between 'frames' followed by an assessment of results.

The target, or threat, detection/recognition process to be built into an aircraft would sensibly work along broadly similar lines to the human system.

- 'Coarse' sensors such as radar warning and intelligence (which may be stale or inaccurate) may give early clues to target location. Additionally, high speed scans of areas of interest, and the use of MTI to highlight movement, would pick up hot or moving objects rapidly: some of these might be targets.
- Areas of interest may be scanned more carefully using magnification, more precise algorithms, correlation between sensors, etc.
- Tracking gates may be set about objects to stabilise them against movement of sightline or background.
- A multi-sensor image pattern of the target may be formed and compared with stored data to classify the target as being a tank, a jeep, a fighter, etc., or possibly to classify as being a particular type of tank, vehicle or aircraft.

The processes will involve correlation and probability; but then, human performance is far less than perfect in such circumstances. The quality of recognition will improve with increased tracking time and best use of sensor data.

The human 'system' does not come ready-made, and practised operators achieve much better detection and recognition performance than those new to the task. The complexity of the task makes this understandable. On the same basis it would be prudent to design an aircraft system both with a basic set of target data, and with the ability to learn, to adapt, and to recognise new targets as they arise. This may not be quite so difficult as it would appear, provided the capability is built-in from the beginning.

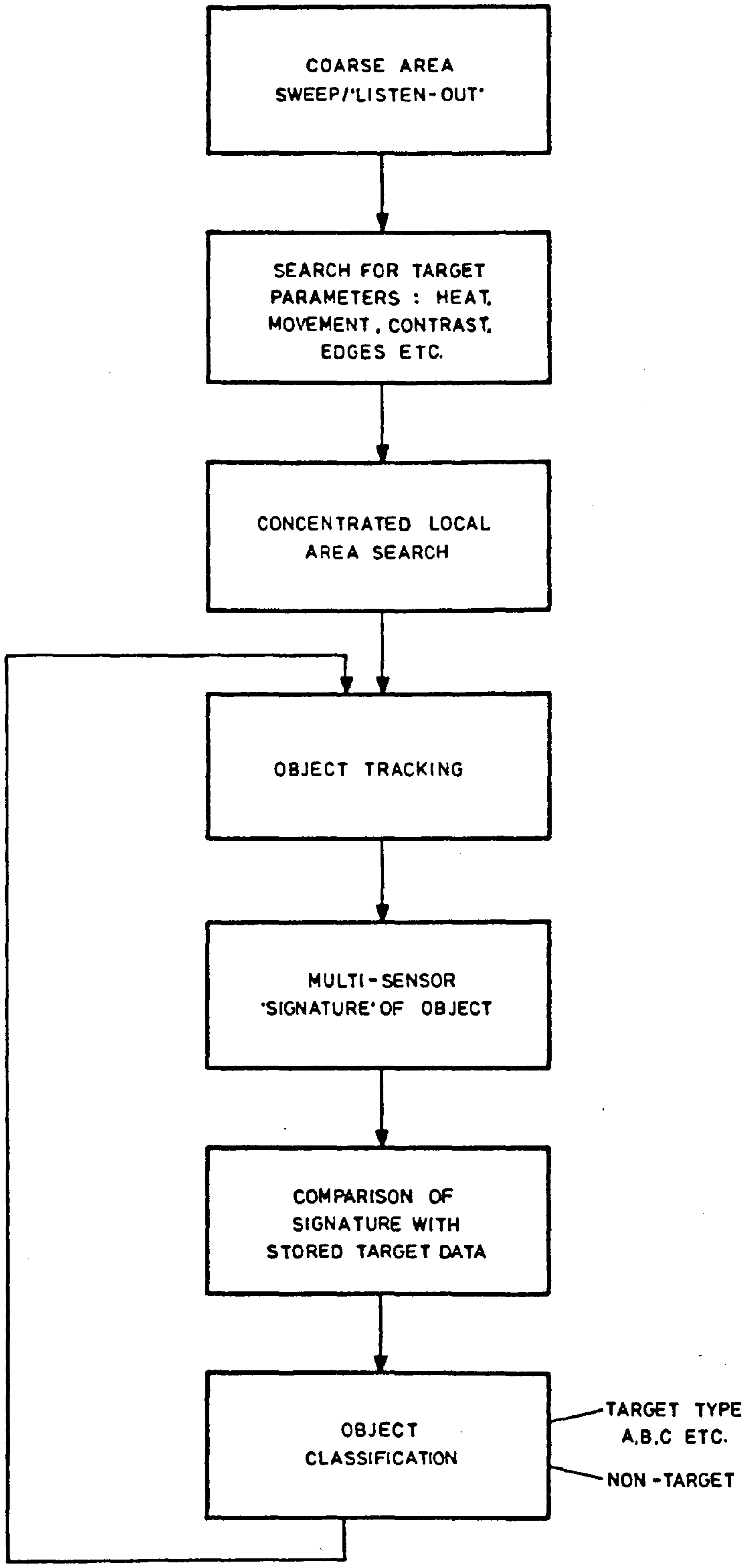


Fig. 5.4 TARGET SIGNAL ACQUISITION

5 PROCESSING

5.6 ACQUISITION TECHNIQUES FOR STATIONARY TARGETS

Many techniques are available, but good automatic performance may require a heuristic design.

Target acquisition by any means (human or automatic) requires that the target should present recognisable characteristics. For example, man-made objects often exhibit straight edges connected by areas of constant intensity. One technique used with TV or FLIR imagery is to search for these characteristics. The pixels in each raster line can be searched for sudden changes of, or plateaus in, intensity. The line may then be divided into 'segments' of interest, which are stored for comparison with adjacent lines and successive frames. The display field store may also be viewed in the Z-axis intensity; if the intensity of each pixel is stored as a number, then a scene may be reviewed to observe all objects presenting a prescribed intensity level. A 3-bit intensity word enables eight independent intensity level views to be presented, but of course they could be presented in selected pairs, threes, etc. Different intensity levels offer useful data; for example there is a correlation between signal magnitude and distance. (Scene dynamic range is typically about 1000:1, compared with a CRT range of some 20:1.)

Colour contrast is used very successfully by the human eye to highlight targets against their background. The human eye can distinguish 130 hues across the visible spectrum; a machine is unlikely to match that performanc in real-time. More useful, particularly at night, is the ability to determine the wavelength corresponding to the maximum intensity of thermal radiation from a body. This wavelength corresponds directly to black body radiation temperature.

Using Wien's Displacement Law:

$$\lambda_{max} \times T = 2.9 \times 10^{-3} \text{ mK}$$

Using spectral filters in the thermal band therefore, corresponds to dividing a viewed scene into temperature bands. In addition to the obvious merit of discriminating hot objects, this technique can be coupled with those above to look for edges, contrast plateaus, etc. at a variety of temperatures. Additionally, many man-made surfaces are flat and tend to reflect sky or sun radiation, which are both characteristic; sky temperature may vary between ambient clouds (237 K) and deep space (4 K), while sun temperature is of course, some 6000 K. A flat surface reflecting a clear night sky, for example, will emit a complex pattern comprising internally generated thermal energy and reflected sky energy, both radiations being differentially absorbed by the atmosphere. The contrast between this pattern and its background may act as a means of detecting and recognising an object.

The features which comprise a particular man-made object may be considered under several headings: it will emit and reflect radiations; it will present an intensity of signal relating to range; its dimensions will have a limit in ratio (length over height); it may have hot spots in particular areas and at known temperatures; and so on. Using these and similar techniques it is possible to detect non-moving targets against a background, and to classify them broadly into categories with some degree of assurance. The recognition process inevitably involves correlation (as in humans) between observed and stored characteristics.

Systems presently divide into automatic target cuers, and semi-automatic systems. In the second case, a human operator may detect an object of possible interest, and cause the system to track it. This has potential advantage if the system has a recording medium; the characteristics of an object which is later confirmed as a target can be learnt by the machine. Since an object will present an almost infinite number of aspect angles and dynamic contrasts, a heuristic technique may be preferable, or a valuable adjunct, to a pre-programmed system. Once the system has learnt those characteristics which correspond to targets of interest, it may then acquire targets automatically. By this means the system can learn about new targets, new camouflage techniques, etc.

This topic has merely touched upon some of the techniques for acquiring stationary targets; many system designs are currently in development and some are in early service. The next 5 to 10 years should see a vast increase in operational systems which until now have been held back by the large amounts of storage and computing needed to operate in real-time.

5 PROCESSING

5.7 COGNITIVE MAPS

The core of the threat assessment system should be a 3-D map similar in concept to that stored and used in the human brain.

Humans generate an internal cognitive map, which is a 3-D mental image of the space surrounding them. By this means it is possible to walk through unknown woods, or a strange town, taking a roughly circular path to arrive at the starting point. Similarly, it is possible to catch a ball without watching hands, since their position is known "instinctively".

Humans walking along a busy street carry a mental zone in front of them such that avoiding action must be taken if another human or an object enters that zone. The zone accommodates intentions to turn right or left, and objects moving at different relative speeds.

This zone has been investigated by researchers, who find it to be broadly elliptical in plan, with the walker roughly at one focus and walking in the direction of the other focus. A similar system is operated when driving cars; the elegant sophistication of the human system can be observed at busy roundabouts.

Clearly it would be difficult to match the flexibility of the human cognitive map in an avionic system, but the principles may be emulated since the machine objectives will be similar to those of a human in combat.

A 3-D map may be established in memory, with aircraft sensors indicating current position, height, track and velocity vector. Intelligence data may be written into the map, while sensors can present data for 'imprint' on the map.

A zone, or zones, of interest may be generated on this avionic cognitive map, according to the systems knowledge of aircraft speed, track, height and objectives. Similarly, threats/targets will have a threat zone - moving and elliptical in the case of enemy aircraft. When zones overlap a threat exists, with degree of overlap relating to degree of threat.

This avionic cognitive map should form the core or hub of the threat assessment system, against which all other data may be related to measure its threat significance. The map will require inputs not only from navigation sensors but also from intelligence and from threat sensors: such information should be presented in the form of threat location speed, track, etc., and threat type, so that the appropriate dynamic threat map may be generated. While part of this cognitive map may be displayed, display is not an essential feature, and indeed much of the map information may not be amenable to display. A simple plan display might prove useful to the pilot, but in essence the system should operate automatically.

The figure shows a conceptual plan display, with the avionic system(s) at the bottom, heading towards the top of the page. The aircraft is surrounded by two zones, a dotted threshold zone and a solid threat zone, which will be examined more closely in the next topic.

The missile system to the right would present an impending threat. The size of the enemy threat zones will take account of current aircraft height and vector. A climbing attitude will reduce the gun and missile threat circle diameters, for example, but may increase the fighter threat zone. Ideally, the system should be able to relate terrain at ground level to threat, since threats may be reduced in effectiveness by flying down a convenient valley, for example. However it is not likely that such sophisticated capability will be totally realisable in the immediate future, although a terrain contour map could be used to cue the pilot that such choices exist.

If the aircraft system, depicted as S on the diagram, steered to port to avoid the missile threat, S's two zones would also swing to the left and the guns (Z) would appear as threats.

System knowledge might include T_1 and T_2 as targets to be engaged, in which case the system would compute a path through the defences, as shown. F, the enemy fighter, will have moved from its current position as T_2 is attacked and will be in front of S, which is therefore unlikely to be taken by surprise.

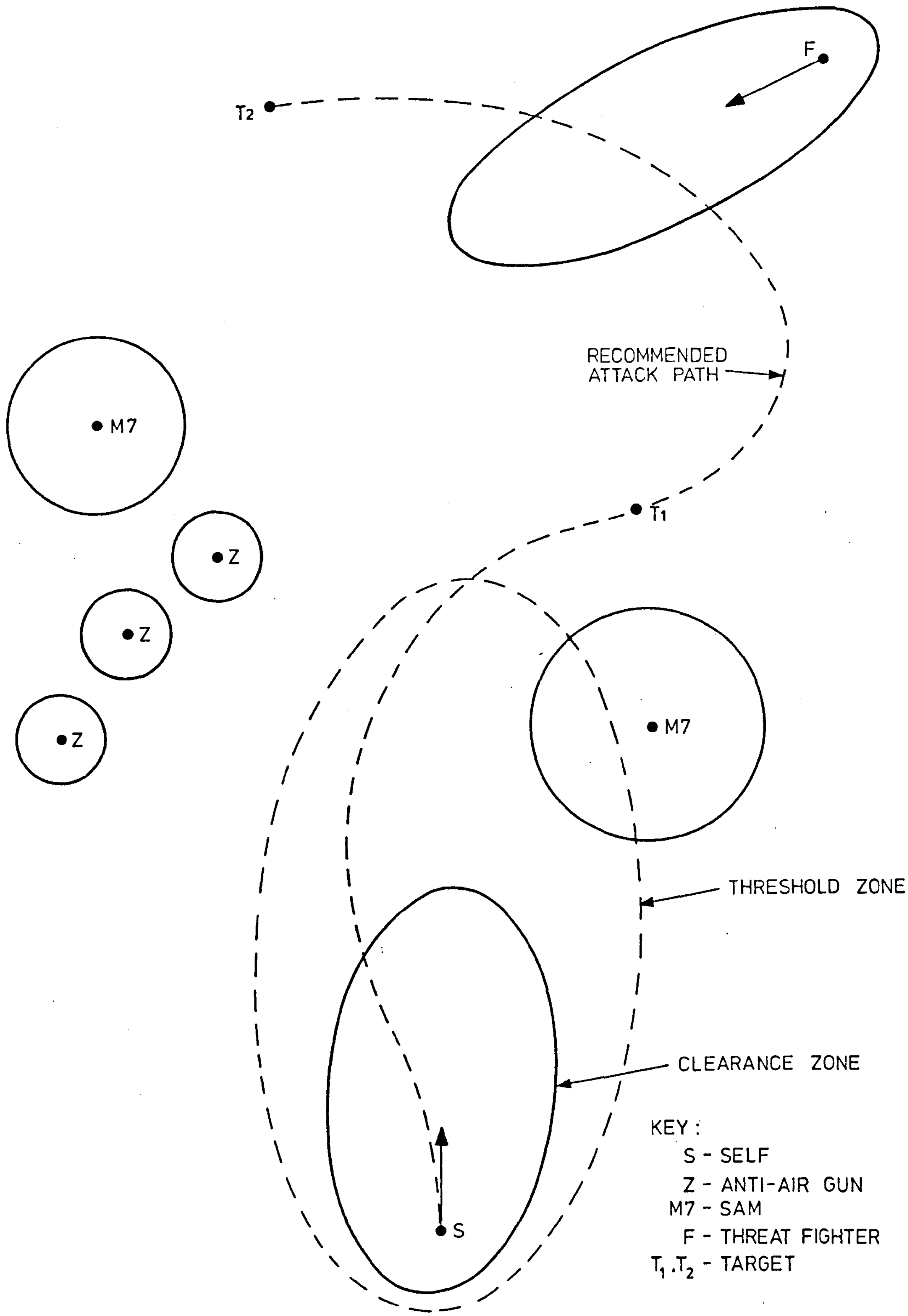


Fig.5.5 COGNITIVE MAP CONCEPT - PLAN ONLY

5 PROCESSING

5.7(a) THREAT ZONE CHARACTERISTICS

The threat and 'own-aircraft' zones in the cognitive map will be designed to give threat advice relating to both imminence and potency of threat. The system will generate 'safe' paths between threats.

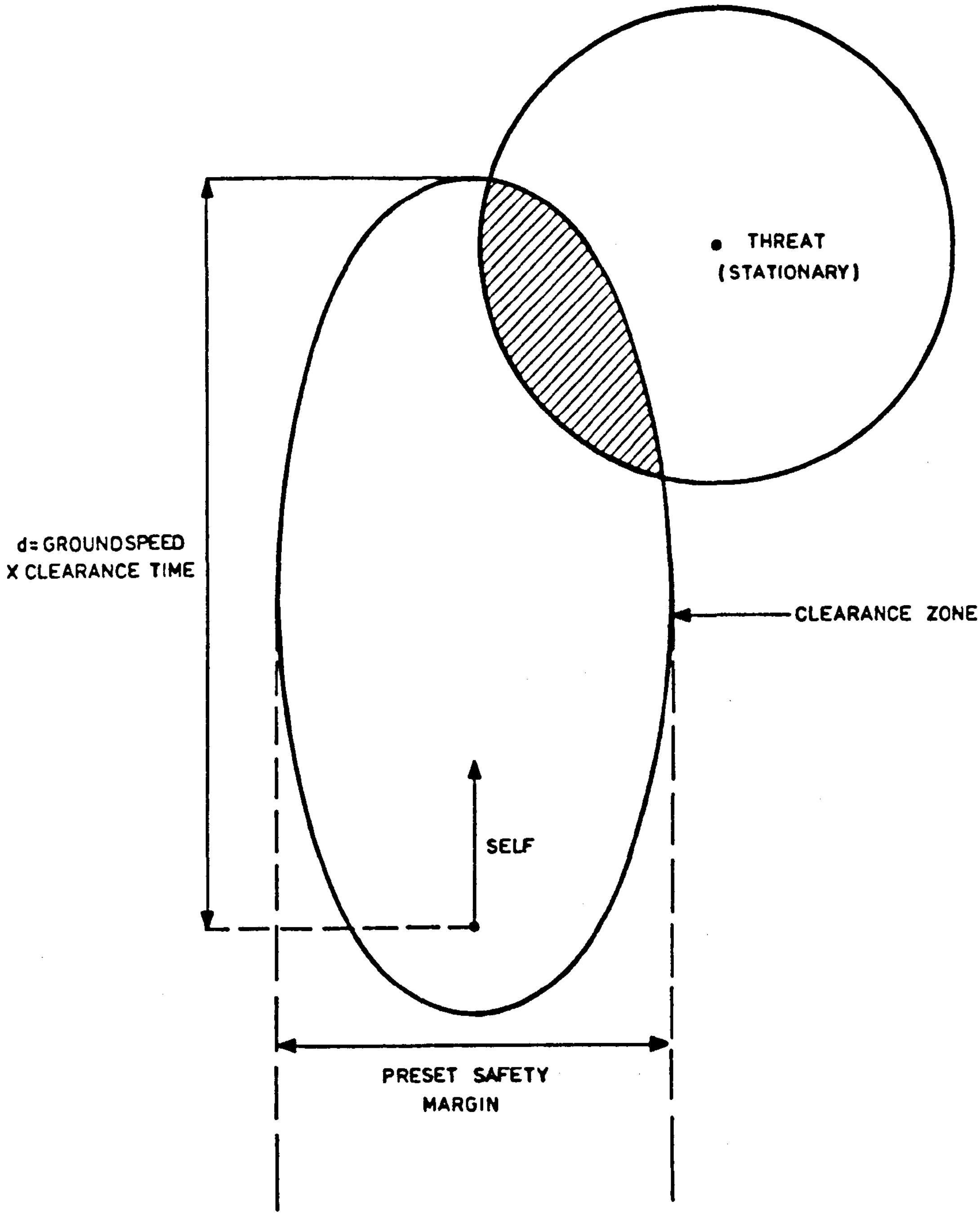
The previous topic presented a cognitive map concept. The zones which might be generated within this map must be related to threat urgency; this is taken to mean some measure which accounts for both imminence and degree of risk. For example, a longer period spent in some zone X might be less of a risk than a shorter time in zone Y, representing a highly accurate weapon.

The threat zones of the cognitive map therefore represent lethality volumes in three dimensions while kill probability at any point within the zone is held as a fourth parameter.

The 'self' ellipses consist of a larger, dotted, threshold zone and a smaller solid clearance zone. Overlap between the volumes of the threshold zone and a threat zone would cause a 'minimum risk flight path' routine to search for a path which would avoid overlap between the threat and clearance zones. In effect the routine would probe between the known threats for a minimum risk path to the target. Risk levels may be selectable to reduce time to reach target.

The major axes of the threshold and clearance zones will also be pilot-selectable, most probably on a time basis. The threshold might be set by the pilot for, say, 1 min 30 s, and the clearance zone for, say, 45 s. The zone major axes would therefore increase with increasing speed.

Zone minor axes serve to provide a safety margin to either side of the aircraft where possible, and to allow choice of the minimum risk path where risk is inevitable.



DEGREE OF RISK IS A FUNCTION OF THE SIZE OF THE COMMON VOLUME, AND THE LETHALITY PROBABILITIES WITHIN THAT COMMON VOLUME APPROPRIATE TO THE PARTICULAR THREAT.

THE PRESENT SAFETY MARGIN REPRESENTS PILOT CAUTION.

Fig. 5.6 ANALYSIS OF DEGREE OF RISK

5 PROCESSING

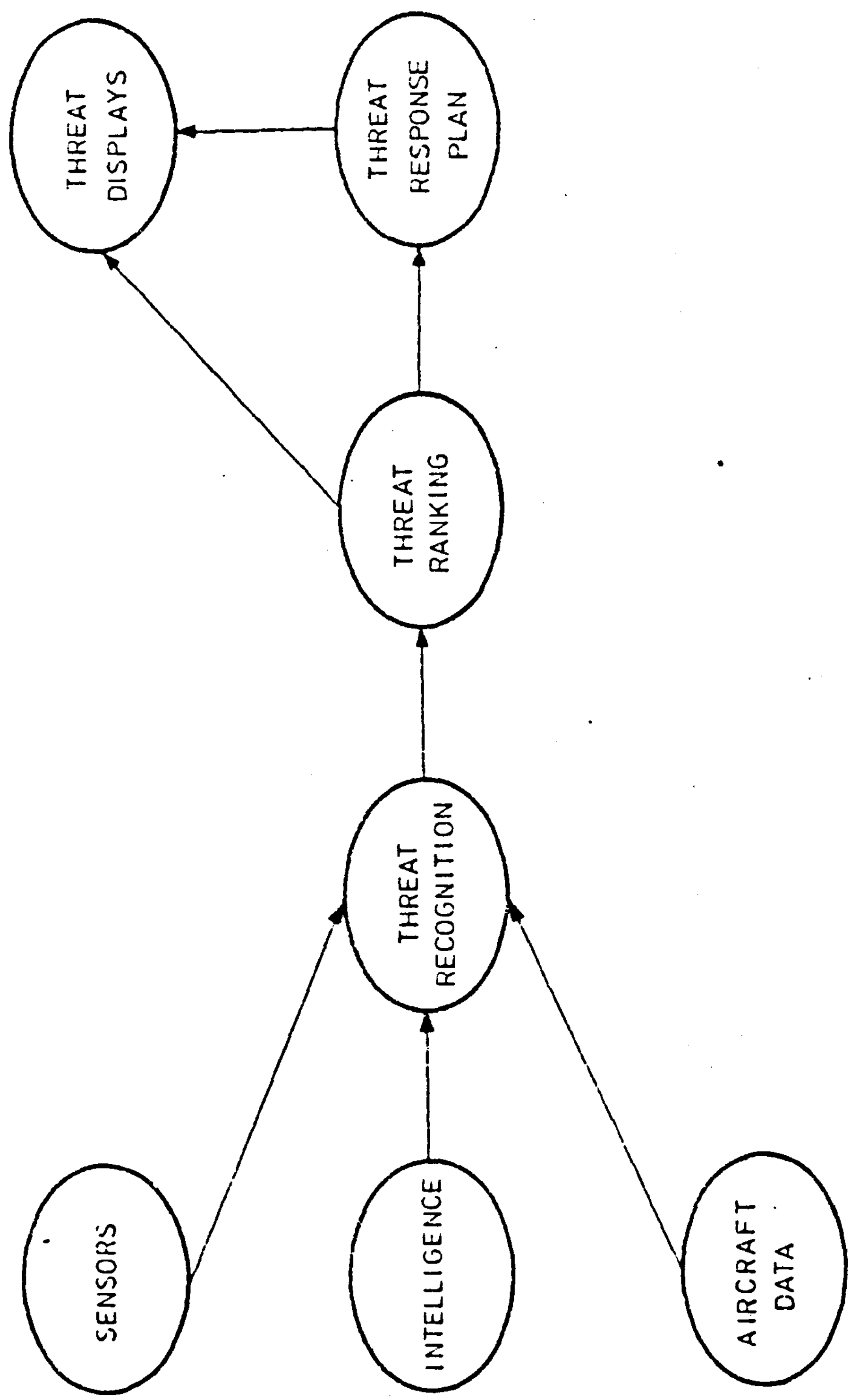
5.8 TARGET/THREAT RECOGNITION - OVERVIEW

The recognition systems must be flexible and adaptable if they are to achieve a reasonable performance at useful ranges.

Targets and threats are closely related, and in consequence, their respective detection and recognition share common facilities. One distinct difference between targets and threats concerns their recognition range. There is evidently greater emphasis on the need for long-range threat recognition; ideally, threats should be recognised at sufficient distance to avoid them, whereas targets may need to be recognised only at sufficient range to press home an attack. Of course, some threats are targets, and vice versa. The point may be illustrated by considering an attack against a 'tank' which turns out to be a tracked-SAM carrier; clearly, a pilot would prefer to differentiate between the two beyond SAM effective range, rather than beyond a (lesser) anti-tank weapon release range.

Bearing this close but distinct relationship between target and threat recognition in mind, the figures show two successive levels of Data State Design (DSD) for threat assessment only. At the first level, the system is presumed to include the processes of recognition, evaluation and response to threats. At the second level, the formation of a threat register can be seen; threats are ranked according to their relative significance before presenting them to the cognitive map where the bases for threat reduction can be established. Reduction can be achieved by evasion, or active countermeasures (chaff, flares, jamming and emission control, etc.) or both. Intelligence data can be used in real -time to indicate the suspected location of threats to direct countermeasures, and even possibly to adapt the threat models which will form the basis for threat recognition-by-comparison.

The processes involved in recognising targets, and on which the credibility of the processing task will be based, are not necessarily familiar. Following topics explore some aspects of target recognition with a view to indicating the magnitude of the processing task.



" TO RECOGNISE, EVALUATE AND RESPOND TO THREATS "

Fig. 5.7 THREAT ASSESSMENT SYSTEM - FIRST LEVEL

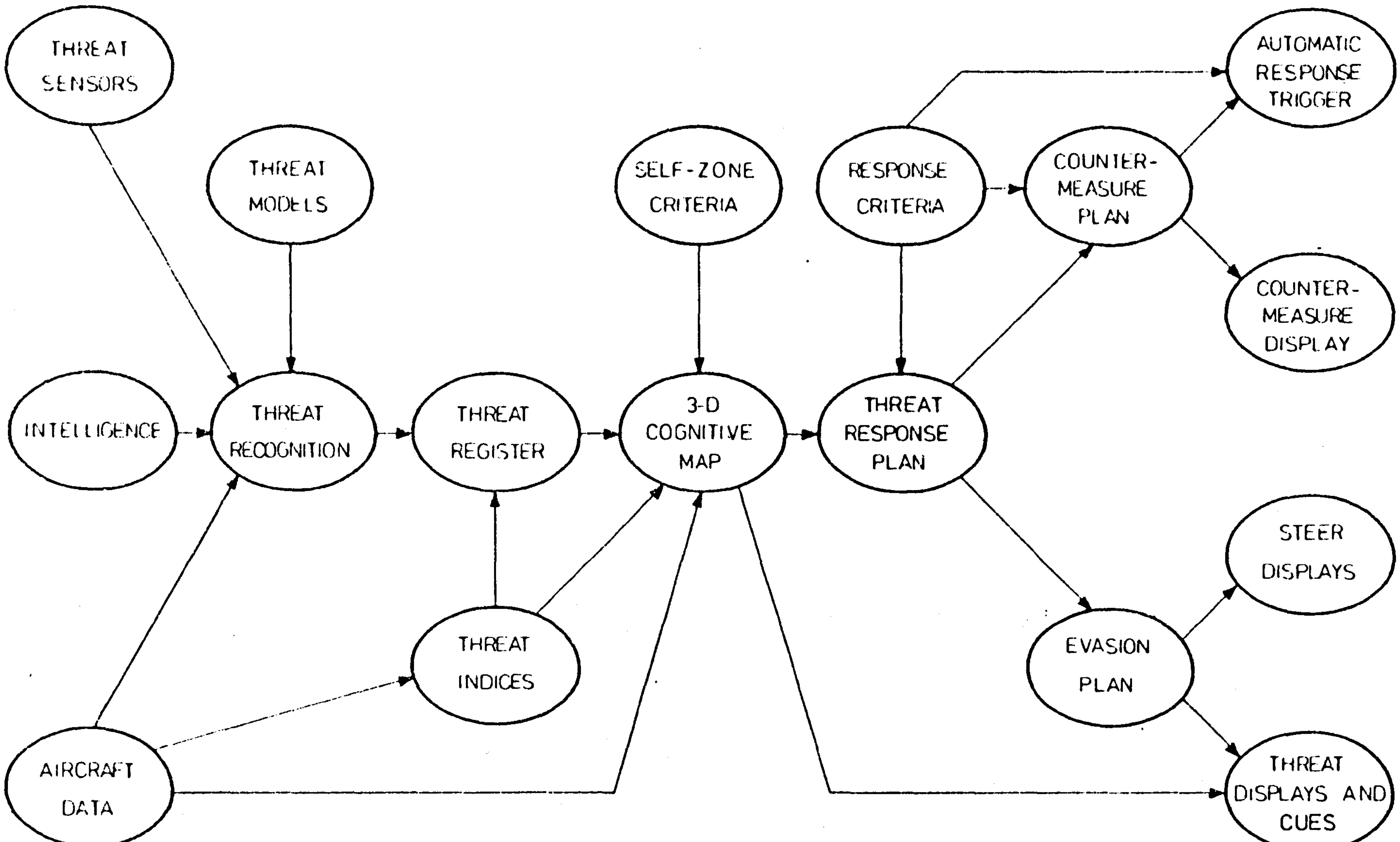


Fig. 5.8 THREAT ASSESSMENT SYSTEM - SECOND LEVEL

5. PROCESSING

5.8(a) LEARNING TO DETECT AND RECOGNISE

Recognition is required at long range as well as closer to a target; the techniques are quite different, so the system must operate in two different ways. Long-range target recognition lends itself to learning, while short-range multi-sensor correlation is more likely to be pre-programmed.

The threat assessment system will have (at least) two different classes of target model to which it will refer when trying to detect and recognise targets. The first model will be particular to its respective sensor, and will consist partly of a long-range representation of targets, as that sensor is expecting to see them. Figure 5.9 (a) shows a pictorial representation of radiation received from an object which has been scanned many times; each line represents a large number of superimposed scans at particular angles. For ease of viewing, successive traces are shifted horizontally so that the image slants across the picture. The height of each trace represents intensity. A store would hold this data in digital x, y and z components. Clearly, the picture could be represented in store with considerable precision - it was in fact derived from just such a digital representation. However, this data represents the radiation received from one aspect angle, against one background, at one time. Radiation from a tank say, differs with angle, background, surface finish and camouflage, time of day, relative angle of sun, weather, ambient temperature, internal energy sources, and so on. Figure 5.9 (b) shows successive recordings of radiation received from an apparently stable object against a relatively low-noise background. All this data could be superimposed to form a single line of the Fig. 5.9 (a), but information would be lost in the process; it is the variability of this object's radiation that characterises it.

A design compromise has therefore to be faced between compiling a vast store of detailed recordings of each of the likely targets, or of compressing the data and losing information in the process. Since the first option is impractical, the most sensible approach is to compress data carefully so as to lose predominantly irrelevant data.

For example, the data in Fig. 5.9 (a) might be stored by statistical means. Two orthogonal distributions could be formed, one along the slant and one across it, both passing through the peak. These would be stored as polynomials. Additionally, the variability represented in the lower figure could also be represented as a distribution, in this case uni-dimensional, and similarly stored, together with radiation polarisation where useful. The process must still be repeated at various aspect angles and under a variety of conditions, but the use of distribution theory and polynomial representation would contain the size of storage and the speed of access needed, and would allow the system to provide quantitative statements of confidence in its recognition of targets.

The process of teaching the sensor to recognise targets can now be considered. The sensor's processing system would be provided with a learning mode, which would be run initially against scale models at various aspect angles, against various backgrounds, etc. This learning process would result in a primitive set of polynomial distributions which would then be refined in the air by flying the sensor over real targets, initially using a helicopter but later using a fixed-wing aircraft. In operational use, this second learning mode would enable pilots to mark and record 'new' targets for intelligence analysis and future recognition both by their own and other aircraft.

So far, target recognition has been considered as a long-range/low resolution task: better performance might be expected at closer range/ higher resoltuion. In threat assessment terms this is questionable, since recognition at short range might be too late. Ideally, threats should be recognised at sufficient range beyond their lethality zone to allow evasion. Typical ground-based gun and missile ranges are given in Section 1, from which it can be seen that lethality out to 3 to 6 km can be expected at low altitude; these figures could be doubled to allow sensible evasion margins. For a human operator to recognise the image of a tank at, say, 9 km, would require an imager resolution of about 50 μrad which, at FLIR wavelengths, and allowing for sight-line jitter, would require apertures of over 40 cm - large by current standards. Thus a two-stage recognition process is envisaged, with closer recognition providing confirmation of the long range interpretation. (Long-range sensors would, of course, be correlated to provide best long-range recognition performance.)

The second class of target model will be pertinent to the closer range recognition process, since it correlates relatively fine-grain data from a number of sensors.

Thus a tank might be recognised by its aspect ratios (length: breadth: height) derived from a thermal imager; by the presence of tracks observed by an active radar with signal processing tuned to detect them; by the detection of a single barrel-like protrusion derived from a millimetre wave sensor; and so on. These target parameters would then be matched against a series of target models to select the best fit. This correlation model of the tank would be derived from intelligence in the first instance, and by observation of similar tanks or mock-ups in the field. A learning mode for this type of recognition can also be envisaged. The model's efficacy presupposes that the sensors can provide the appropriate input data, at least in part. It is a design choice as to the best location within the system of, for example, edge and contrast detectors for aspect and (hence) aspect-ratio determination, but generally sensors will be tuned to detect particular characteristics of targets, while the process of correlation between sensors will be centralised.

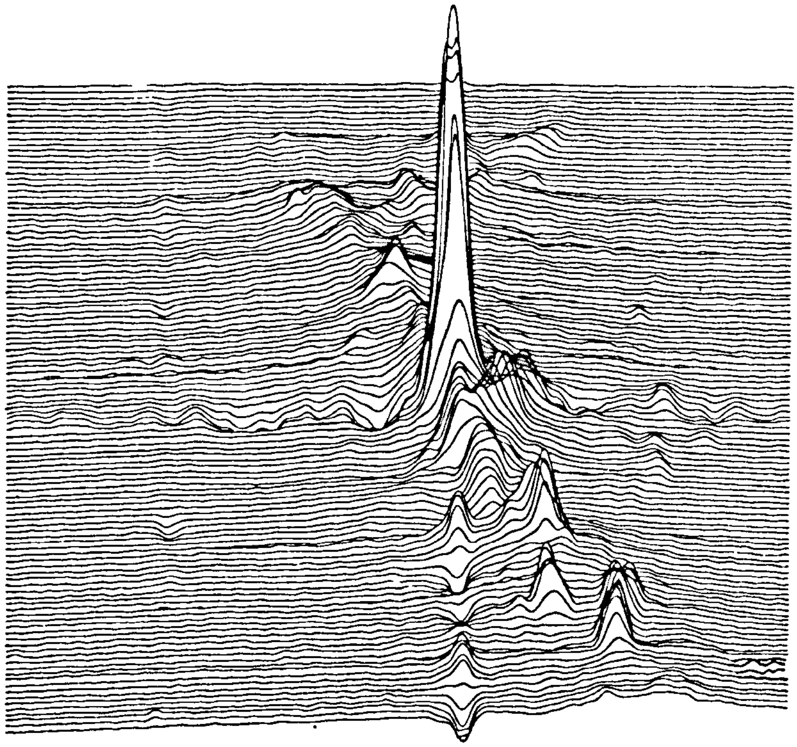


Fig. 5.9(a)

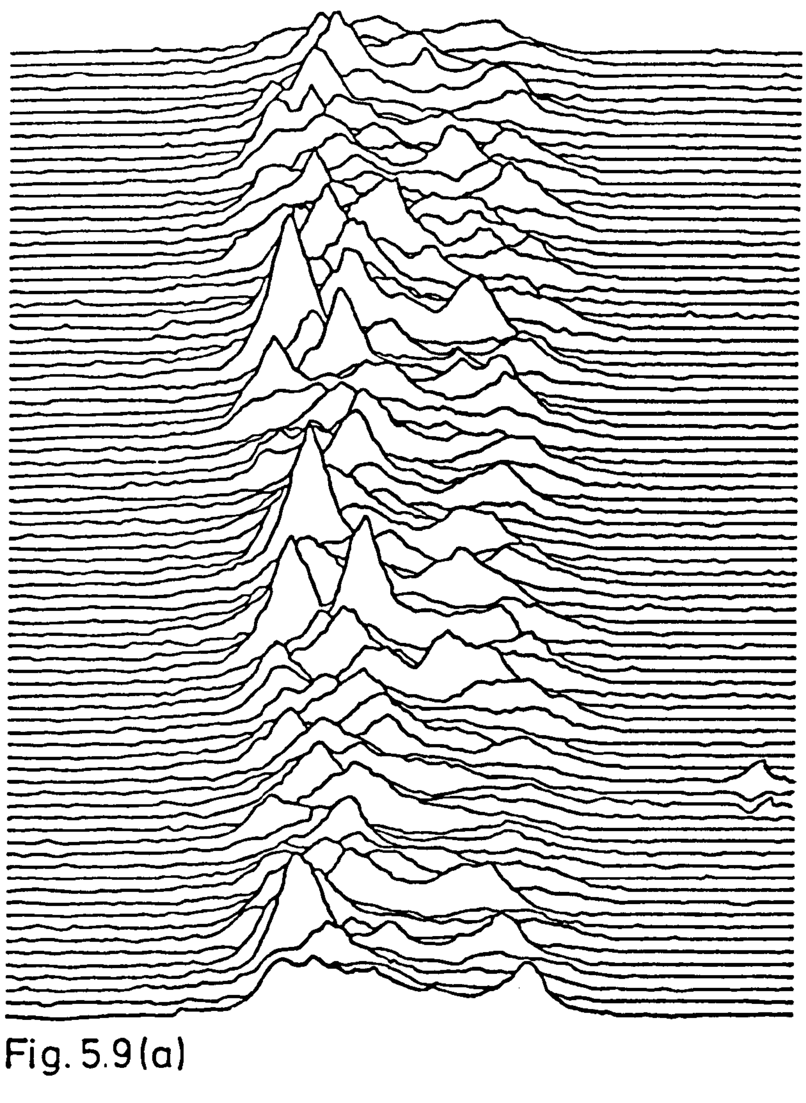


Fig. 5.9(a)

5 PROCESSING

5.8(b) TARGET RECOGNITION TECHNIQUES

Target recognition will be essentially a stochastic process involving correlation between observed data for a potential target and stored data for several target classes. Data can be manipulated and tested to give as high a degree of confidence in the recognition process as time (and the designer) might consider necessary. A simple example is provided which mimics human performance, with level of confidence or uncertainty being displayed as appropriate.

Target recognition consists of comparing observed target parameters with some stored models or memory of targets and finding an acceptable level of correlation between observation and model.

The process can be carried out on a single scan basis but is more likely to give sensible results using multi-scan comparison.

Target parameters can be examined in two ways: questions can be formed to which the answer is YES or NO (does the target have a . . . ?); or quantitative queries can be presented (what length-to-width ratio does the target have?). The second group can be rephrased as YES/NO questions and this type of question can be used to form a 'tree' or network of questions leading to the correct solution.

Unfortunately, perfect correlation between model and observation is unlikely to be achieved and some form of stochastic scoring system therefore becomes necessary in order to provide confidence in the recognition performance. Many such systems can be devised; the following example is original, but only one of very many. Its principal merits are simplicity, high computational speed, and a human-like response when choice is either obvious or uncertain.

Table 5-1 shows five elementary models of likely ground targets; Jeep, Lorry, Armoured Personnel Carrier (APC), Tank and 4-missile SAM battery. Not all of the discernible characteristics are equally useful in classifying targets. For instance, a single barrel is a more valuable clue to recognising a gun than, say, the position of a hotspot on a vehicle which could be cold. Weightings have therefore been attributed to the various parameters according to some *a priori* understanding of the 'uniqueness' of the recognition parameter.

The four columns A, B, C and D represent sample observed parameters; in no case is there complete agreement between the observations and a model vehicle.

PARAMETER		MODEL VEHICLE									
		JEEP	LORRY	APC	TANK	SAM	WEIGHT	OBSERVATIONS A	B	C	D
BARREL-LIKE PROTRUSIONS	1				✓		80	✓			
	4					✓	100		✓		
TRACKS				✓	✓	✓	65	✓	✓		✓
VENTS	REAR		✓				40			✓	
	SIDE L	✓				✓	25				
	SIDE R			✓	✓		25				✓
ASPECT RATIOS	4/3	✓				✓	30	✓			
	5/3		✓	✓			30		✓	✓	
	6/4				✓		50				
HOT SPOTS	REAR		✓	✓	✓	✓	15	✓			
	SIDE L	✓					30			✓	
	SIDE R				✓		30	✓			
	FRONT	✓	✓			✓	25				✓

COMMERCIAL-IN-CONFIDENCE

The correlation coefficient (r) between each observed set (A, B, C and D) and the models which the set might represent has been produced using the relationship:

$$r = \frac{1}{N} \sum \left(\frac{x - \bar{x}}{\sigma_x} \right) \left(\frac{y - \bar{y}}{\sigma_y} \right)$$

where:

$\bar{x}$ and $\bar{y}$ are the respective means of observed and corresponding model values

and:

σ_x and σ_y are the standard deviations of all the x and y values.

For example, observation A scores x = +80, and y = -80 (inverse correlation) for Jeep, Lorry, APC, and SAM, but x = +80 and y = +80 for Tank (i.e. positive correlation).

The results of this process are as follows:

			Order
A	r = +0.806 -0.186 -0.318 -0.806 -0.956	Tank APC SAM Jeep Lorry	1 2 3 4 5
B	r = +0.966 -0.423 -0.748 -0.966 -1.000	SAM Tank APC Lorry Jeep	1 2 3 4 5
C	r = +0.610 -0.610 -1.000	Lorry Jeep/APC Tank/SAM	1 2/3 4/5
D	r = +0.832 -0.832	APC/Tank/SAM Lorry/Jeep	1/2/3 4/5

On the basis of a single scan, it could be concluded that:

- A is clearly a tank
- B is clearly a SAM
- C is probably a lorry, but there is some doubt
- D is neither a lorry nor a jeep, but is likely to be a tracked vehicle.

However, 'better' information will accrue by associating scan-to-scan results. Suppose that six successive scans show the following 'orders of likelihood'.

Scan No.	Jeep	Lorry	APC	Tank	SAM	
1	5	4	2	1	3	
2	5	4	1	2	3	
3	5	4	3	2	1	
4	5	3	2	1	4	
5	5	3	4	2	1	
6	5	4	3	1	2	
Totals	30	22	15	9	14	= 90

It could be reasonably concluded that the column with the smallest total indicates the target - after all, a Tank was selected the first three times and the second three times. Some statistical criteria can be applied to test this 'obvious' result. The Coefficient of Concordance is given by:

$$W = \frac{12S}{m^2(n^3-n)} \quad (0 \leqslant W \leqslant 1)$$

where:

m is the number of scans

n is the number of models (Jeep, Lorry, etc.)

S is the sum of the squares of the differences between observed and expected rank totals (18 in this case).

whence:

$W = 0.739 \qquad O \leqslant W \leqslant 1$

which suggests that 'Tank' is a fair assumption on the data available. However, SAM was chosen twice and APC once, so what confidence can there be that it was not all random chance? Using SNEDECOR's F Test, it can be shown that the data could quite easily have been chance; more scans are required before 'Tank' is confirmed or replaced as a 'confident recognition'.

Thus a suite of tests can be established which can be used to process the data to give high levels of confidence that a recognition is possible, or that an object is not one of a predetermined set of known targets. Moreover, the system can be made adaptive, as the last paragraph implied by deferring judgement until sufficient data has built up a satisfactory degree of assurance.

5 PROCESSING

5.8(c) TARGET RECOGNITION BY BEHAVIOUR AND ASSOCIATION

Targets have recognisable performance parameters and they tend to form groups. Both factors may assist in recognition and identification.

Contemporary recognition systems often base their analyses exclusively on the radiation received from a target. Various combinations and distributions of electromagnetic radiation will characterise targets up to a point, but other approaches may contribute also.

Human recognition is largely concerned with observing an object in relation to its background or in relation to other objects: a moving object is very much easier to detect than a stationary one; an aircraft's range can be estimated by relating its angular subtense to that of the trees near it, and so on. Indeed, once denied such clues, human performance falls off rapidly.

It is not difficult to build similar capabilities into a threat assessment system once their potential is appreciated.

An imager may detect movement relative to background either by locking to an object and observing ground motion, or vice versa. A digitised version of the image may be analysed pixel by pixel in a relatively straightforward but lengthy process so as to detect frame-to-frame movement. This technique also enables lock-follow in passive sensors.

Once movement is detected, velocity can be estimated and the object may be compared with simple target performance models (as a precursor to recognition) to eliminate impossible solutions.

More subtle perhaps is the use of association. It is known from intelligence sources that certain types of target associate with other vehicles. A SAM-battery has an associated array of vehicles for workshops, power supplies, accommodation and cooking. A possible SAM detection may be more credible if such association is observed. Association can also be used to estimate range, which is generally difficult over uneven terrain, by comparing the dimensions of known objects with those of a nearby recognised target.

Negative associations also exist; but are generally less useful. Lorries and jeeps do not float and so cannot be superimposed on water. Some tanks, paradoxically, can float and their so-doing may be used for recognition by association.

5 PROCESSING

5.9 THE OUTLINE PROCESSING SYSTEM

The processing task will be large by current standards. Learning modes and operational reprogramming will need to be built in. High integrity performance will be required in some areas. While the task appears feasible, there are many grey areas where more work will be essential to arrive at an operational system.

Task Groups

The third level DSD opposite gives some impression of one processing organisation; other arrangements could be conceived. There are six major processing groups:

- FLIR image processing
- Radar image processing
- Multi-sensor recognition/correlation
- Threat ranking
- Response planning
- Hazard detection.

(Radar, laser and missile warning sensors will also have their own processors.) There are significant tasks associated with transferring image data from radar and FLIR to the multi-sensor recognition/correlation process; such data could take the form of simple video, to be digitised in the threat processor, or could be digitised at source. Scene correlation requires that the data from each sensor be allocated to the 'correct' store location *vis-a-vis* the other sensor; in other words, the two scenes have to be registered, each to the other.

These data then have to be refreshed, superimposed, discarded and manipulated so that comparisons can be made between observed patterns and stored models. This will be a time-consuming process, and research work suggests that the process of highlighting patterns and matching them with models is unlikely to be achieved in a typical frame time (1/25 s for a FLIR), using contemporary serial airborne computers. A combined approach appears viable, employing array processors on the one hand, to speed computation, and operating on every third or fourth frame only, for example. Recognition times of several seconds can be anticipated, since the need to compare between successive samples will be necessary to establish confidence. This essential uncertainty in recognition time must be allowed for in the real-time configuration.

The FLIR and radar are expected to have their own signal processing capability, not only because it will provide useful long-range and reversionary capability, but also because the specialist manufacturers will be best able to design to their particular requirements; they will also have a strong selling motivation. For these reasons, both sensors are expected to provide some limited stand-alone performance, with imagery and target/threat recognition and overlay outputs. Transfer of these latter parameters into the multi-sensor processor has to be achieved; one technique would be to interlace them with the scene data (either digitised or video).

Threat ranking should be straightforward, at least in processing terms, once the ranking criteria are established. It may well be best to leave some ranking criteria to crew selection. Others will be derived from intelligence sources concerning threat lethalities and engagement zones; these may need rapid revision in the light of fresh experience, so new data must be readily assimilated by the ranking processor when in operational use.

Similarly, threat models will require rapid update and the ability to accept new categories; learning modes have been discussed which will dictate the need for comprehensive automatic statistical analysis, reduction and storage of large quantities of signal data.

Response planning will require high integrity and reliability, to avoid nuisance trips. Alternatively, speed of response must be the essence of its performance. The countermeasure plan is the simpler part, since its characteristics can be established and verified with little risk. The evasion plan evidently requires careful design and testing, coupled with an understanding of its limitations *vis-a-vis* local terrain. Terrain avoidance data would clearly be valuable in this respect; a millimetre radar might provide terrain avoidance cable detection, and mapping facilities.

The manoeuvre monitor is envisaged as an aid in combat or evasion, to warn the pilot that he will hit ground unless he changes his current manoeuvre. Its principles are established.

Areas of Uncertainty

There are many grey areas where further work is required to establish techniques and validate performance. A reasonable assessment would be that the processing task is large, but seems achievable using innovative as well as contemporary methods and designs. Included among the grey areas are:

- Model-characteristic analysis and storage, with respect to the various combinations of aspect, incident, radiation, internal radiation, background terrain, obscuration, etc. which will be met in practice.
- Pattern recognition algorithms with acceptable run times and sufficient power to differentiate between similar - though vitally different - targets.

- Statistical algorithms which reconcile recognition confidence with processing time, and which perform in a way which non-specialists can readily appreciate.
- Adaptive pattern tracking algorithms.

With these and other problems unresolved, it would be imprudent to speculate concerning numbers for computing storage, cycle times etc. In practice it will be possible to trade these factors against performance, confidence and time delays, so there seems little doubt that such systems will work; how well, will require demonstrator tests to establish.

Preliminary Processing Concepts

Some preliminary figures can be estimated - see Topics 5.1 and 5.2. Frame stores would be required, both in the sensors and in the multi-sensor processor (MSP). That in the radar might form part of a scan converter. The FLIR design is difficult to predict, owing to the uncertain timescales for the introduction of high-density focal plan imaging matrices. However, the FLIR will require a frame store of some kind in order to search for thermal patterns. The sizes of such frame stores may be limited by packing the data using the methods already described, such as spline approximation. Typical frame store dimensions might then be:

- FLIR - 1 Mbit
- Ground search radar - 1.2 Mbit

If the sensors compress their data to limit storage requirements, this compressed version of the images may be conveniently transferred to the MSP for correlation. The MSP frame store is seen as a 3-dimensional (3-D) storage block: X and Y co-ordinates would represent spatial angles from the aircraft, while the Z co-ordinate would contain FLIR and radar data appropriate to each X-Y co-ordinate; the stores would inevitably be larger than those in the imaging sensors.

Threat extraction would then consist of searching through this store for patterns which compare in some degree with stored, multi-sensor models. The time taken for this exercise may be reduced if the search were concentrated around particular co-ordinates - hence 'co-ordinates of interest' would be transferred from single sensors to range detection and tracking in the MSP. The processes are nonetheless seen as lengthy.

Threat recognition/correlation would combine the information from all sources to achieve the best recognition confidence available within the system. There would therefore be inputs not only from the threat extraction process but also from laser warning, radar warning, etc. These various sources would be compared and an assessment made of their combined viewpoint, which implies weightings and stochastic analyses. The sources also have widely differing resolutions; radar warning might give relatively crude spatial resolution, but precise recognition data compared, say, with FLIR. The processing tasks associated with this work would use established principles and are not foreseen as a major hurdle once the various weightings and combinatory logic are established during development.

Threat ranking would consist of a straightforward computing task of ordering the recognised threats according to their relative significance, where significance is a combined measure of imminence and potential lethality, taking account of current track, speed and height. Hysteresis would need to be built-in to the procedure to reduce oscillation between threats of similar and changing significance. The 3-D cognitive map would present a less conventional problem however. This map is seen conceptually as a 3-D store representing the space in front of the aircraft. Ground will fill the bottom planes. Drawn into the map would be: the most significant threats taken from the threat rank register; together with their engagement volumes and lethality data; two 'self' zones representing own-aircraft track and clearance volumes; and planned target co-ordinates. The map would be 'sliced' along the aircraft velocity vector to produce a 2-D display to the pilot, showing threat areas varying with attitude, height, speed, etc.

The threat response plan would be designed to reduce the threat by evasion, by countermeasure, or both. The evasion plan would consist of algorithms designed to construct a minimum risk route between the threats, towards the target. Such algorithms can be complex (see Topic 3.4(a)) and may incur protracted run times, so that this task is seen as likely to require a dedicated processor and store. However, the techniques are well established, so that no new ground need be broken. The outputs from the evasion plan would consist of an overlay showing the route plan on the 2-D map and a steer display, most probably head-up. The degree of steer signal authority would require careful study. At this stage it is considered as simple left/right advice, but pitch and speed signals may be necessary too.

The evasion plan would provide a suitable input point for hazard avoidance signals which may require to override other steer advice. Hazards may be detected by sensors - cable warning or radar principally - or by a manoeuvre monitor. The resulting hazard warning may be a candidate for the central warning system.

The countermeasure plan would be designed to:

- Curtail electronic emissions from active sensors and communications which might attract anti-radiation missiles
- Provide advice on the firing of chaff, flares and on jammer initiation
- Trigger such active countermeasures automatically when enabled by the pilot and required by the situation.

These requirements suggest that there would be a need to establish priority within the processing complex, so that lengthy recognition processes did not impede timely countermeasure response.

The conceptual processing system therefore emerges as an arrangement of inter-related activities, some of which are predictable in content and duration whilst others will be stochastic and unpredictable. The latter must not be allowed to impede the former, and so the conceptual design tends towards a multi-processing organisation, perhaps with micro-processors, programs and data being functionally divided rather along the lines suggested by the conceptual diagram of this topic. Image recognition will require very fast processing performance.

Work associated with many of the processes described above is in the research and development phases at present. Much more work will need to be done before a proven system emerges.

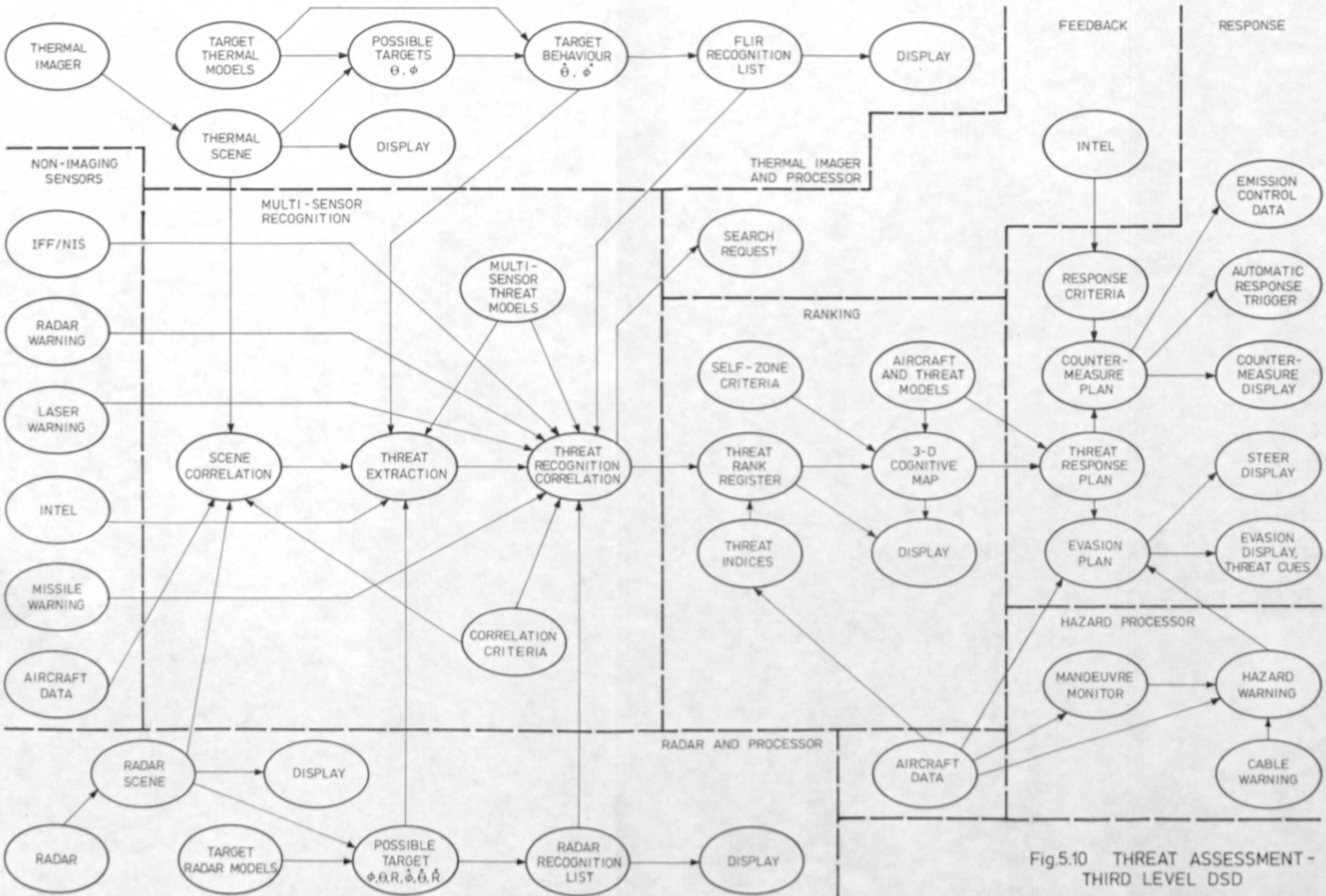


Fig.5.10 THREAT ASSESSMENT - THIRD LEVEL DSD

6 THREAT WARNING SYSTEM - OUTLINE DESIGN

6.1 INTERFACES AND COMMUNICATIONS

The design keynote is confidence in target/threat assessment; this implies retention of high sensor entropies, which in turn results in the need for high data rates.

Figure 6.1 and the Key to Fig. 6.1 show the principal interfaces of a typical design, before physical grouping and partitioning; they should be viewed with Fig. 5.10.

Clearly, there is a close functional relationship between the radar, its processor and scan converter, the FLIR(s) and its (their) processor(s), the target/threat processor and the electronic displays and controls. A major design option concerns the transfer of thermal and radar scenes to the target/threat MSP. Figure 5.10 showed the scenes being transferred for centralised correlation. There are three options:

a) Dispense with scene correlation, and simply correlate data from the FLIR and radar processors, which will highlight co-ordinates where targets are recognised by single sensors.

b) Transfer digitised frames of data - practicable if each sensor has its own recognition performance, requiring integral frame stores and data compression.

c) Transfer the data as raw video - after scan conversion for the radar.

It will also be noticed that the interface chart, Fig. 6.1 shows one MSP for recognising both threats and targets; this is logical provided the recognition processes are similar, which seems likely.

However, a crossover in performance is anticipated; single sensor recognition is expected to operate more successfully at longer ranges, and with lower confidence levels, than multi-sensor recognition. Confidence is the key to the design implied at Fig. 5.10; by correlating unprocessed data it is hoped to retain the essential characteristics which will enable high-confidence, shorter-range target recognition. Using Option (a), much of the basic data is lost in single sensor processing, so that, at closer ranges, correlations may simply not occur at all.

For example, a camouflaged tank with a hot spot may not be recognised by radar and would not appear in the radar recognition list; in consequence there would be a total lack of radar input using Option (c), while the FLIR would continue to give some single sensor recognition performance.

For this reason of confidence, Option (a) as shown in Fig. 6.1 is unlikely to provide a good solution.

Option (b) seeks to take advantage of the digitisation and data compression which will be necessary in the radar and FLIR(s). If this pre-processed data were to be transferred directly to the MSP, that unit would not need to duplicate the exercise, which will present a significant task. Fig. 6.2 shows one way in which Option (b) might be mechanised, and Fig. 6.3 shows the equivalent interface diagram, which is evidently simpler than that of Option (a). The high-speed (capacity) bus would multiplex data from all the sensors on to one bus. Clearly, the bulk of data would be that provided by the radar and FLIR(s); other data would be inserted between frames passed from these sensors, and, if necessary, between scans.

One reason for considering a separate bus for target/threat data is the nature of the warning sensor outputs. This is sporadic in nature, with periods of long 'silence' being interspersed by bursts of rapid and important data. To mechanise sporadic data through the avionics bus would incur a heavy, inefficient overhead since the sensors would require frequent interrogation, largely giving nil returns.

Option (b) could also be mechanised using a lower capacity bus and two direct digital links between the radar and FLIR and the MSP for the transfer of image frames.

The potential disadvantages of Option (b) are:

- High capacity digital links will be required in the aircraft
- The digitisation and compression performed in radar and FLIR might be unsuited to the MSP tasks; reduced signal entropy and hence spatial resolution, might prejudice the MSP performance.

Hence a different mechanisation seems preferable; this is shown at Fig. 6.4. The high capacity links simply pass raw video, and the duplicated digitisation overhead is accepted. The MSP is used as a gateway between the target/threat and avionics buses, and the target/threat bus extracts the inefficient overheads caused by warning sensors from the avionics bus. The MSP could also act as its own bus controller. The associated, and relatively simple, interface diagram is at Fig. 6.5.

On present evidence the scheme represented by Fig. 6.4 is the preferred conceptual design. It will not require 'special' digital links; video could be passed via coaxial cable or optical fibre. The MSP need not be 'tuned' to the output formats of radar and FLIR. There is no unnecessary loss of signal entropy prior to image signals reaching the MSP. The use of a separate bus for target/threat data reduces avionics bus loading, as does the use of the MSP as a gateway.

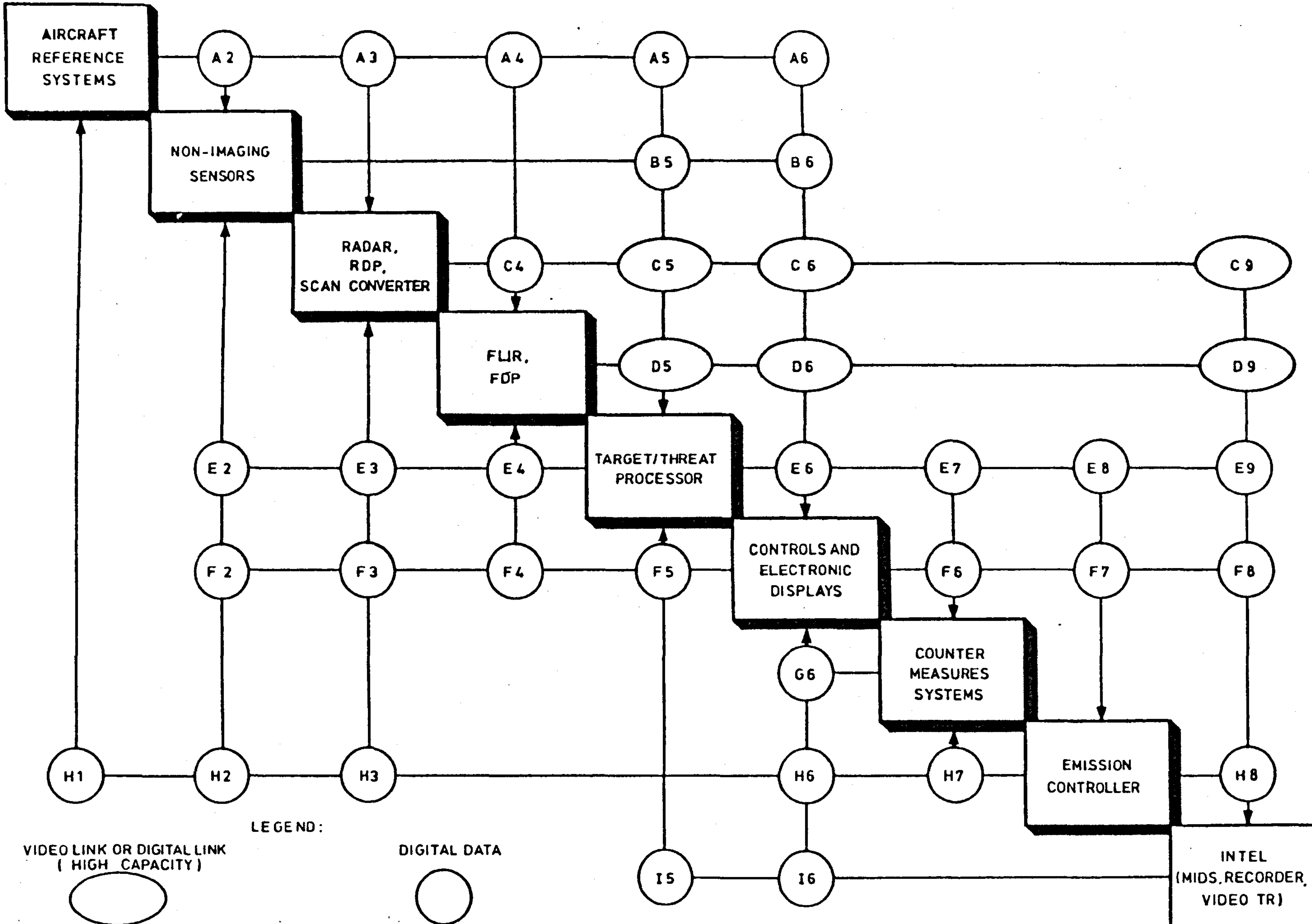


Fig 6.1 THREAT ASSESSMENT SYSTEM INTERFACES

KEY TO FIG. 6.1

A2	Altitude, heading, attitude, present position
A3	Altitude, heading, track, speeds, attitude, present position
A4	Altitude, heading, track, speeds, attitude, present position
A5	Altitude, heading, track, speeds, attitude, present position
A6	Altitude, heading, track, speeds, attitude, present position
B5	Laser, radar, missile, cable warning data
B6	Laser, radar, missile, cable warning data
C4	Synchronisation
C5	Video or digitised frame
C6	Video + plots and tracks + target threat recognition lists + synchronisation
C9	As C6 for recording
D5	Video or digitised frame
D6	Video + plots and tracks + target threat recognition lists
D9	As D6 for recording
E2	Mode Control (automatic)
E3	Mode Control, fire search request, pointing commands
E4	Mode Control, fire search request, pointing commands
E6	2-D Cognitive map display + target/threat recognition lists + steer signals
E7	Chaff, flare and jammer commands
E8	Emission suppress command
E9	Intel data for recording or transmission
F2	Mode control (manual)
F3	Manual radar controls - including mode, marker, pointing angle, scan pattern, sweep, etc.
F4	Manual FWR controls - including brightness, contrast, reverse contrast, pointing angle, magnification;
F5	Mode select, including learn/run, directed search, etc.
F6	Manual controls
F7 F8	Manual controls
H1	Suppression of rad alt, doppler etc

KEY TO FIG. 6.1 (Cont'd)

H2	Suppression of active cable warning etc.
H3	Radar transmission suppression
H6	EMCON indicator
H7	Active radiation suppression
H8	JTIDS/COMMS suppression
I5	Response rules, target/threat locations and parameters
I6	Intelligence text

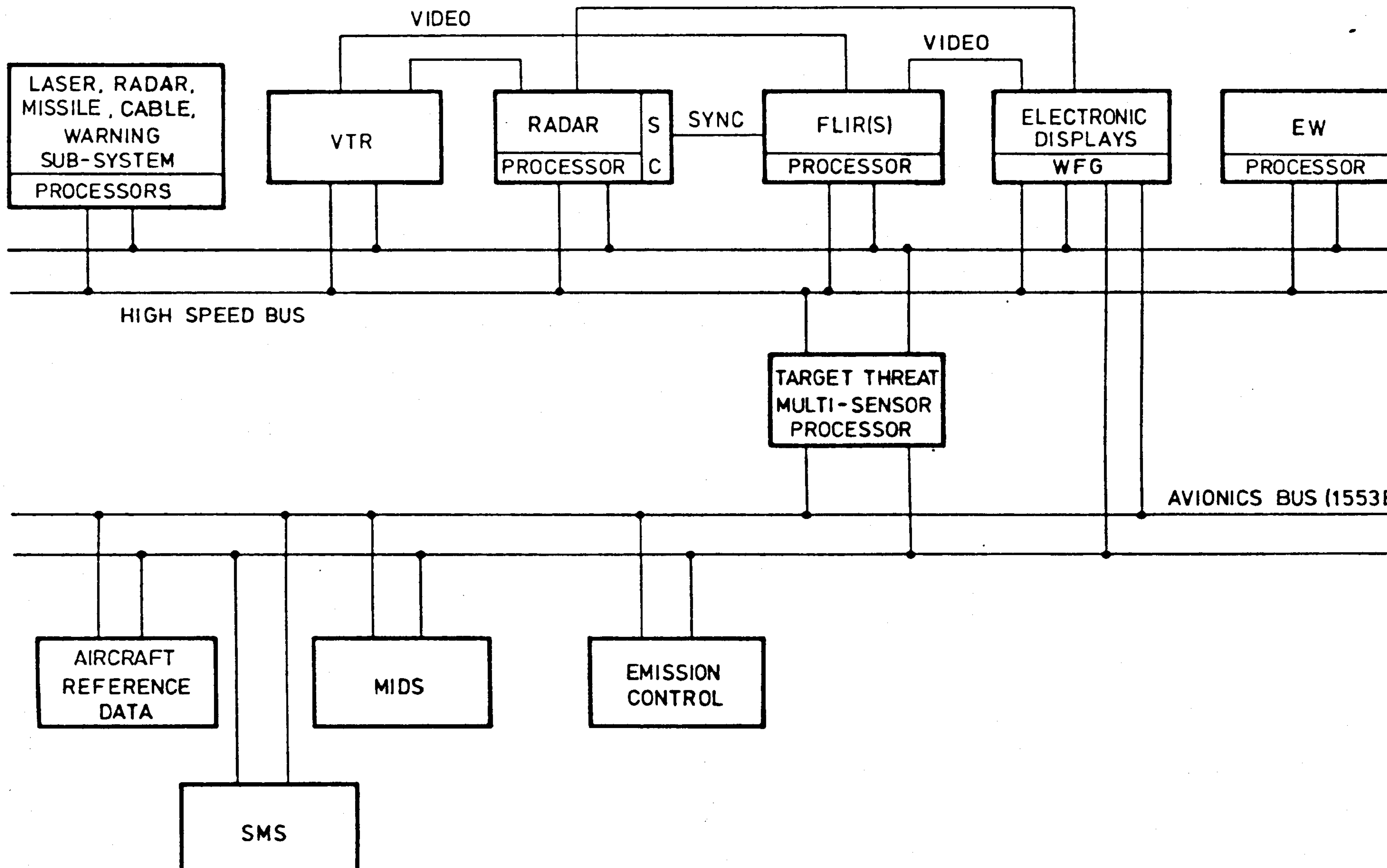


Fig.6.2 THREAT ASSESSMENT SYSTEM MECHANISATION-EXAMPLE

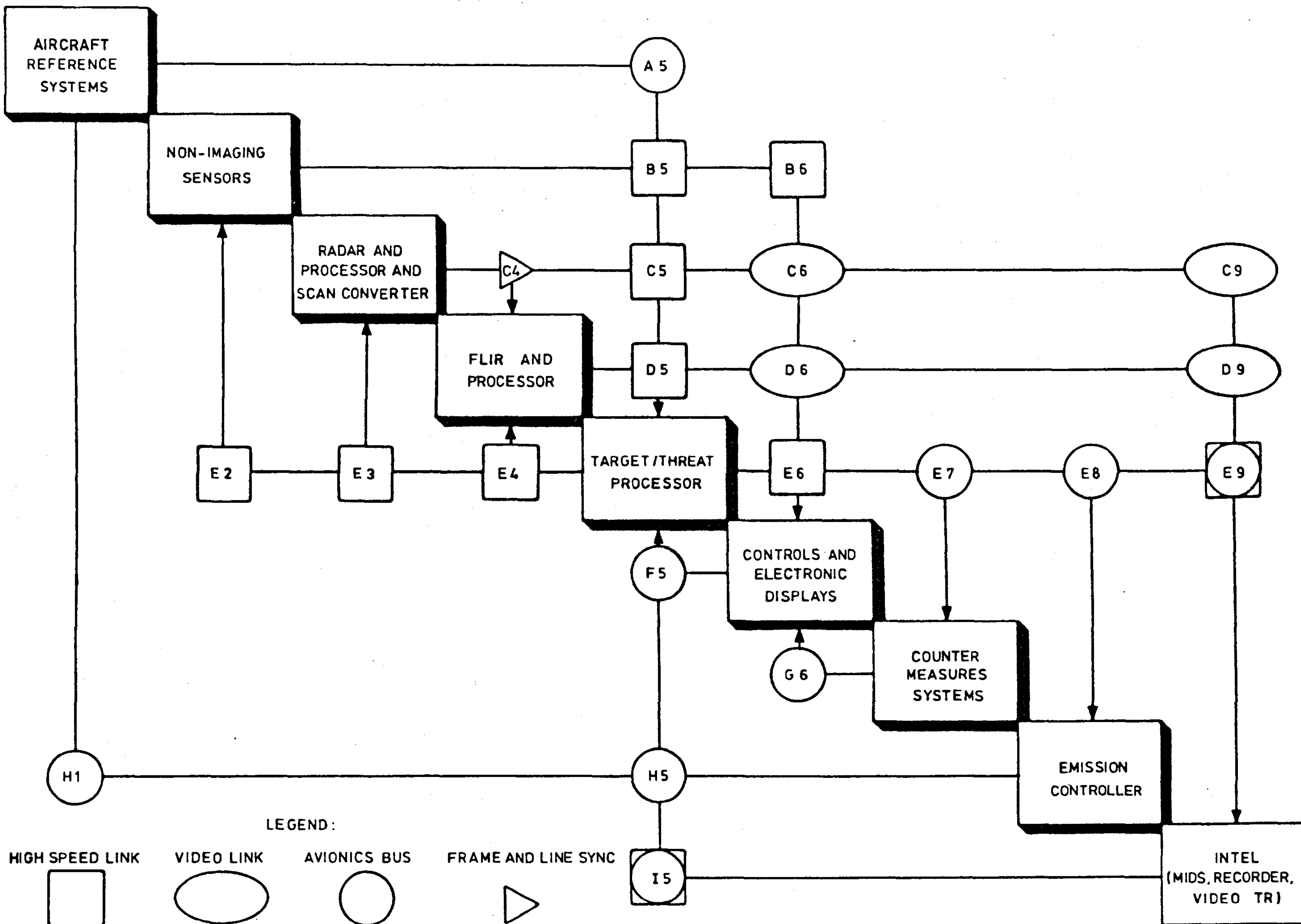
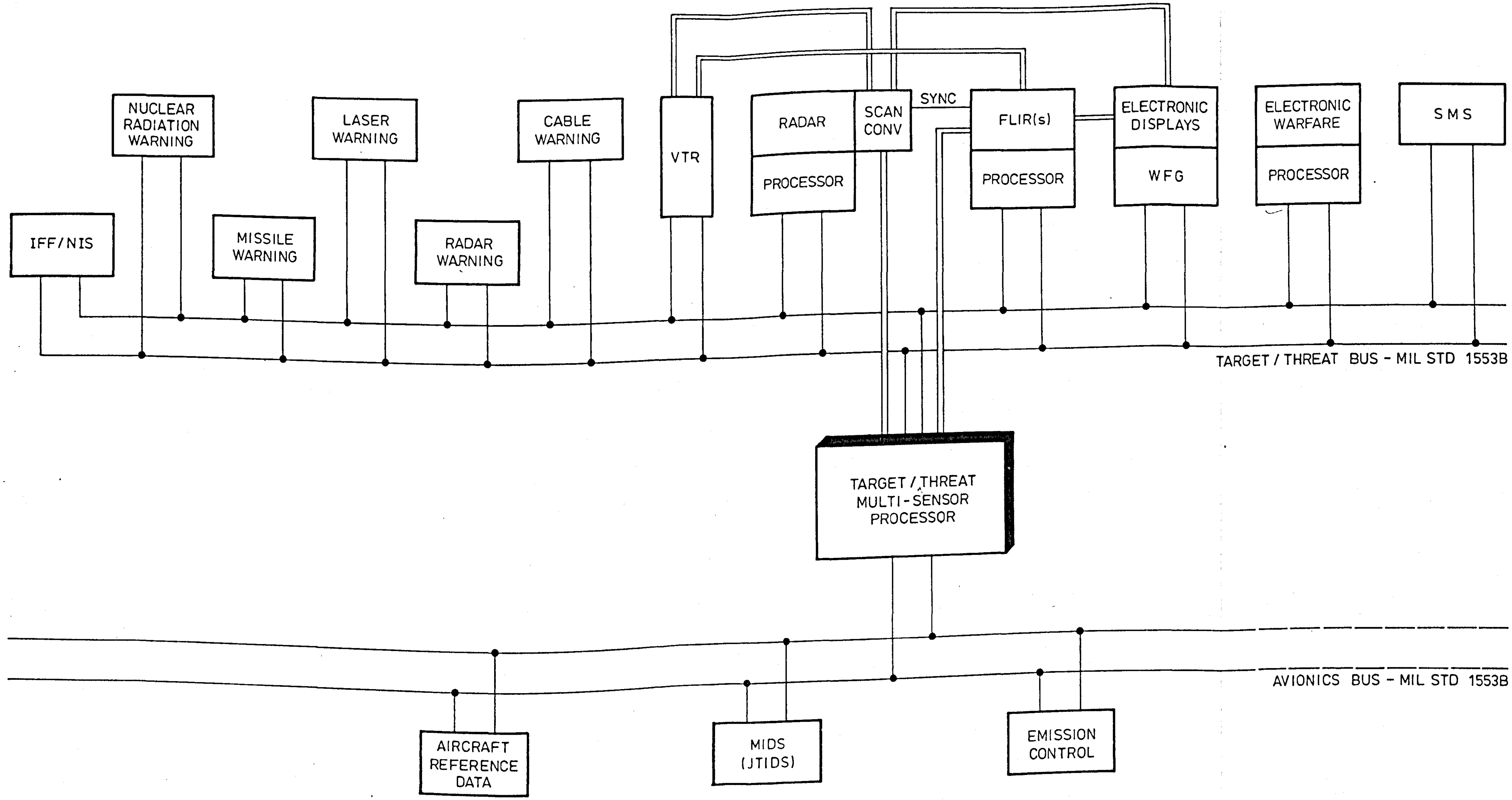


Fig.6.3 MECHANISATION EXAMPLE



KEY :-

VTR - VIDEO TAPE RECORDER
FLIR - FORWARD-LOOKING INFRA-RED
WFG - WAVEFORM GENERATOR
MIDS - MULTI-FUNCTION INFORMATION DISTRIBUTION SYSTEM
JTIDS - JOINT TACTICAL INFORMATION DISTRIBUTION SYSTEM
═══ - VIDEO LINK - COAX OR FIBRE

Fig.6.4 THREAT ASSESSMENT SYSTEM MECHANISATION - EXAMPLE

THE HUMAN ELEMENT IN C^3I

D.K. Hitchins

Racal Decca Systems Ltd

INTRODUCTION

The idea that command, control, communication and intelligence (C^3I) systems can be likened to a body is not new. Proponents of C^3 Counter Measures (notably Major General Doyle E Larson USAF until recently Commander, Electronic Security Command) often view enemy C^3I systems as having brains, nervous systems, muscles and of course vulnerable points which can be attacked to best effect.

C^3I systems bear the mark of their designers and their simian heritage in many apparent ways. For instance the bandwidths of our communications systems are generally comprised of channels designed to accommodate human aural or visual bandwidths.

In more subtle ways, however, we can learn from the biological and social evolution which has shaped our internal construction, our individual characteristics and our group behaviour patterns. Since C^3I systems are essentially supporting human interactions, such understanding would seem a promising way of deriving successful designs.

The aim of this paper is to suggest a new perspective and to offer a new mental model of C^3I systems which may enhance our approach to system design.

A C^3I DESIGNERS VIEW OF THE HUMAN STRUCTURE

Even at a cursory level, parallels can be seen between our developing C^3I designs and the way in which the human frame has evolved. Perhaps the most interesting aspect, conceptually, is the balance which has evolved between opposing pressures. For example, the body has three different structural techniques to enhance survivability with hardening of the central control and sensor processing (cranium), protection of the trunk communication system (spine) and of the respiration and energy transfer centres (lungs and heart) and multiple redundancy of purifying systems such as kidneys. However these essential features are not allowed to impede the essential mobility and flexibility of the design, since these two also favour survival. Such conceptual analogies can be useful in forming a modified mental model of C^3I system designs, which in nature seem invariably to compromise between opposing objectives. Some further examples will illustrate the point.

Speed of Response versus Central Control

From touch sensor through the brain to motor takes between 0.2 and 0.4 seconds for a fit adult. Since this is inadequately slow in emergencies, the spinal reflex arc links sensor to motor directly, without prior reference to the brain which is informed in parallel. Figure 1 shows both the spinal reflex arc (Robert Barrass (1)) and the command and control analogy. For the human, stepping on a thorn results in a pre-programmed fast response to limit damage with the brain being informed in slower time. Similarly the surface-to-air sensor and missile may need to respond too rapidly to allow time for direct higher control.

Groups of instructions may apparently be learnt within the human central nervous system so that a pianist or typist for example, may execute a string of actions as a single command and thus overcome the delays which would otherwise be incurred. In C^3I terms this technique would be similar to Standard Operating Procedures (SOPs). An alert mechanism (adrenalin) heightens responsiveness to danger and prepares the human for fight or flight in a similar way to Alerting and Early Warning in C^3 systems.

Integration versus Sensor Optimization

Body subsystems are so well integrated that it is not sensible in many cases to define an interface. The spinal cord and the brain form a single processing and communication system while eyes and ears may be viewed more as extensions of the brain rather than separate features. Their colocation near the brain at the highest point for best visibility and simplest sensor fusion reconciles the dictates of wide base for triangulation, day/night performance and shortest signal path for rapid response.

Processing

The brain comprises two mirror image halves, each capable of independent action but coordinated through a joining nerve "cable", the corpus callosum. The information content of the brain has been estimated at some 10^{14} bits (Carl Sagan (2)) and conscious thought, contained in the cerebral cortex, is divided between the two hemispheres. The right hemisphere is mainly responsible for pattern recognition, intuition and creative insight, while the left hemisphere is the seat of reason, analysis, critical thinking and understanding of speech. Together these two halves provide the means for generating and testing ideas, generally without an overlap in function which could lead to disagreements and unresolvable internal conflicts.

Training and Development

The system when initially delivered has basic autonomic and reflex responses, but is so designed that it may develop its full capacities after an in-service training period of some ten or more years. The design would lack flexibility and adaptability if it were delivered with too much pre-programming as is the case with lower life-forms. Strangely we expect our C^3I systems to be effective from the beginning of their lives and are disappointed when unreasonable expectations are not realised.

LEARNING LESSONS FROM THE HUMAN DESIGN

The analogy between human body is so consistent that it is tempting to carry it further. Indeed the activity can become so rewarding that it ceases to be the comparison of analogies. Instead when seeking new ways to tackle a C^3I design task, it can be fruitful to examine the principles employed in the human design which has, after all, evolved over tens of millions of years. Some examples will indicate the concept.

Pain Gate Theory and Communications Minimization

It has long been known that response to pain in humans is very much related to situation. For example, swimmers have suffered such severe injuries as the loss of legs due to shark bites without feeling pain at the time.

Patrick Wall, Professor of Anatomy at University College, London, has recently developed the Pain Gate Theory which provides a consistent explanation of this phenomenon; Figure 2 refers. Pain signals from, say, the foot are passed to the brain via the pain gate only if:

- o The brain is prepared to receive them
- o Local influences permit.

Thus a runner may ignore, or genuinely not be conscious of, blistering of the feet when he is in competition, and indeed can run virtually "on autopilot" provided his rythm remains undisturbed. Meanwhile the brain is concentrating its not unlimited capacity on calculating strategy and tactics, and on conscious self-conditioning.

Is there a similar mode of operation in military C^3I systems? To a less sophisticated degree the military tactic of minimizing communications is a similar concept. The minimization instruction is employed under stress to reduce communications to those essential for operations; less urgent background communications are either deferred or sent by slower means. However, unlike the Central Nervous System, the C^3I processing system is not generally reconfigured to concentrate power on the urgent problem in hand. Perhaps this is behind the apparent ability of humans to respond to crisis situations without collapse due to total overload.

Figure 2 also shows a diagram which mimics the pain gate. The system is multi purpose, but is shown only in its Emergency Processing Mode. A multi-task processor is shown, with groups of Input/Output (I/O) ports, each with dedicated functions. An emergency signal from the Sensor in Area E reaches the multi-task processor only if:

- Local Emergency Control deems the signal to have exceeded a threshold and closes switch E1

- The multi-task processor has prepared itself to process any signals from Zone E and has set switch E2 accordingly

This decision, whether or not to look towards Zone E, is part of an adaptive polling process, in which the overall situation as determined by the Situation Assessor adapts the polling schedule.

Emergency Alarms are hard-wired to one of the Multi Task Processor output ports, and the initiation of an alarm is fed back to the situation assessor. An Executive Controller is shown as a means of scheduling, interrupting and reconfiguring the Multi-Task Processor; in practice the Executive Controller could form part of the Multi-Task Processor.

Figure 2 does not represent a typical contemporary C^3I design. Are there any lessons to be learnt? Perhaps: for example, the bundle of cables from the emergency sensors is unlikely to be affected by cross-talk, since communications are not operative unless an emergency both exists and is of interest. In a networked system the required communications capacity would be reduced. Of significance in large C^3I systems, the central I/O processing task would be greatly reduced, leaving Multi-Task Processors more time to concentrate on relevant activities. Perhaps most important of all, the design concept prevents the situation assessor from being distracted by information which is not relevant to the immediate problem of survival.

Sensor Design

The human eye is highly evolved and has a variety of interesting characteristics, including a curious susceptibility to impulsive jamming. The same optics in the eye are shared by a variety of sensor elements (rods, cones, etc) providing:

- High resolution over the limited foveal area

- Peripheral flicker vision alerting (PFVA) to detect and call for concentrated attention to movement outside the high resolution area. (PFVA is of course the cause of some limitations in the design of multiple raster-scan displays where looking at one display allows the flicker from adjacent displays to be evident)

- Separate colour sensing

- Separate low and high ambient light sensors.

Additionally, the eye-brain combination provides:

- Analogue to Digital conversion at the retina

- Specialized detection of vertical, and horizontal edges

- The ability to link up features and, in some little understood way, recognise patterns.

- Distance measurement by binocular vision textural changes, focal distance, perspective and rates of change of sight lines to various objects.

The size of the eye is also relatively small, which reduces the probability of accidental damage: operation at, say, infra red frequencies would increase eye area and de-optimize the general purpose day/night design. Again, perhaps the most interesting design features are the optimization between opposing requirements, and the degree to which eye/brain performance develops after birth and is tuned to the environment rather than fully developed in utero.

Complementary Communications

The human body employs two principal communication systems: neural control by the transmission of nerve pulses and hormonal or chemical control. These two control mechanisms are quite different both in their relative speeds and in their functions; but it seems to be a standard philosophy of human system design that balance (homeostasis) is achieved by opposing forces or effects. Neural control travels at about $100ms^{-1}$ and is used for real time control mainly of skeletal muscles. Hormonal Control is carried via the blood and is therefore available in every tissue and to every cell. Action depends upon target cells responding to the chemical agents while non-targets are not affected. Hormonal control can operate in the time it takes for blood to circulate the body - some 30s - or may operate over much longer periods stretching to years (1). Hormonal control regulates alert responses, combats invasion by foreign substances, regulates growth and development, etc.

The C^3I equivalent of the neural control system is evident in near-realtime C^3I design although the human control system is very much more sophisticated and employs non-linear feedback "mechanisms". Hormonal control has some counterparts as in broadcast alerts and damage repair organizations, but generally the elegance of hormonal control has yet to be matched in C^3I designs. However, if we were to push the analogy a little, and note that blood is, inter alia, the carrier of energy to the muscles and organs then we could suggest that mains power supplies might form a useful non-secure communications bearer at least within an HQ since almost everyone has connections to the supplies. Street-lighting has for many years been switched on and off by signalling superimposed on the power lines, so the concept is far from alien.

C^3I AND CYBERNETICS

Comparisons between human body design and C^3I systems can be continued almost idenfinitely and to advantage, but it is equally revealing to consider C^3I systems or even fleets and armies as organisms with communication equipments transferring information between groups of users/operators.

The human element in modern C^3I systems is most apparent in the command function. Analysis of C^3I must, therefore, start at the commander, his particular purpose, environment and constraints. The total C^3I system can be viewed as an extension of the commander in much the same way as the remote manipulating equipment used for handling radioactive isotopes in a laboratory is an extension of its operator's functions.

Command is perhaps the least understood element in C^3I. It is necessary first to consider the decision making process, a small part of command. Military decision-making is a cyclic process as is illustrated in Figure 3.

The decision cycle is based on situation and resource data from lower formations, and on intelligence which may arrive from a discrete source. Commanders' decisions result in orders and information to subordinates, and may involve liaison with other commanders; see Figure 4. Command occurs at many levels, so that the structure must have height, see Figure 5. Each commander between highest and lowest levels has superiors and subordinates, which leads to a pyramid structure or, viewed from 'above', a spiders web with the commander-in-chief at the hub. See Figure 6.

Each of the decision cycles must interact with several others, and usually at three levels. At the same time, each decision cycle has its own environment, speed of response, areas of responsibility and even its own personality. Taken as a whole, the interacting decision cycles can be seen as nodes in an organism whose purpose is to provide information to, and obey the orders from, a commander.

Unfortunately the representation of Figure 6 is limited; the true representation is 3-dimensional and would involve many crossing lines.

N^2 Charts

Figure 7 shows a simple example of N^2 chart (Robert J. Lano)(3)). N^2 charts are particularly valuable for both the representation of interfaces and the analysis of C^3I structures, electronic and human. Figure 7 shows how the N^2 chart works. The chart is divided into N rows and N columns (N^2 rectangles). Leading diagonal squares represent entities, other squares represent interfaces. The interface from father to mother - No. 1 - is "love and stability" while from child to father (No. 8) come "love and problems". As important as annotated interfaces are blanks; why does father have no interfaces with teacher?

C^2 Hierarchies and N^2 Charts

Figure 8 shows a generic C^2 hierarchy committed to N^2 chart, and the figure can be seen to integrate the concepts of Figures 4 to 6 inclusive without overlapping lines. Interfaces can be seen to arise in pairs; B1 pairs with A2, E4 with D5, etc. These pairs represent the lower "legs" of Figure 4. Formations at the same level liaise and exchange intelligence - D2 and B4 - and also form interface pairs. Missing interfaces indicate a hierarchy; Air Bases report exclusively to Air Operations Centres which report in turn to the Joint Operations Centre (JOC). Multiple interfaces indicate a node; the JOC has interfaces with all the single-service Operations Centres and constitutes a node although liaison between operations centres provide some reversion should the JOC become inoperative. Indeed speed of response requires that operations centres talk directly to each other, in direct analogy to the spinal reflex arc of Figure 1, with the JOC as the higher centre of thought being informed in parallel. Blocks in which all the interfaces are filled represent functionally - bound groups which are best considered as a single formation; air defence with its C^2, Airborne Early Warning, fighters, SAM and ground sensors tends to form functionally bound group in military operations for example.

A generic N^2 chart for Air Defence is shown in Figure 9; two tightly bound functional blocks are shown, one for the Air Defence Ground Environment (ADGE) and the second for airborne forces. By one of the basic theorems of sociobiology, the more closely related an individual is to others around him in the population, the more likely he is to act altruistically and cooperatively towards them (Nigel Calder (4)). The N^2 chart, which in Figure 9 appears to represents machines but in reality represent interpersonal relationships transmitted by landline or radio links, highlights closely related groups; this concept will be enlarged upon later. The figure also in this instance reveals the nodal nature of the Airborne Early Warning (AEW) aircraft, through which all air-ground operational communications are seen to flow. Were such an organization to be instituted, the AEW would need either replication or protection to reduce system vulnerability.

C^3I Design Drivers

The design of C^3I system is influenced considerably by notions of functional binding, loose-coupling between unlike formations, and the avoidance or protection of nodes. Several key themes compete to influence both the overall and the detailed design, including:

- Survivability
- Security
- Flexibility and Expandability
- The Man-Machine Interface (MMI)
- Operational Availability
- Mobility
- Responsiveness/Performance
- Communication Capabilities and Interoperability
- Capital and Life Cycle Costs
- System reconfiguration, etc.

Reconciling C^3I system architecture with these criteria is the art of system design. The operator has to remain the focal point however.

GROUP BEHAVIOUR IN C^3I

Having examined lessons from cybernetics, there may be value in examining groups of people interacting in pursuit of an overall objective.

Information Reduction

A central theme in C^3I design is that of information reduction, which we may view initially at the individual level to establish a potential model for group behaviour.

If we presume that an individual has a considered response time of, say, 0.5 seconds (from sensor through brain to motor), then his "bandwidth" may be estimated at 2Hz. In other words he will respond to varying stimuli at about 2 responses per second.

The input sensor bandwidth to the individual is certainly much higher than 2Hz when we consider vision, hearing, smell and touch; the individual can discard, condense or review extraneous data as a background activity, and can concentrate on the matter in hand. So it must be a C^3I system, where input bandwidth from sensors and communications generally far exceeds the combined capacities of the operators.

The key to understanding the process of selecting relevant information is to be found in the Decision Cycle of Figure 3 and in the operators organization into role cells in order to support elements of the decision cycle. The elements of the decision cycle can be mapped on to an HQ organization in generic terms as follows:

Function	Principal HQ Group
Identify Strengths and Weaknesses	Intelligence, Operations
Assess Threats	Intelligence
Establish Options	Operations
Analyse Options	Operations
Select Preferred Option	Commander
Allocate Resources	Commander
Initiate Attack/Respond to Threat	Operations
Assess Situation	Commander

Supporting but essential groups include logistics, engineering, communications and plans, which present information - often of a constraining character - which limits the options.

The organization will generally comprise a number of such groups organized into cells which operate upon the information base, and which intercommunicate in support of the decision cycle.

It is likely that the cells will devolve into a 3-tiered structure as shown in Figure 10; this seems to be a "comfortable" structure for organizations, possibly related to the son, father, grandfather structure of family life. Alternatively, it may be thought of as redolent of the 3-tiered human brain, with the commander making 'high level decisions at the "fight or flight" level, operations staff conducting the routine of warfare management and specialist staff interpreting and relaying orders, situation reports and requests between lower units and the HQ information base. Taking the cell-organization and the 3-tiered structure together presents an overall picture of HQ organization. The cell structure can be almost tribal in nature, with (for example) Intelligence rarely communicating with Engineering who in turn keep themselves at a distance from Logistics. On the other hand, the 3-tiered structure evidences a specialist-manager-executive sandwich, and there may be three or more shifts operating over 24 hour periods. Reconciling these territorial and family imperatives is fundamental in a C^3I design to ensure harmonious inter-personal relationships.

Span of Control/Understanding

Humans have some little understood barrier beyond the number six. If sets of random dots are presented to a test subject for less than 1/2 second, and he is asked to state the number of such dots, most subjects will answer correctly for up to 5 or 6 dots per set; few subjects can operate consistently above 6. A commander with more than six immediate subordinates tends to neglect some and concentrate on others, particularly if the subordinates have different functions; it has been suggested (3) that the cause is man's limited spans of interest and attention and limited depth of understanding. It may also be that 7 has in the past been a typical family unit, requiring each member to have close ties with about 6 other family members. Certainly groups of seven or more people tend to split into subgroups.

Span of Control can be used as a simple guide to sizes of forces. If a general has 'S' subordinates, who each in turn have 'S'

In military circles it would be improper of the operations officer at HQ to query a logistics depot; that is the prerogative of the HQ Logistics Officer, and properly so since he will put the response in perspective, being of the same "tribe". Instead, the operations officer obtains logistics information from the HQ logistics officer who has compiled that information from "raw" depot data, other sources and his own experience. Thus military protocol, which exists for sound reasons, can be properly observed and reinforced in a function - mapped distributed system, using the important concept of data ownership (see also - Doherty and Derbyshire (5)). Each workstation, besides its own security fences has a DBMS of containable dimensions; each can be developed in relative isolation from the rest; failure in one workstation has no effect on others; and so on. Function mapped systems are however not only simpler to develop; they promise to be much more effective too because they mirror the users' organization and because each workstation is local, under immediate control, and simply "more human"!

C^3I - A GLIMPSE INTO THE FUTURE?

L.F. Richardson was a British Meteorologist interested in war (2). For the years between 1820 and 1945 he collected data on hundreds of wars, which were published in "The Statistics of Deadly Quarrels." Richardson was interested in the time interval between wars that claimed a specified number of victims. He defined an index M, the magnitude of a war, measured by the number of consequent deaths. M = 3 corresponds to (10^3 =) 1000 deaths; M = 6 corresponds to (10^6 =) 1,000,000 deaths; and so on. Richardson found that the more people killed in a war, the less likely it was to occur. His graph is shown at Figure 13, where (by the totally invalid process of extrapolation) it can be implied that the destruction of the total world population (M = 10) will not be reached for about 1000 years. Starting from 1820 this suggests that M = 10 will not be reached until the year 2820. The advent of nuclear war seriously prejudices the extrapolation and can only turn the rising end of the graph towards the horizontal. Incidentally Richardson also extrapolated towards M = 0, suggesting that the graph represented a continuum, with individual murders occurring about once every 5 minutes worldwide. The concept of a continuum may not be trivial; war may indeed be a group expression of individual motivations.

C^3I To Prevent/Moderate War

Richardson's work does remind us of the underlying inevitability of war. It has been suggested that current interest in C^3CM - C^3 Counter Measures - is grossly misplaced at least in respect of strategic HQ, since enemy strategic HQs will form vital links for the channelling of essential situation data upwards to the political controllers. It is certainly true that war results in more deaths as it becomes more impersonal as Richardson's curve implies. Perhaps then future C^3I systems will see a re-emergence of more personal forms of communication including TV so that communicants can appreciate not only tone of voice and graphics but also visual cues as well. Perhaps by showing the carnage of battle directly to those responsible for its initiation, we can restore some of the natural reflexes which inhibit us from inflicting damage beyond the point of observed submission.

CONCLUSION

The aim of the paper has been to promote a new perspective on C^3I designs. The modified mental model presented has allowed the extraction of certain points of interest as follows:

- The close relationship between sensors and C^3I is such that they should be considered as comprising one system, as by analogy the eye is part of the brain
- C^3I systems take many years to develop and of necessity much of this development must occur, as by analogy with the human infant, post delivery
- Study of biological design can reveal useful C^3I design alternatives as for example the analogy between pain gates and C^3I alerting mechanism, and the potential of power supplies to carry information throughout a wide subscriber population
- Analyses of group behaviour show that our military organisations form both nodes and functionally bound groups. These groupings may be expressions of our deep-rooted territorial imperatives and should be given due weight in C^3I designs by the adoption of philosophies such as function mapped architectures
- In the final analysis, decisions are made by individuals who can operate at no more than 2 decisions per second. No matter how much money is invested in enhanced communications, data processing and display technology, the finite human capability remains the limting factor
- The concept of strategic C^3CM may not always favour the best interests of antagonists, since successful C^3CM operations may prevent governments from becoming aware in time of submission by either of the opposing military groups
- L.F. Richardson's model of the inevitable escallation of war emphasises the need for ever better C^3I systems, so that we may limit the seemingly inexorable trend towards overkill in conflict.

REFERENCES

1. Robert Barrass, 1979, "Modern Biology", 226-237
2. Carl Sagan, 1980, "Cosmos"
3. Robert J. Lano, 1980, Operational Concept Formulation
4. Nigel Calder, 1978 "Spaceships of the Mind".
5. Doherty and Derbyshire, 1984, A Distributed Approach to CCIS design.

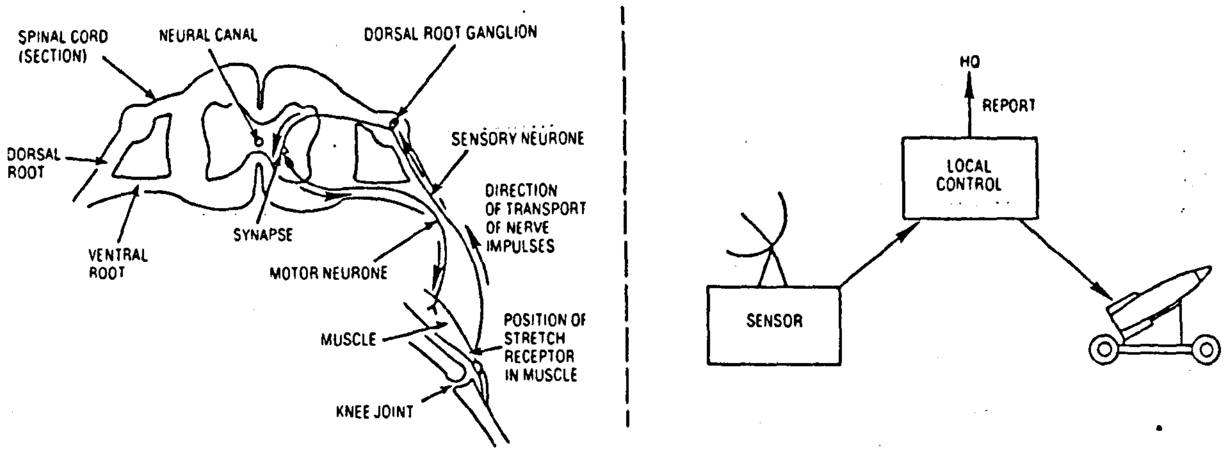
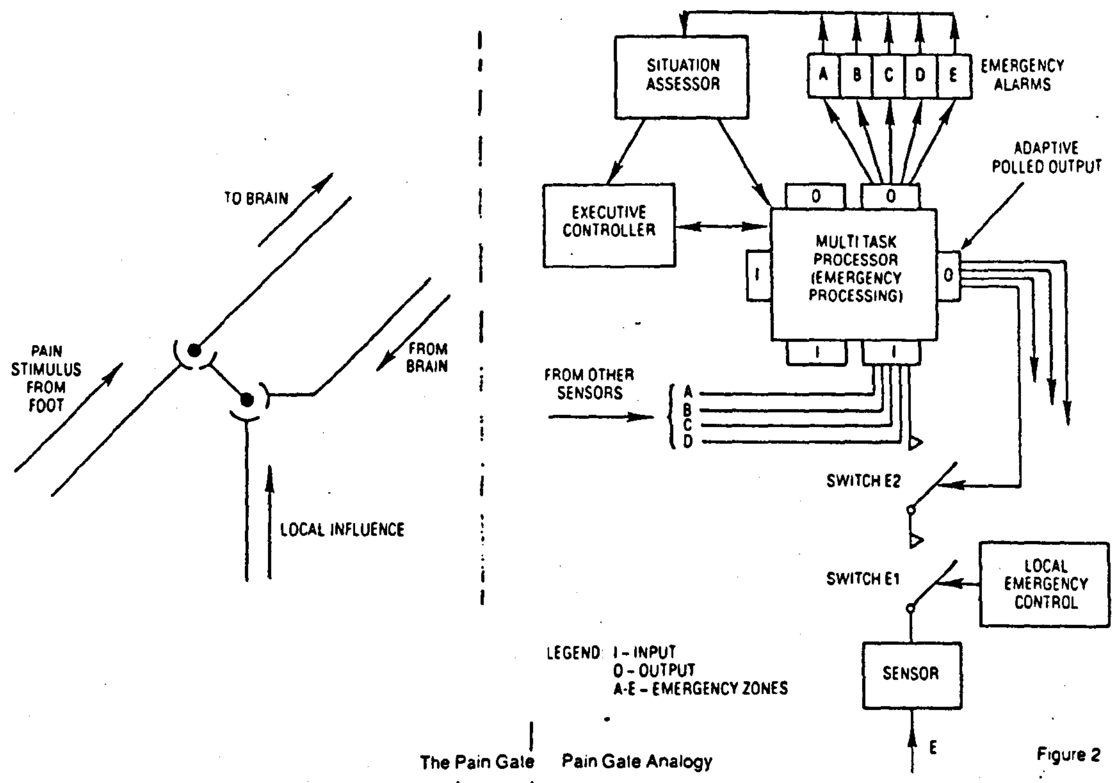


Figure 1 Local Fast-Response Control



The Pain Gate | Pain Gate Analogy

Figure 2

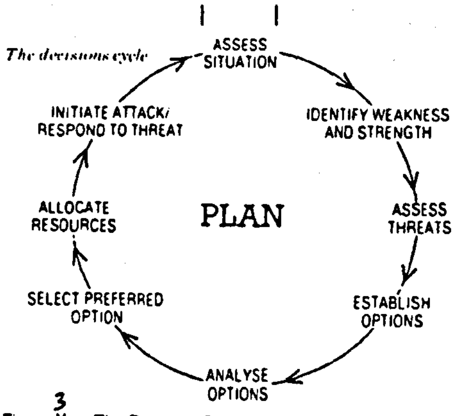


Figure ~~X~~ 3 The Decision Cycle

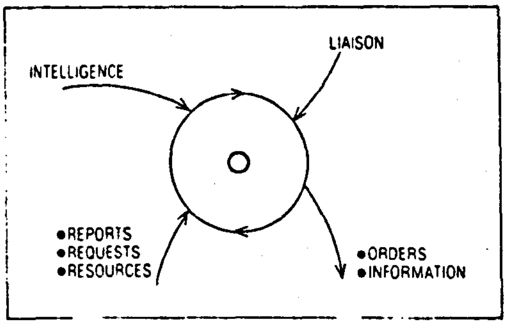


Figure 4 Vertical Command Structure

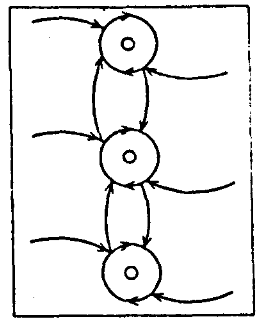


Figure 5 inputs and Outputs

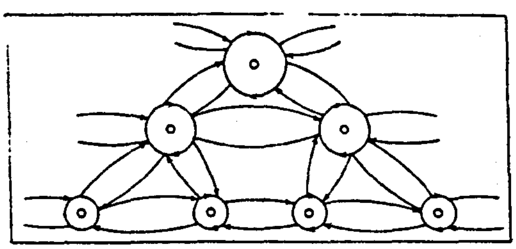
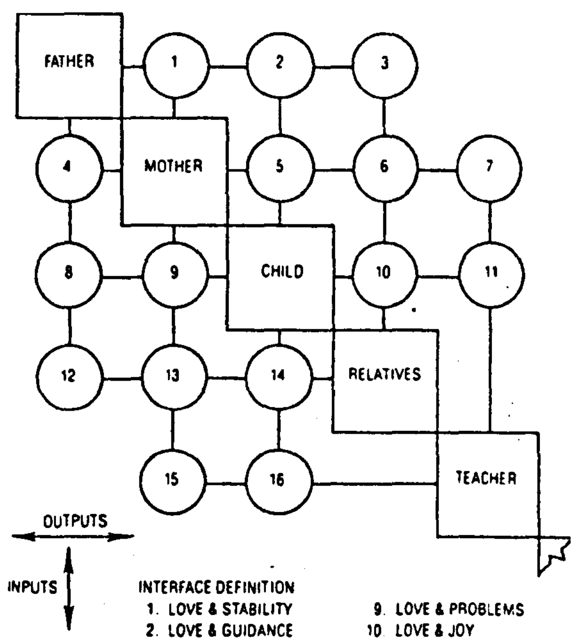


Figure 6 Pyramid Command Structure



INTERFACE DEFINITION

1. LOVE & STABILITY
2. LOVE & GUIDANCE
3. LOVE & TIME
4. LOVE & DEVOTION
5. LOVE & UNDERSTANDING
6. LOVE & TIME
7. CONCERN & QUESTIONS
8. LOVE & PROBLEMS
9. LOVE & PROBLEMS
10. LOVE & JOY
11. QUESTIONS & PROBLEMS
12. LOVE & PROBLEMS
13. LOVE & ADVICE
14. LOVE & PRESENTS
15. STATUS & PROBLEMS
16. WORK & DISCIPLINE

Figure 7 Family Relationships on an N² Chart

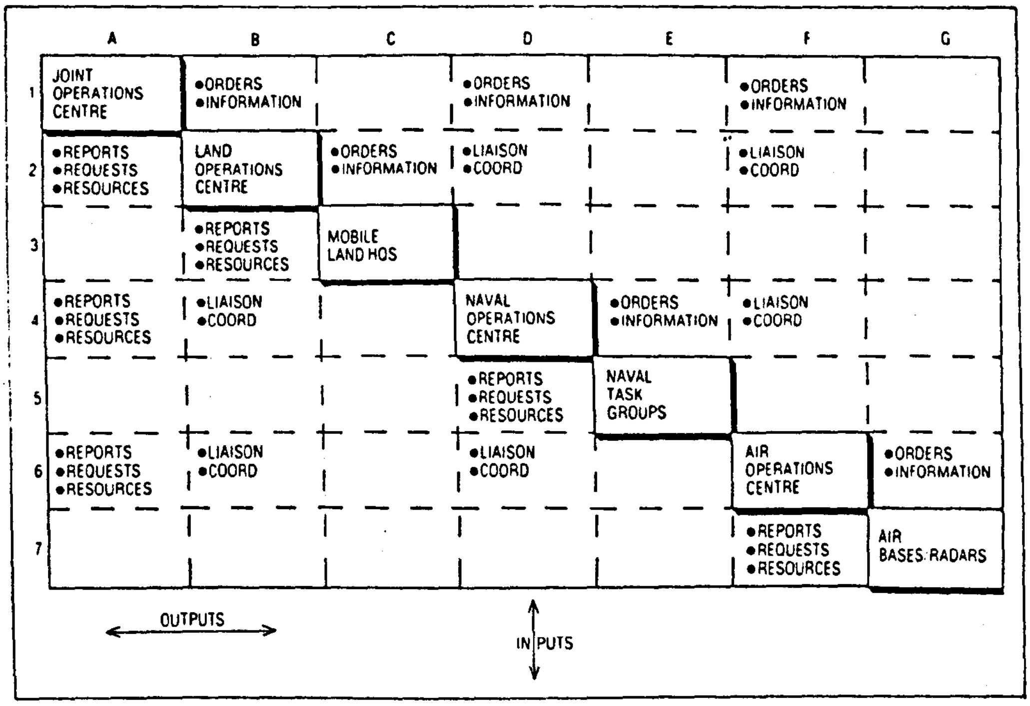
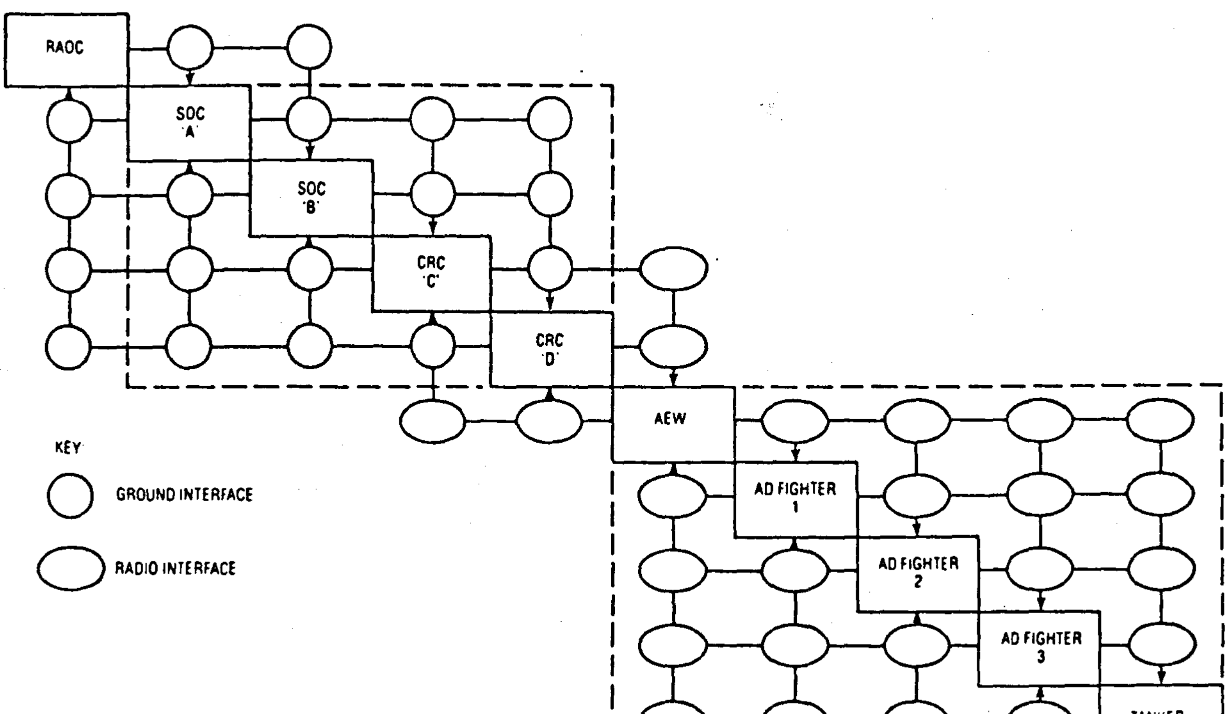


Figure 8 Generic C^3 I Architecture



NOTE THIS DIAGRAM IS EMPLOYED TO DEMONSTRATE THE N^2 TECHNIQUE
IT DOES NOT PURPORT TO REPRESENT A FUTURE SYSTEM

Figure 9 Air Defence N^2 Chart

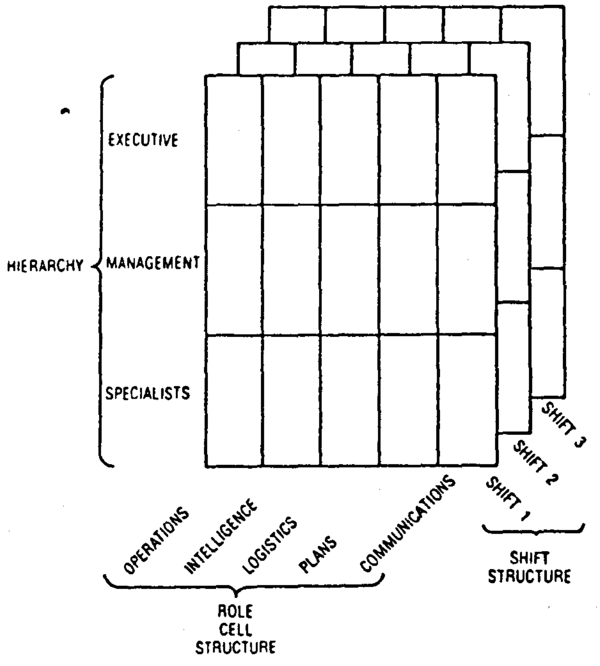


Figure 10 C³I Organisation Structures

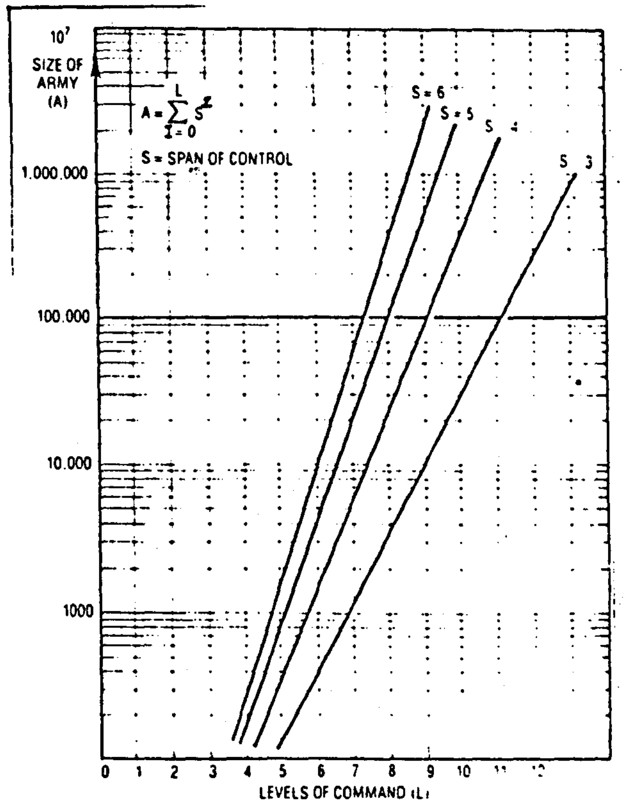


Figure 11 Graph Showing (Simplified) Relationship between Size of Army and Span of Control

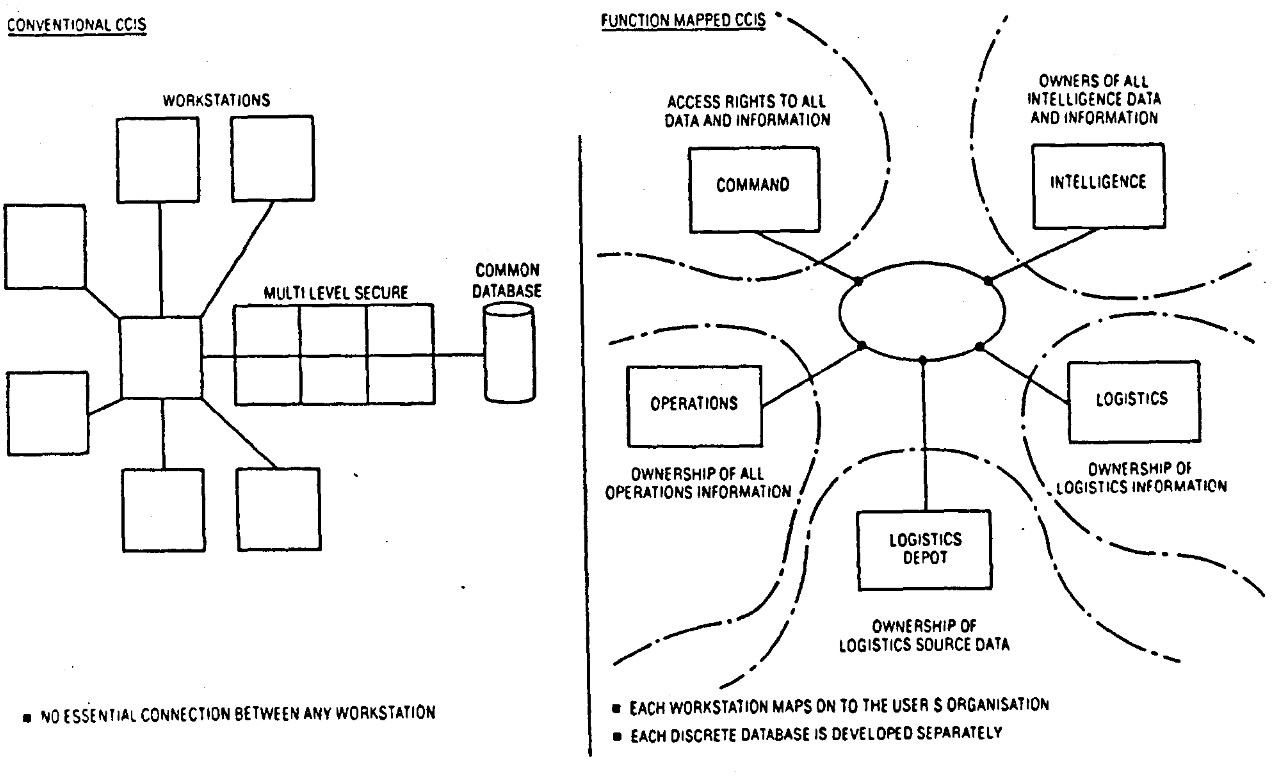
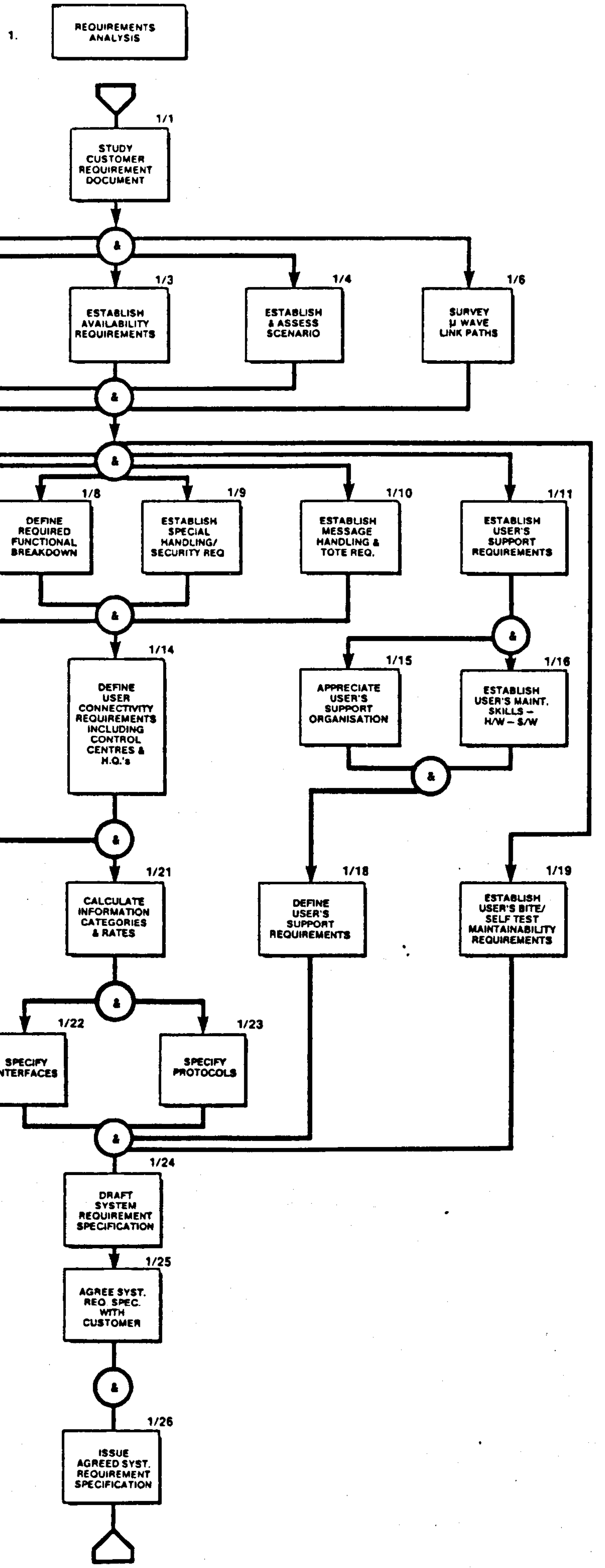
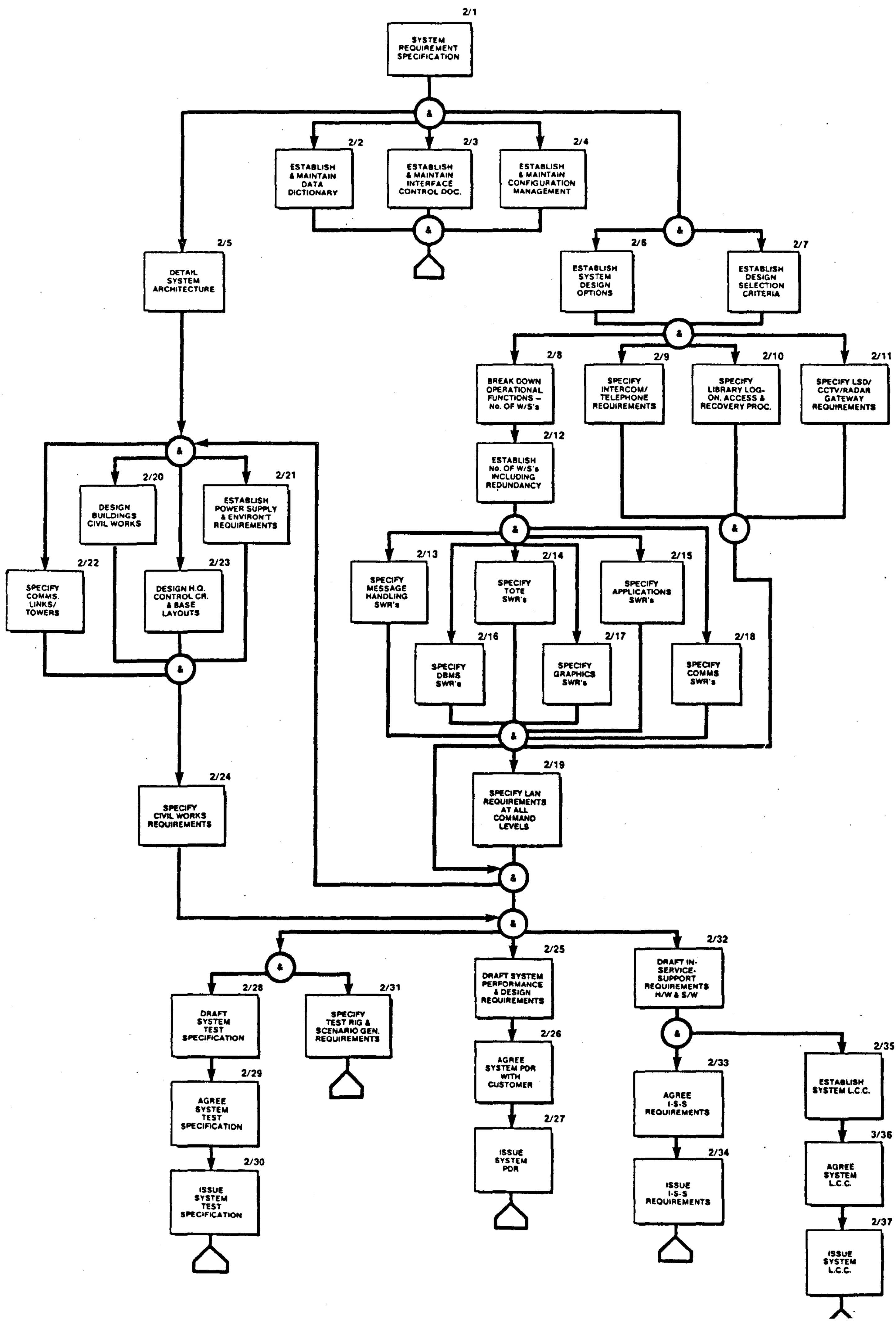
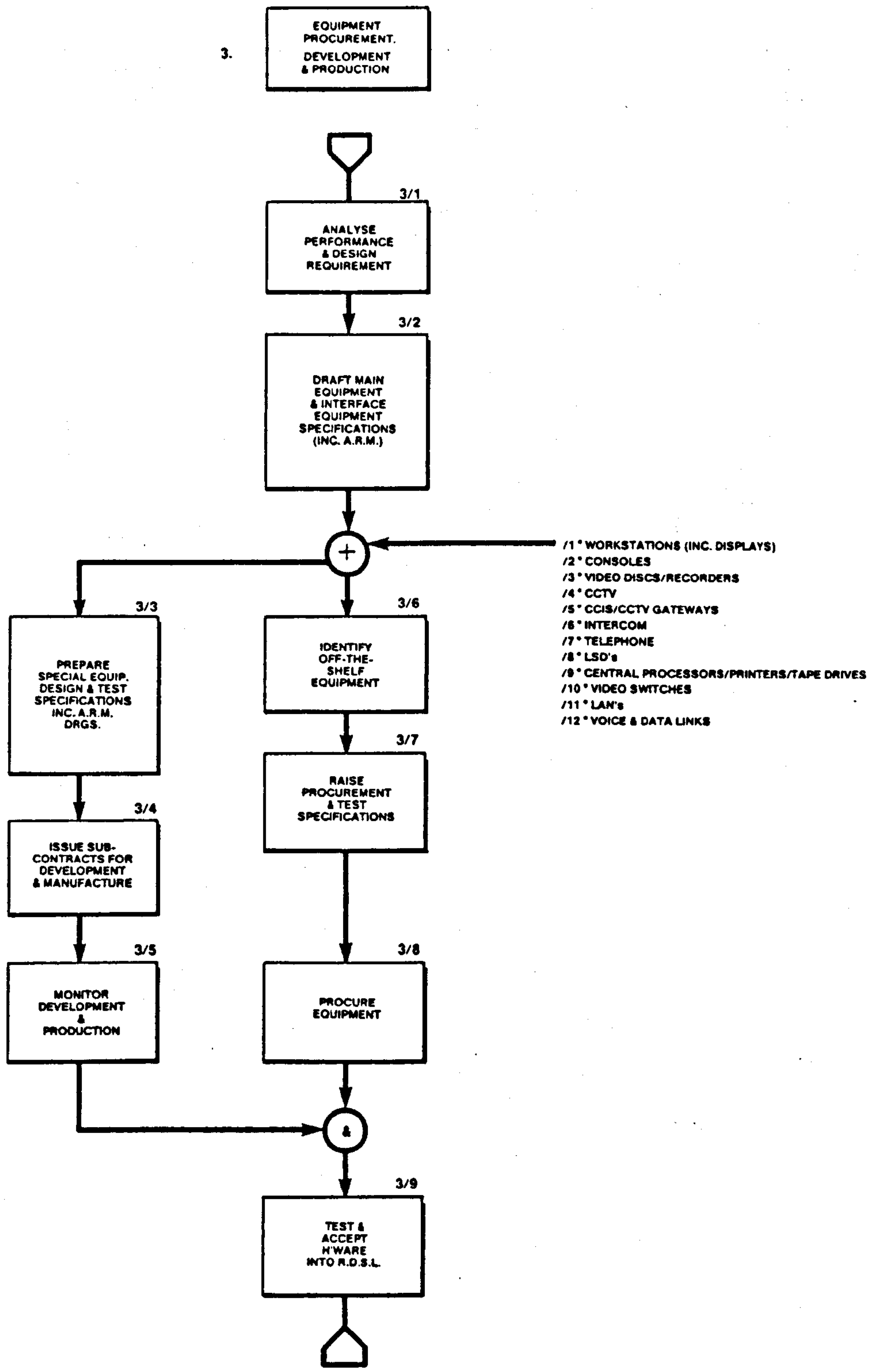


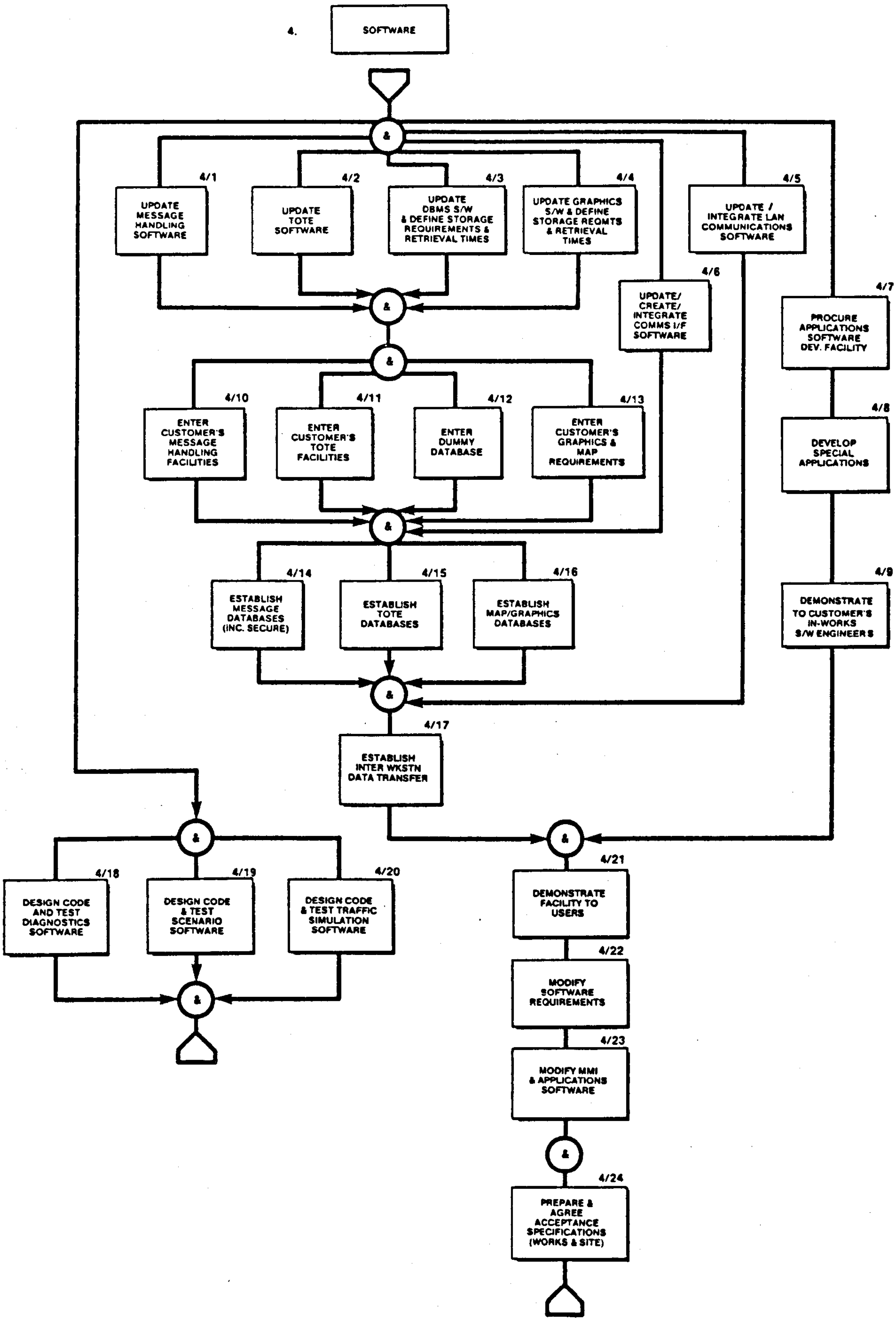
Figure 12 Function Mapping and Database Ownership


1.
REQUIREMENTS ANALYSIS
1/1
STUDY CUSTOMER REQUIREMENT DOCUMENT
&
1/5
SURVEY BASE INT./EXT. COMMS LINKS
1/2
ESTABLISH MAJOR DESIGN DRIVERS
1/3
ESTABLISH AVAILABILITY REQUIREMENTS
1/4
ESTABLISH & ASSESS SCENARIO
1/6
SURVEY μ WAVE LINK PATHS
&
&
1/7
APPRECIATE USER'S ORGANISATION
1/8
DEFINE REQUIRED FUNCTIONAL BREAKDOWN
1/9
ESTABLISH SPECIAL HANDLING/ SECURITY REQ
1/10
ESTABLISH MESSAGE HANDLING & TOTE REQ.
1/11
ESTABLISH USER'S SUPPORT REQUIREMENTS
&
&
&
1/13
IDENTIFY USERS EXIST. FACILITIES (INCL. POWER)
1/12
DEFINE SYSTEM BOUNDARIES
1/17
ESTABLISH USER PROCEDURES
&
1/20
DEFINE OVERALL SYSTEM ARCHITECTURE
1/14
DEFINE USER CONNECTIVITY REQUIREMENTS INCLUDING CONTROL CENTRES & H.Q.'s
&
1/21
CALCULATE INFORMATION CATEGORIES & RATES
&
1/22
SPECIFY INTERFACES
1/23
SPECIFY PROTOCOLS
1/15
APPRECIATE USER'S SUPPORT ORGANISATION
1/16
ESTABLISH USER'S MAINT. SKILLS – H/W – S/W
&
1/18
DEFINE USER'S SUPPORT REQUIREMENTS
1/19
ESTABLISH USER'S BITE/ SELF TEST MAINTAINABILITY REQUIREMENTS
&
1/24
DRAFT SYSTEM REQUIREMENT SPECIFICATION
1/25
AGREE SYST. REQ SPEC. WITH CUSTOMER
&
1/26
ISSUE AGREED SYST. REQUIREMENT SPECIFICATION

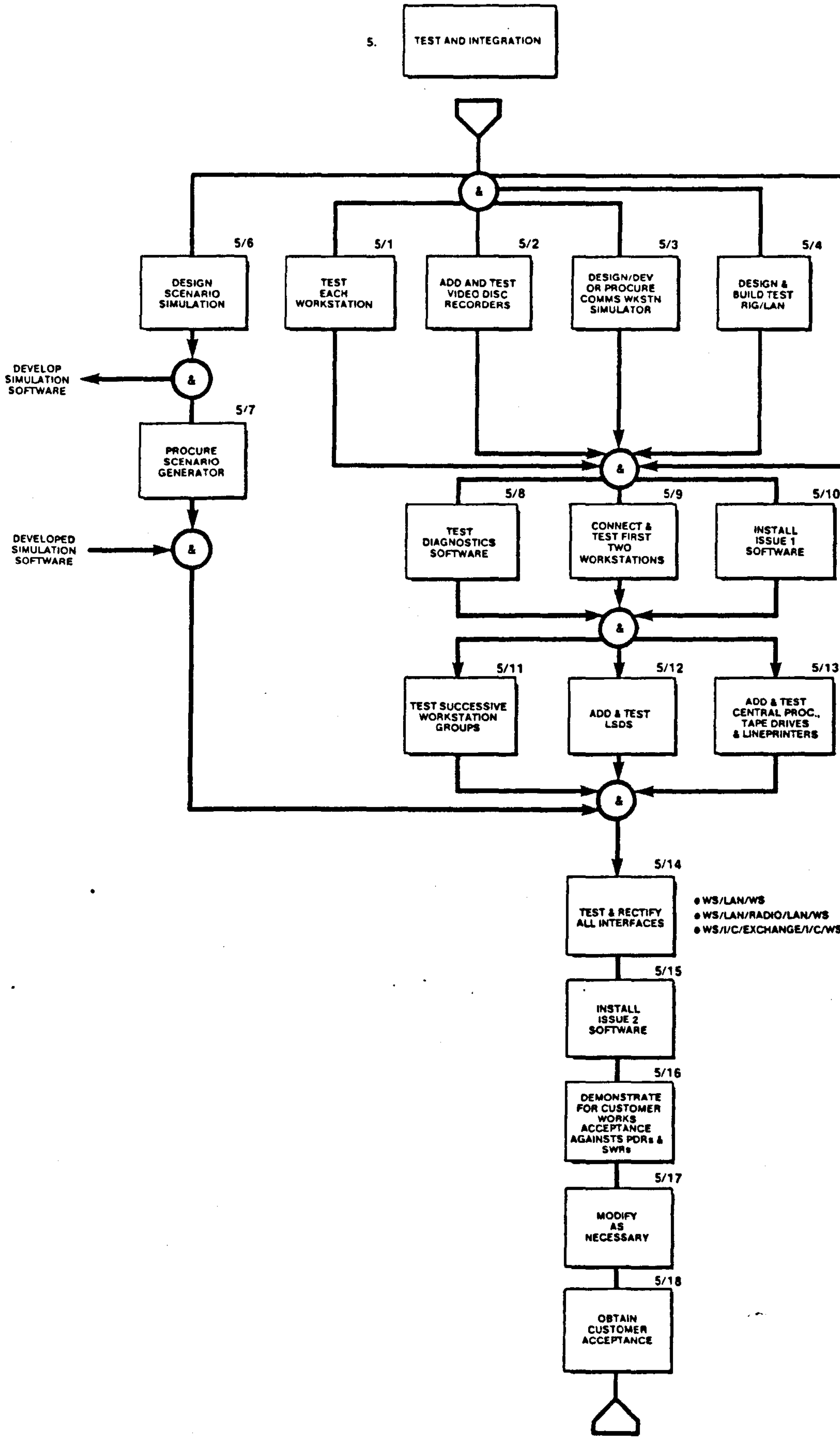
2. SYSTEM DESIGN

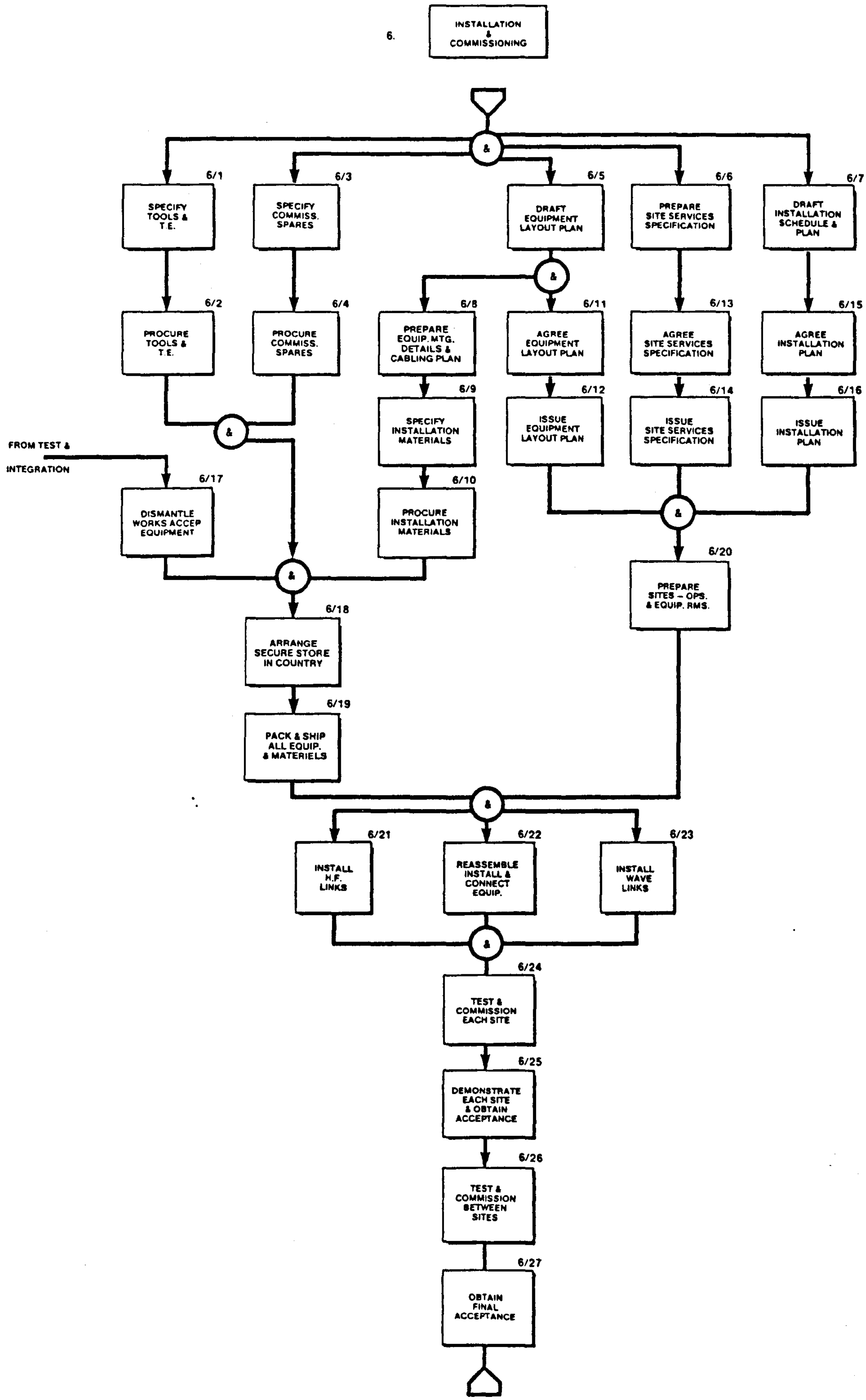


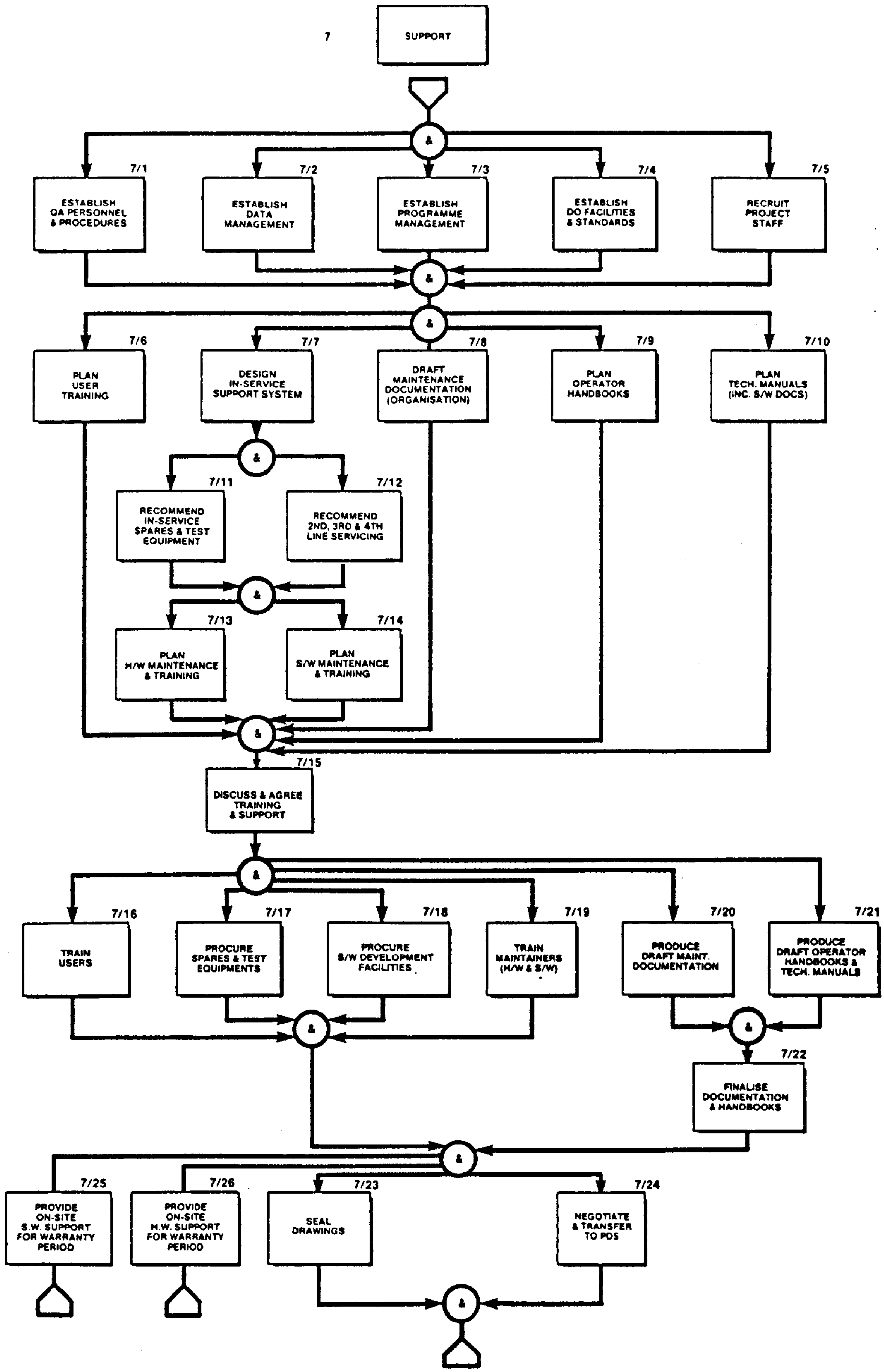

3.
EQUIPMENT PROCUREMENT. DEVELOPMENT & PRODUCTION
3/1
ANALYSE PERFORMANCE & DESIGN REQUIREMENT
3/2
DRAFT MAIN EQUIPMENT & INTERFACE EQUIPMENT SPECIFICATIONS (INC. A.R.M.)
/1 * WORKSTATIONS (INC. DISPLAYS)
/2 * CONSOLES
/3 * VIDEO DISCS/RECORDERS
/4 * CCTV
/5 * CCIS/CCTV GATEWAYS
/6 * INTERCOM
/7 * TELEPHONE
/8 * LSD's
/9 * CENTRAL PROCESSORS/PRINTERS/TAPE DRIVES
/10 * VIDEO SWITCHES
/11 * LAN's
/12 * VOICE & DATA LINKS
3/3
PREPARE SPECIAL EQUIP. DESIGN & TEST SPECIFICATIONS INC. A.R.M. DRGS.
3/4
ISSUE SUB-CONTRACTS FOR DEVELOPMENT & MANUFACTURE
3/5
MONITOR DEVELOPMENT & PRODUCTION
3/6
IDENTIFY OFF-THE-SHELF EQUIPMENT
3/7
RAISE PROCUREMENT & TEST SPECIFICATIONS
3/8
PROCURE EQUIPMENT
&
3/9
TEST & ACCEPT H'WARE INTO R.D.S.L.



4.
SOFTWARE
4/1 UPDATE MESSAGE HANDLING SOFTWARE
4/2 UPDATE TOTE SOFTWARE
4/3 UPDATE DBMS S/W & DEFINE STORAGE REQUIREMENTS & RETRIEVAL TIMES
4/4 UPDATE GRAPHICS S/W & DEFINE STORAGE REQMTS & RETRIEVAL TIMES
4/5 UPDATE / INTEGRATE LAN COMMUNICATIONS SOFTWARE
4/6 UPDATE/ CREATE/ INTEGRATE COMMS I/F SOFTWARE
4/7 PROCURE APPLICATIONS SOFTWARE DEV. FACILITY
4/8 DEVELOP SPECIAL APPLICATIONS
4/9 DEMONSTRATE TO CUSTOMER'S IN-WORKS S/W ENGINEERS
4/10 ENTER CUSTOMER'S MESSAGE HANDLING FACILITIES
4/11 ENTER CUSTOMER'S TOTE FACILITIES
4/12 ENTER DUMMY DATABASE
4/13 ENTER CUSTOMER'S GRAPHICS & MAP REQUIREMENTS
4/14 ESTABLISH MESSAGE DATABASES (INC. SECURE)
4/15 ESTABLISH TOTE DATABASES
4/16 ESTABLISH MAP/GRAPHICS DATABASES
4/17 ESTABLISH INTER WKSTN DATA TRANSFER
4/18 DESIGN CODE AND TEST DIAGNOSTICS SOFTWARE
4/19 DESIGN CODE & TEST SCENARIO SOFTWARE
4/20 DESIGN CODE & TEST TRAFFIC SIMULATION SOFTWARE
4/21 DEMONSTRATE FACILITY TO USERS
4/22 MODIFY SOFTWARE REQUIREMENTS
4/23 MODIFY MMI & APPLICATIONS SOFTWARE
4/24 PREPARE & AGREE ACCEPTANCE SPECIFICATIONS (WORKS & SITE)


5.
TEST AND INTEGRATION
&
5/6
DESIGN SCENARIO SIMULATION
5/1
TEST EACH WORKSTATION
5/2
ADD AND TEST VIDEO DISC RECORDERS
5/3
DESIGN/DEV OR PROCURE COMMS WKSTN SIMULATOR
5/4
DESIGN & BUILD TEST RIG/LAN
5/5
PROCURE SPECIAL TEST EQUIPMENT
DEVELOP SIMULATION SOFTWARE
&
5/7
PROCURE SCENARIO GENERATOR
&
DEVELOPED SIMULATION SOFTWARE
&
5/8
TEST DIAGNOSTICS SOFTWARE
5/9
CONNECT & TEST FIRST TWO WORKSTATIONS
5/10
INSTALL ISSUE 1 SOFTWARE
&
5/11
TEST SUCCESSIVE WORKSTATION GROUPS
5/12
ADD & TEST LSDS
5/13
ADD & TEST CENTRAL PROC., TAPE DRIVES & LINEPRINTERS
&
5/14
TEST & RECTIFY ALL INTERFACES
• WS/LAN/WS
• WS/LAN/RADIO/LAN/WS
• WS/I/C/EXCHANGE/I/C/WS
5/15
INSTALL ISSUE 2 SOFTWARE
5/16
DEMONSTRATE FOR CUSTOMER WORKS ACCEPTANCE AGAINSTS PDRs & SWRs
5/17
MODIFY AS NECESSARY
5/18
OBTAIN CUSTOMER ACCEPTANCE


6.
INSTALLATION & COMMISSIONING
6/1 SPECIFY TOOLS & T.E.
6/3 SPECIFY COMMISS. SPARES
6/5 DRAFT EQUIPMENT LAYOUT PLAN
6/6 PREPARE SITE SERVICES SPECIFICATION
6/7 DRAFT INSTALLATION SCHEDULE & PLAN
6/2 PROCURE TOOLS & T.E.
6/4 PROCURE COMMISS. SPARES
6/8 PREPARE EQUIP. MTG. DETAILS & CABLING PLAN
6/11 AGREE EQUIPMENT LAYOUT PLAN
6/13 AGREE SITE SERVICES SPECIFICATION
6/15 AGREE INSTALLATION PLAN
6/9 SPECIFY INSTALLATION MATERIALS
6/12 ISSUE EQUIPMENT LAYOUT PLAN
6/14 ISSUE SITE SERVICES SPECIFICATION
6/16 ISSUE INSTALLATION PLAN
FROM TEST & INTEGRATION
6/17 DISMANTLE WORKS ACCEP EQUIPMENT
6/10 PROCURE INSTALLATION MATERIALS
6/20 PREPARE SITES – OPS. & EQUIP. RMS.
6/18 ARRANGE SECURE STORE IN COUNTRY
6/19 PACK & SHIP ALL EQUIP. & MATERIELS
6/21 INSTALL H.F. LINKS
6/22 REASSEMBLE INSTALL & CONNECT EQUIP.
6/23 INSTALL WAVE LINKS
6/24 TEST & COMMISSION EACH SITE
6/25 DEMONSTRATE EACH SITE & OBTAIN ACCEPTANCE
6/26 TEST & COMMISSION BETWEEN SITES
6/27 OBTAIN FINAL ACCEPTANCE


7
SUPPORT
7/1
ESTABLISH QA PERSONNEL & PROCEDURES
7/2
ESTABLISH DATA MANAGEMENT
7/3
ESTABLISH PROGRAMME MANAGEMENT
7/4
ESTABLISH DO FACILITIES & STANDARDS
7/5
RECRUIT PROJECT STAFF
7/6
PLAN USER TRAINING
7/7
DESIGN IN-SERVICE SUPPORT SYSTEM
7/8
DRAFT MAINTENANCE DOCUMENTATION (ORGANISATION)
7/9
PLAN OPERATOR HANDBOOKS
7/10
PLAN TECH. MANUALS (INC. S/W DOCS)
7/11
RECOMMEND IN-SERVICE SPARES & TEST EQUIPMENT
7/12
RECOMMEND 2ND, 3RD & 4TH LINE SERVICING
7/13
PLAN H/W MAINTENANCE & TRAINING
7/14
PLAN S/W MAINTENANCE & TRAINING
7/15
DISCUSS & AGREE TRAINING & SUPPORT
7/16
TRAIN USERS
7/17
PROCURE SPARES & TEST EQUIPMENTS
7/18
PROCURE S/W DEVELOPMENT FACILITIES
7/19
TRAIN MAINTAINERS (H/W & S/W)
7/20
PRODUCE DRAFT MAINT. DOCUMENTATION
7/21
PRODUCE DRAFT OPERATOR HANDBOOKS & TECH. MANUALS
7/22
FINALISE DOCUMENTATION & HANDBOOKS
7/25
PROVIDE ON-SITE S.W. SUPPORT FOR WARRANTY PERIOD
7/26
PROVIDE ON-SITE H.W. SUPPORT FOR WARRANTY PERIOD
7/23
SEAL DRAWINGS
7/24
NEGOTIATE & TRANSFER TO PDS

Managing system creation

D.K. Hitchins, M.Sc., C.Eng., F.I.E.E., M.R.Ae.S., A.M.B.I.M.

Indexing terms: *Engineering administration and management, Project and production engineering, Management*

Abstract: The paper introduces systems methodologies to illustrate the span of systems engineering activities and the system management structure appropriate to creating a complex and enduring system. Information/decision/action (IDA) systems provide a vehicle to demonstrate a variety of methodologies in operation. This class of system includes individual humans, companies, governments, computer-based data systems and, paradoxically, the system design/management team intent on creating the respective systems themselves. Systems are viewed as having seven ages: conception, design, development, implementation, transition utility and senility (which is followed by replacement). System creation invokes five subsystems: the operational facility, two development subsystems concerned with test-and-integration and incompany support, and two deliverable inservice subsystems for user/operator training and through life maintenance. System creation team structure is shown to comprise operations analysis, requirements analysis, system design, equipment engineering, software engineering, test and integration, and installation and commissioning; the systems approach emphasises the first three of these in particular, to reduce the generation of *ab initio* design errors. Functional decomposition and design option tradeoffs are demonstrated by example, and the establishment of end-to-end system dynamics is introduced using the ISO open system interconnection philosophy. System management organisations are outlines, and the allocation of resources by time within the project team structure is explored pictorially. Transition into operational use is highlighted as an area of special concern and, finally, the key issues of the systems management approach are identified.

Introduction

Much has been published on the subject of systems engineering, its methodologies, techniques and benefits. Relatively little has been produced on the management of systems engineering projects; this paper seeks to remedy that shortfall.

The word 'systems' is so widely used that it has become necessary to provide a taxonomy. The broad class of systems considered in the paper is that in which humans or groups of humans gather information, reach agreements and make decisions based on that information and implement consequent actions. This information/decision/action (IDA) classification has been selected not only because it is so wide in application, but also because it contains probably the most difficult categories of system design management, those concerned with multi-user/machine interactions and human factors.

The paper may be read at several levels. For example, it may be viewed as an embryonic company procedure document, or a largely pictorial description of IDA system creation. On the other hand, the principles and graphical techniques employed in the paper can be viewed as simple examples which might be re-employed in a much wider context, from detailed analysis of production organisations to the broad design of management systems.

In a similar vein, the paper could be viewed as arguing the case for systems engineering as a separate discipline, and in a way it does. Nobody questions the role of the architect in civil engineering, although his individual engineering skills are no different from, and may be no better developed than, those of the civil engineering contractors. His major contributions are creativity, structure and balance within constraints imposed by a particular situation. And so it should be for the system designer; he too covers a broad span of disciplines and seeks to bring creativity, structure and balance to design. And he too deserves separate recognition.

Paper 4783A (M3, M4), first received 14th November 1985 and in revised form 9th January 1986

The author was formerly Technical Director of Racal Defence Systems Ltd, Richmond Court, 309 Fleet Road, Fleet, Hampshire GU13 8BU, and resides at Sarum House, 4 Clevedon Court, Frimley, Surrey GU16 5YW, United Kingdom

The paper is presented in 'storyboard'. Each page of text is accompanied by a diagram, which contributes directly and without separate explanation to the topic under discussion. A suitable procedure is to read the first section of the text, examine the diagram and then return to the remaining text for expansion.

1 Information/decision/action (IDA) systems

This paper is concerned with the creation of the wide class of IDA systems which share the characteristic that they all involve human decision making. System creation to support this decision process must reflect the complex nature of the human relationships between the system user and operators. Effective management of system creation is characterised by its own skills and techniques.

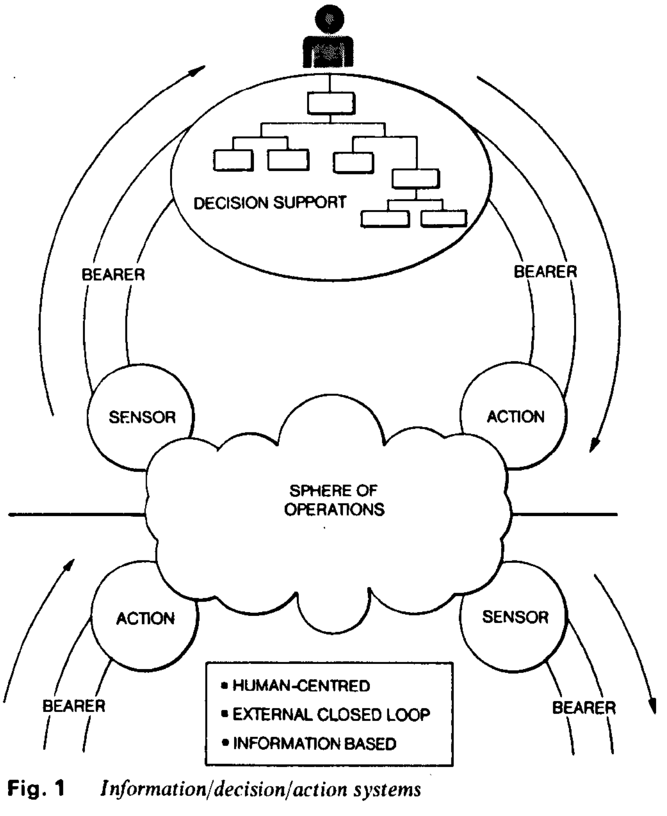


Fig. 1 *Information/decision/action systems*

Fig. 1 shows the information/decision/action class of system with which this volume is principally concerned. It comprises sensing, communication, assessment of the sensed information, human-based decision and consequent controlled action.

The class of system embraces a very wide variety of members, including:

(i) individuals and groups of humans
(ii) piloted vehicles
(iii) companies
(iv) governments
(v) air traffic control
(vi) finance and banking
(vii) management information
(viii) command and control information
(ix) intelligence
(x) IDA system management of creation.

Excluded from the class, by virtue of their fully automatic nature, are process control systems not requiring human decision. However, process control may occur in the subsystems from which information/decision/action (IDA) systems are constructed, particularly in sensing and execution subsystems.

IDA systems are generally highly complex. Complexity arises principally from the intimate relationship which must be established between the IDA systems's components and the human organisation which it is intended to serve. The decision support subsystem in particular can comprise many operators, each with his own formation handling facilities, concerned with different aspects of an overall decision process. The 'real' decision process is occurring, then, between the system operators, while the system in the Figure must provide carefully tailored support to each individual in the decision chain. This paper is concerned with the means for creating that synergy between operator and system in a cost-effective way, using the system management approach.

2 System lifespan

A system will evolve during its lifespan and should be designed with that ability from the outset. Eventually system senility will set in, however; good system design seeks to delay the onset of senility and to support ease of transfer for the collected information from the old to the new system.

Fig. 2 shows eight phases in the life of a system, from conception to eventual replacement. System creation is usually a shorter period than system operation. The creation process must, then, recognise from the outset that the system will evolve during its useful life, and 'design for evolution' is essential.

Transition is the period of introduction of the system into operation. Designers often envisage a system only as they expect it to be during its operational state of utility. Achieving this goal requires special facilities and procedures which may be peculiar to the transition phase, so that the human operators can transfer smoothly and with enthusiasm from the old to the new system.

An objective of system design should be to stay the onset of system senility, the causes of which will be discussed later. The onset of senility is to some extent, however, a self-fulfilling prophecy. Conception, design and development of a new system may be so protracted that senility must be anticipated, and the new system may be ready before the old really needs replacement.

An appreciation of the terminal phases of a system's lifespan is valuable in the system creation process. As the figure shows, the new system will interface with the system it replaces and information will be transferred to the new system during start-up. It is, therefore, sound practice in that context to design for ease of replacement, so that information collected and stored, perhaps at great expense, in the old system can be transferred to, and retained in, the new.

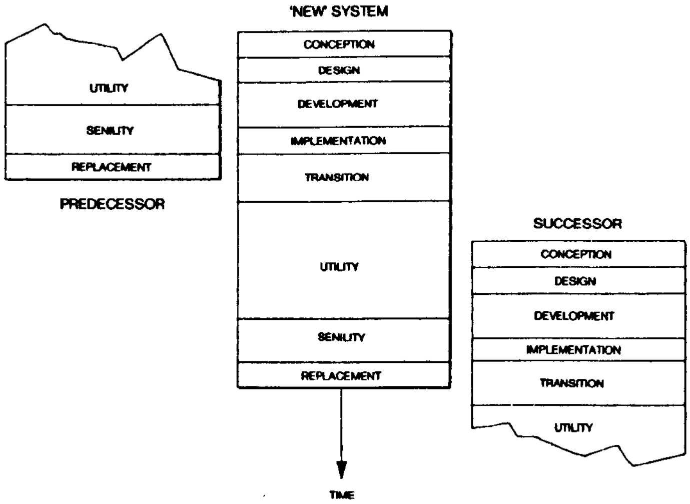


Fig. 2 *System lifespan*

3 System life cycle

System designs contain the seeds of their own eventual decline. The management of systems creation generates system options and balances their merits against life-cycle cost so as to maximise the period of cost-effective system utility.

Fig. 3 consists of a graph with axes of cost per unit time and effectiveness. The progress of a typical system is plotted against these axes. Rising costs, with no capability, start the cycle for development and implementation. During transition the system starts to become effective and the rate of expenditure may fall. The period of utility is marked by rising effectiveness and, in mid term, rising cost of ownership. Finally, effectiveness starts to tail off with costs still rising into senility and ultimate replacement.

The Figure also shows some of the causes of the cost-effectiveness cycle. Designs are usually frozen early in development, and it is common for hardware in particular to be dated even before entry into use. Protracted development and implementation also result in a mismatch between the evolving user requirement and the more static

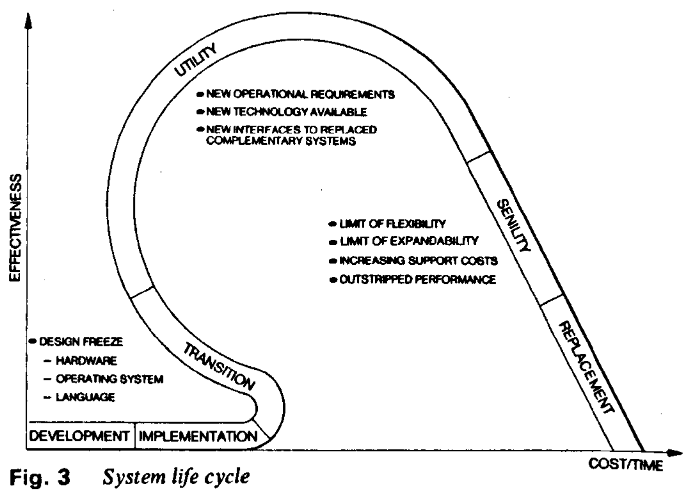


Fig. 3 *System life cycle*

design, so that design changes can be needed during transition as well as the period of utility. The new system will also interface, in the wider system context, with complementary systems which are at different phases in their life cycles, and new replacements will arise as situations change.

These changing requirements can be accommodated only by providing inherent and accessible flexibility and spare capacity in the system design. Even these pledges to adaptability must all be redeemed eventually, at which time the system can no longer evolve in step with the changing wider system and must eventually decline.

The system design approach recognises these evolutionary factors and trades off performance and continued performance, on the one hand, against life cycle cost, on the other. The generation of system options and the tradeoff process are a fundamental responsibility of systems management.

(*d*) The Inservice User Training System (to train operational users) comprising trainers, simulators and analysis elements, and usually separate from the operational system *per se*.

Of special note is the composition of the operational system; it should comprise the five integral subsystem as shown in the Figure, so that it contains its own reversion (secondary system to operate when the primary system fails) and its own transition features, as well as the ability to detect faults and exercise operators.

A design which is intended to be life-cycle cost effective will capitalise on the potential similarities between the inservice and the incompany support systems and between the user training and the test and integration subsystems. For example, the customer's need for automatic test facilities in service can influence the choice of incompany test facilities, so that test access, test techniques, test software and test hardware are compatible. This approach offers shorter timescales and reduced costs, as well as more harmonious transition and eventual contractor support.

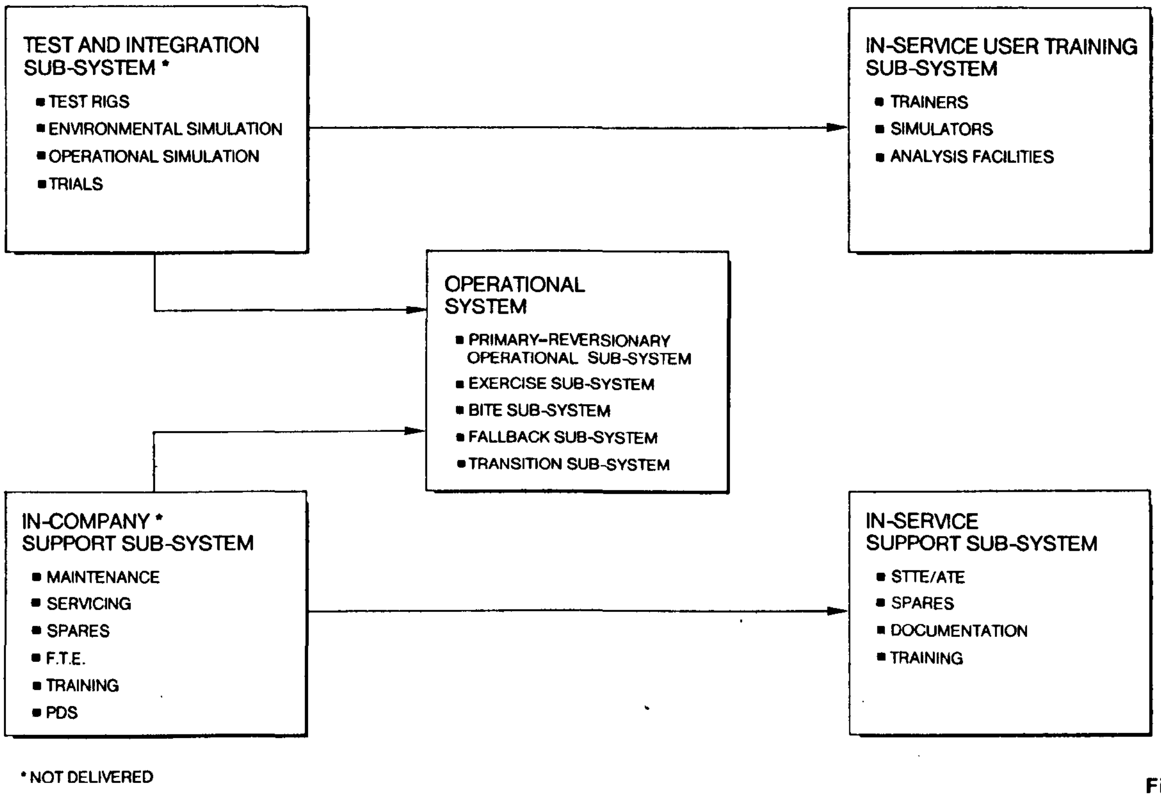


Fig. 4 *Project subsystems*

4 Project subsystems

Each system is comprised of at least five related subsystems, and the operational system should contain its own reversionary, fallback and transitional capabilities. Inservice and incompany subsystem compatibility can be promoted to maximise life-cycle cost effectiveness.

Systems and their eventual users need to be supported. Consideration of a system project shows that it comprises at least four major subsystems in addition to the operational system required by the customer. These four are:

(*a*) The Test and Integration Subsystem needed to prove the developing operational system's elements

(*b*) The Incompany Support Subsystem needed to maintain both the test and integration facilities and the developing operational equipments

(*c*) The Inservice Support Subsystem which provides spares, test facilities, training, documentation and development facilities for the user's operational and inservice training systems

5 System project team structure

System creation comprises seven phases which can be reflected in the team organisation. The resultant teams are bound to each other through the formal use of specifications which are developed, agreed and maintained in cooperation with the customer. The team management is ideally established so that system design and development must be proven to the satisfaction of an independent test authority within the project, supported by the QA authority.

The system creation process comprises seven major phases. Project organisation provides seven equivalent functional groups. These tightly bound functional groups provide discrete outputs as shown in Fig. 5. Coupling between the groups is provided by formal specifications for each of the five subsystems shown in the previous topic.

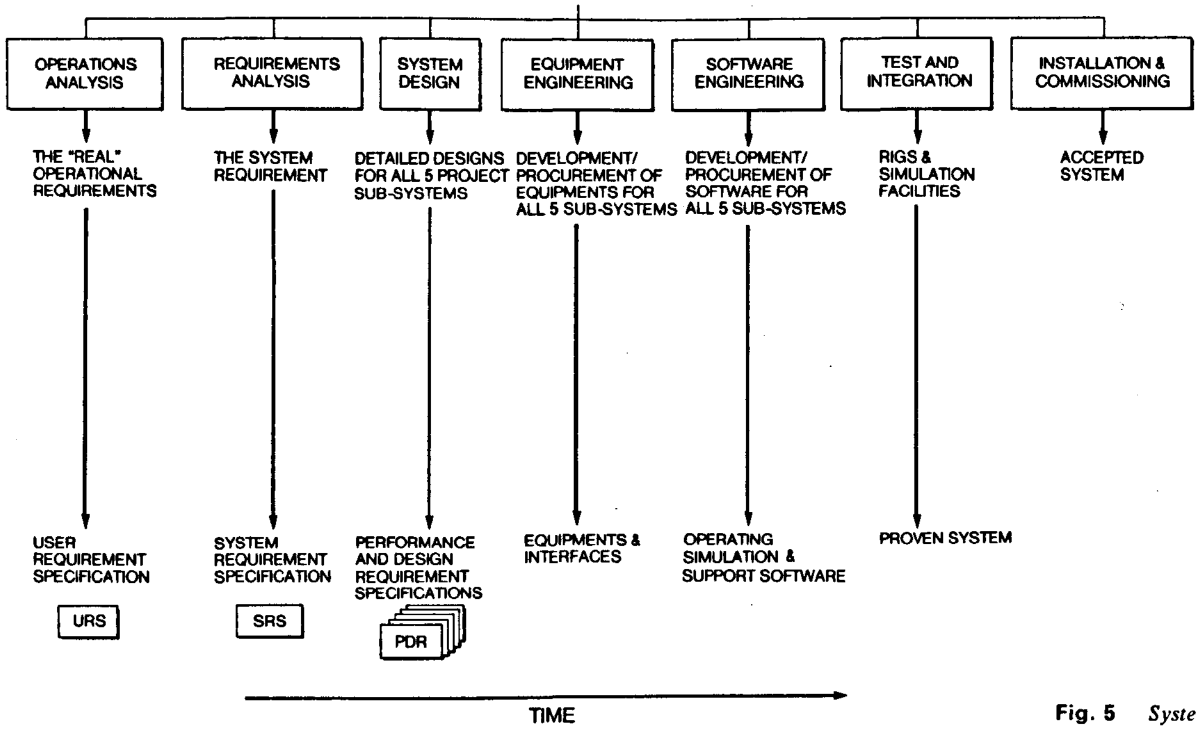


Fig. 5 *System project team structure*

The progressive elaboration of specifications is an essential if unromantic basis for effective system creation; the specifications are agreed stage by stage with the customer and evolved as both the users and the system creator's ideas develop.

The user requirement specification (URS) delineates the user's overall requirement regardless of solution (solution transparent). The system requirement specification (SRS) defines those parts of the user's requirement which the system will satisfy, again in solution-transparent terms.

System design results *inter alia* in performance and design requirements (PDRs) which specify, in detail, (but again solution transparent) the hardware, software, interface and performance parameters of the individual elements of the system. These PDRs are used as the basis for bidding, development and test. All of the specifications come together as the basis for customers' acceptance of the system.

The project may be thought of as moving in time from left to right on the Figure. Thus, operations and systems requirements analysis should be early transitory phases. Personnel may also move with the project by transferring between teams. In particular, those responsible for system design should, where practicable, prove their design in test and integration to the satisfaction of an independent authority provided within the overall management structure, usually the Test and Integration Manager with the Quality Assurance Manager.

6 Systems engineering activities

Management of system creation emphasises operational analysis, requirement analysis and system design and it is essential to provide objective traceable justification for system design. Most of the truly expensive mistakes are made early in a system project, which amply justifies the emphasis and requires that the system designer and manager have to be masters of a variety of skills, especially that of compromise.

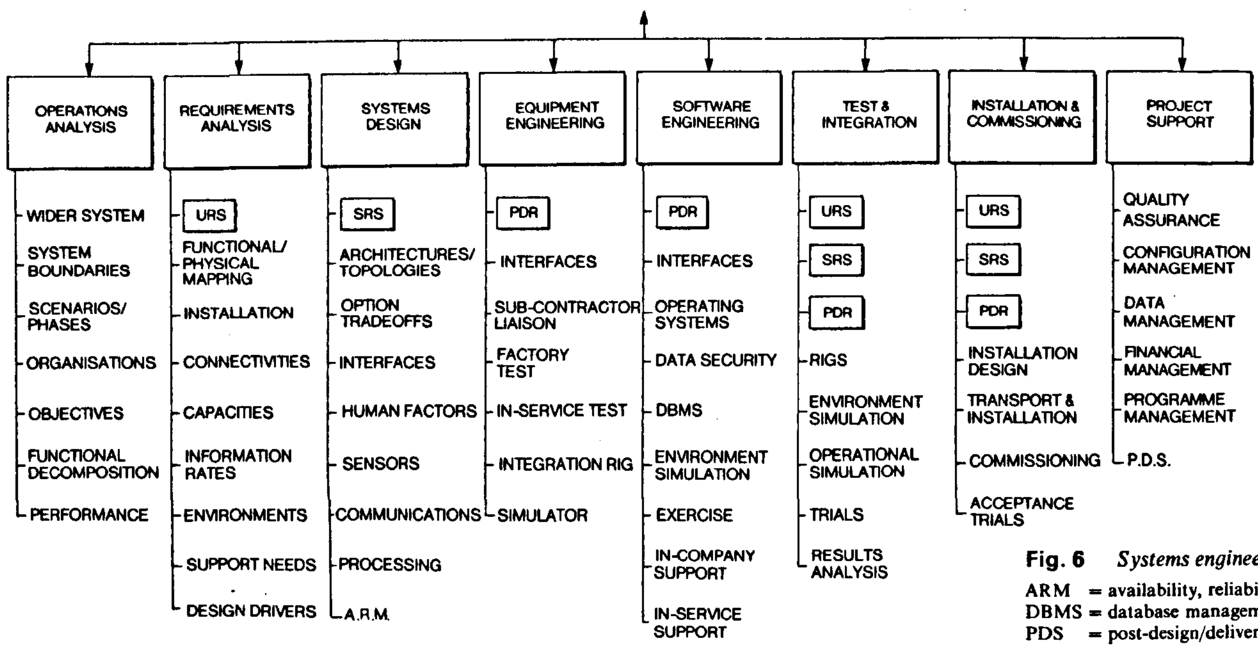


Fig. 6 *Systems engineering activities*

ARM = availability, reliability, maintainability
DBMS = database management system
PDS = post-design/delivery services

Fig. 6 shows a complete systems project structure with project support added to the seven chronological functions of the previous topic. Project support exists throughout the project life.

Operations analysis, requirement analysis and systems design can be seen as containing 'wider system' activities, and emphasis on these three areas is characteristic of effective systems creation. Following topics will highlight management methodologies in these areas particularly.

The activities shown under the various headings are generic as far as is practicable. However, the application of IDA systems requires highlighting of the human factors and information handling areas in particular. The system designer, in particular, has to master a variety of trades including:

(i) architecture analysis
(ii) computer science
(iii) control theory
(iv) human factors
(v) software structures
(vi) logistics and maintenance.

His principal skill must be in none of these areas, however; it must be in the ability to weigh and balance the many factors impinging on, and influencing, a design. The good system designer will not leap at the first solution he comes across. Instead he will:

(*a*) generate optional system designs which meet the system requirement
(*b*) compare each of the options against the required design characteristics
(*c*) select and justify the preferred solution.

Most of the major mistakes in system design are made at the start, hence the system management emphasis on operational analysis, requirements analysis and system design.

7 System tasks and skills

The work to be done by the systems functional teams can be logically 'decomposed' in stages. As an example, Fig. 7A shows stage one for operations/requirements analysis and systems design and Figs. 7B and 7C show stage two, the beginning of formal work breakdown, which highlights the system skills associated with each particular system.

The activity areas indicated in the preceding Section must be systematically structured so that the creation process is controlled and balanced. One technique for developing a logical workflow is the R-Net (requirement network) publicised (but not fully documented) by Robert Lano of TRW in a series of lectures he gave for the Technical Marketing Society of America in 1979, in London. Fig. 7A shows a high-level R-Net for operations and requirements analysis combined, and for system design. The R-Net is entered at the top and branches marked '&' must all be pursued, while branches marked '+' are alternatives; there are no '+' branches in these examples. The R-Net is much simpler than, say, a PERT Chart; items in parallel between two vertical &s can be executed together, serially or in any order. The Figure is concerned simply with their logical relationship.

Figs. 7B and 7C are also example R-Nets, but at the next level of detail. They too address operations/requirements analysis and system design, and are topographically mapped identically to the first Figure. Each block in the second-level Figures is marked with a work breakdown structure reference number which refers to a full description of the work to be done, skills required, resources required, outputs etc.

Examination of the R-Nets shows not only the tasks to be carried out, but also presents the skills which are particular to the system design process. The Systems Creation Manager is responsible for promoting and encouraging a dynamic, creative, even lateral thinking, environment during the early system design phases so that a full and balanced set of design options can be generated, reduced, massaged and detailed. Thereafter, the continuing gradual evolution of the preferred system design will require that the original concepts and tradeoffs be traceable, which, in turn, means that system design will persist at some level throughout the project's life.

Following Sections will highlight areas and methodologies of special relevance to system design.

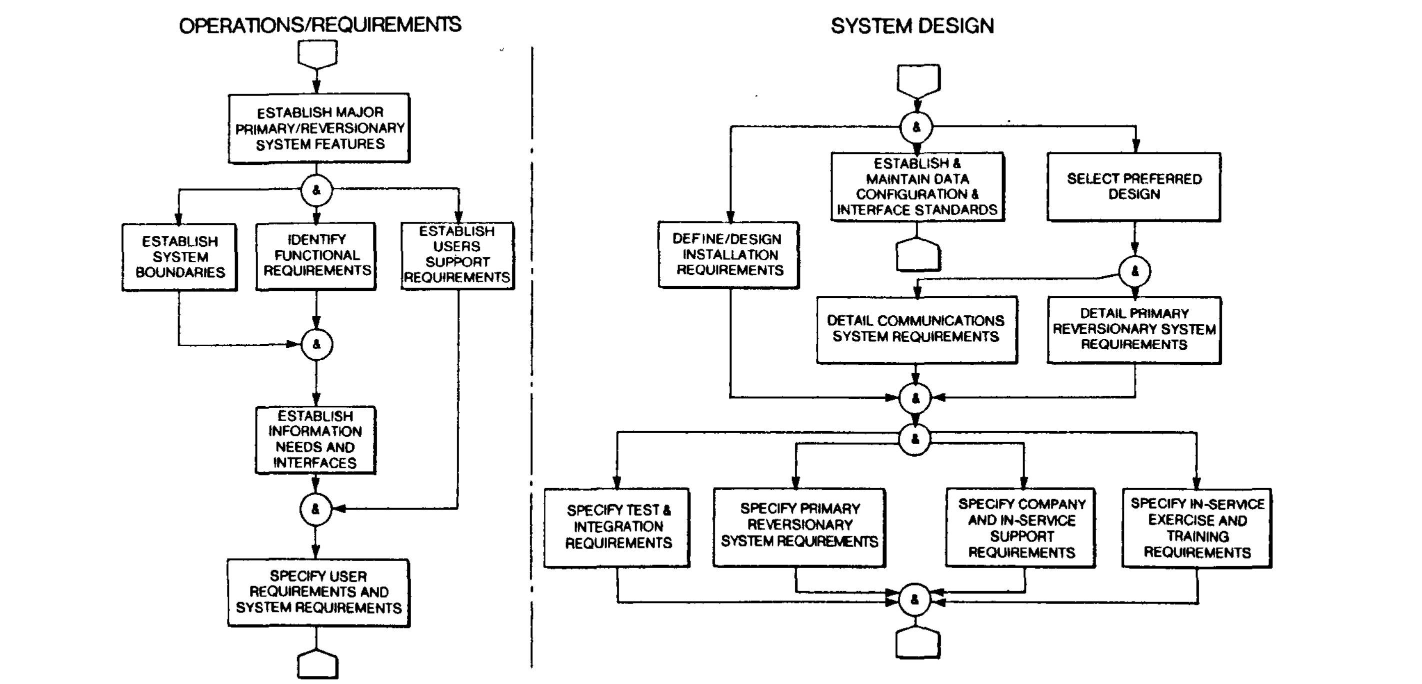


Fig. 7A *Tasks and skills: first level*

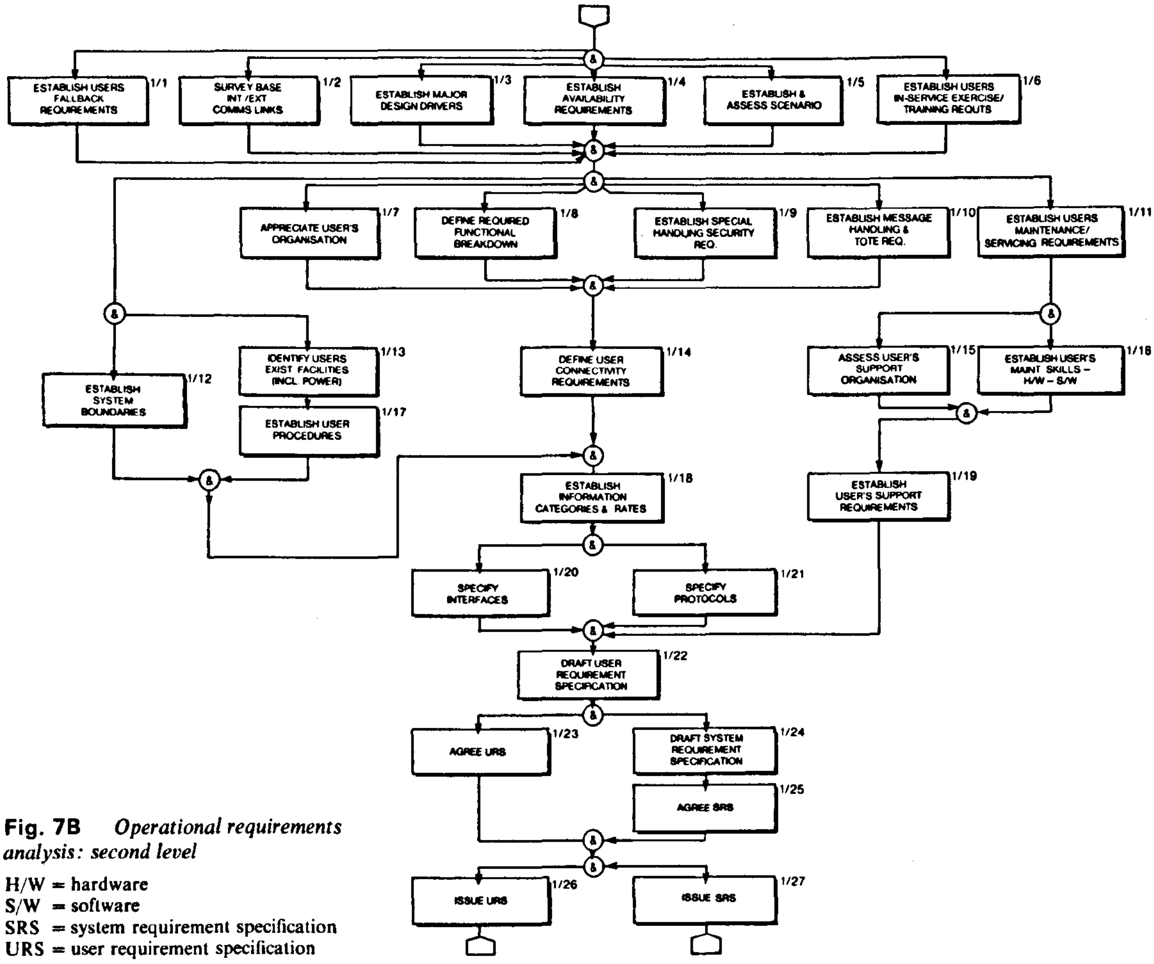


Fig. 7B *Operational requirements analysis: second level*

H/W = hardware
S/W = software
SRS = system requirement specification
URS = user requirement specification

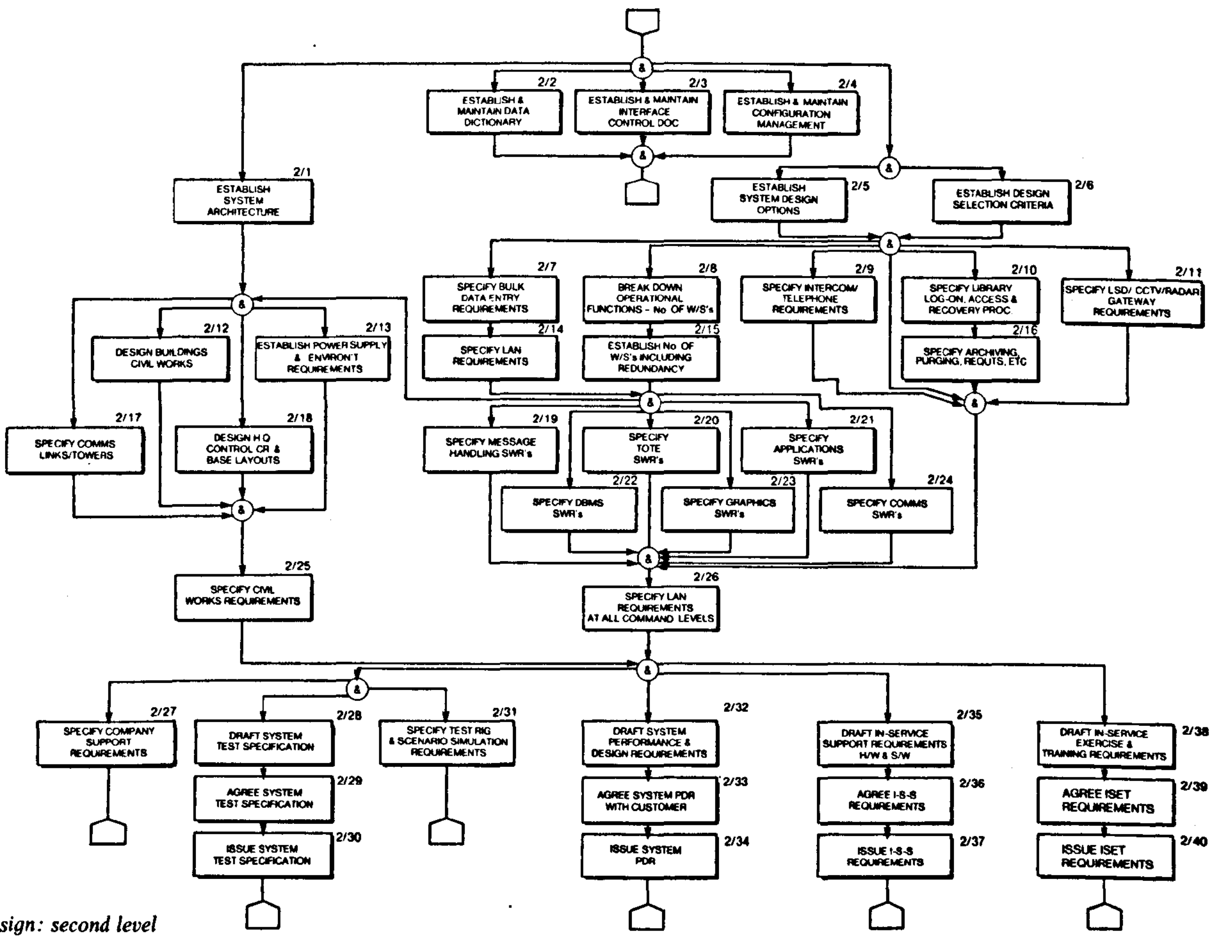


Fig. 7C *System design: second level*

HQ = headquarters
CCTV = closed-circuit television
DBMS = database management system
ISET = inservice exercise/training
ISS = inservice support
LAN = local-area network
LSD = large-screen display
PDR = performance & design requirement
SRS = system requirement specification
H/W = hardware
S/W = software
W/S = work station

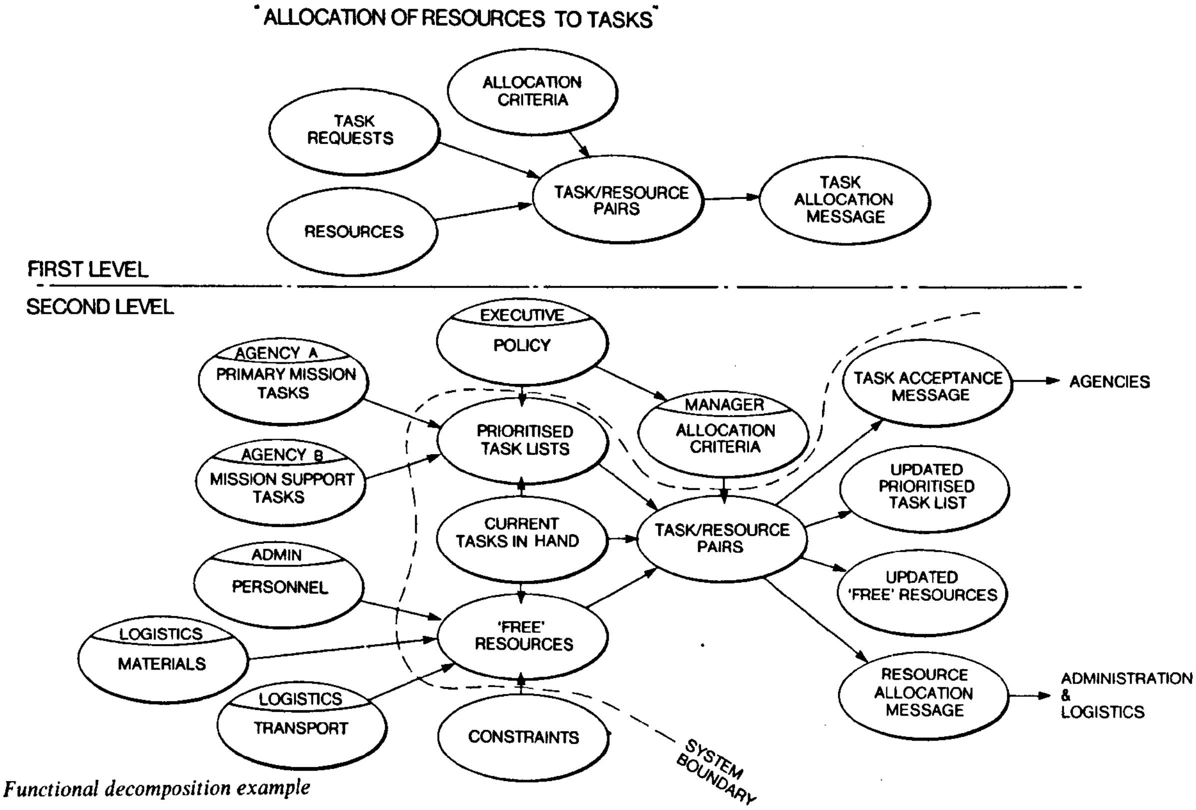


Fig. 8 *Functional decomposition example*

8 Functional decomposition

A common language is essential for effective dialogue between end-user and analyst/designer who must analyse the primary, reversionary and fallback systems on the same basis to ensure compatibility, and in a way that the end-user finds simple to understand.

Functional decomposition, part of operations and requirements analysis, is the systematic, progressive detailing of a function without regard to physical interfaces or other constraints. It is essentially a top-down creative process, which should proceed hand in hand with the potential end-user because it describes either his current tasks or future wishes.

As functional decomposition brings together end-user and analyst, a common language is needed which is at once easy to follow and quite specific.

The example shown in Fig. 8, data state design, presents data only in the bubbles and unspecified processes in the arrows. Successive detailing of a function, as shown for the simple case, 'allocation of resources to tasks', eventually leads to individual data pools which may represent, in some cases, database entities. However, it is important to realise that the IDA system requirement is comprised of only a part of each function; the dotted line separates out the system functions from the overall user requirement while keeping their relationship intact.

Fallback will probably require that each user function continues at some level, even when the end-system is 'down'; the data state design is also used as the basis for designing the reversionary (limited system performance) and fallback (no system performance) modes of operation so that these are compatible with the primary mode. Fallback in an IDA system generally consists of the operators working without part or all of the system. For fallback, they will need nonvolatile information during downtime. They will similarly need an alternative communication medium to receive and send volatile information. Fallback information system design may be layered to accommodate differing degrees and durations of main system failure.

9 Design drivers/option tradeoffs

Each IDA system is unique in realisation. Each design is influenced from the start of its creation by a variety of mutually inconsistent design criteria. An objective, visible and traceable means of comparing system-design options and balancing the many design influences is essential.

Each IDA system may have the same basic elements, but each is unique in realisation; the interfaces with the wider system and the characteristics of the particular requirement ensure that uniqueness. Classic systems approach requires the generation of objective measures of system excellence. Fig. 9 shows some of these principal design influences or design drivers for IDA systems.

Utility is concerned with performance of the primary, reversionary and fallback systems. Availability to the end-user is based on high system reliability and ease of maintenance. Adaptability, the staff of longevity for the system, is comprised of flexibility for evolution and expandability for growth. Interoperability is concerned with communication (in its true sense) between end-users in the wider system. Usability is the compatibility in design between the end-user in his working environment and the system. It includes the psychology of the end-user's perception of the wider system. Survivability for systems in harsh or military environments is a powerful design influence; it comprises avoidance of detection, self defence, when detected, and damage tolerance, when damage is sustained. Security is important for personal, financial, banking and military information systems.

Comparison of design options is a necessary feature of the systems approach. This is often undertaken using conventional weighting and scoring methods, which suffer from two principal deficiencies: the weightings are subjective, and the simple algebraic summation of weighted scores for different criteria is often invalid. The technique shown in Fig. 9 has been developed by the author to overcome these limitations; it is called rank matrix analysis (RMA). RMA ranks design options criterion by criterion and statistically analyses the resultant rank

Design drivers		Display system options (example only) *a* 1 alpha	*b* 2 alphas	*c* 1 alpha + graphics	*d* 1 graphics	*e* 2 graphics	Row sums
Utility	Performance	5	4	2	3	1	15
	Fallback/recovery	$4\frac{1}{2}$	2	2	$4\frac{1}{2}$	2	15
Availability	Reliability	4	2	3	5	1	15
	Maintainability						
Adaptability	Flexibility	5	4	2	3	1	15
	Expandability						
Interoperability	Communications						
	Compatible protocols						
Usability	Human factors	5	3	4	2	1	15
	Man–machine interface	3	4	5	1	2	15
Survivability	Avoidance of detection						
	Self defence						
	Damage tolerance	$4\frac{1}{2}$	3	2	$4\frac{1}{2}$	1	15
Security	Data						
	Physical						
	Rank sum	31	22	20	23	9	105
	Preferred solution	5th	3rd	2nd	4th	1st	

Coefficient of concordance = 0.5102
Probability of random occurrence < 1%

Fig. 9 *Options and tradeoffs*

matrix for significant divergence from random (null hypothesis). Ranking is often the only objective comparison possible between options at the early stage of design. In the example, optional operator display configurations are compared and the pattern is shown as very unlikely to be random. Option *e*, the preferred option, is undoubtedly the most expensive, but cost effectiveness must be judged system-wide and over the full life cycle, not piecemeal, subsystem element by element. In practice, RMA and conventional weighting-and-scoring methods are often best used together, and their results compared.

10 Architectures and interfaces

System performance, survivability and resilience are founded on the basic architecture, which is comprised of nodes and links. Design of a flexible architecture is at the heart of effective system creation, but is often dismissed as obvious or unimportant. The most far-reaching consequences can accrue from adopting a poor architecture at the start of system creation.

System structure, architecture, topology, etc. are the basis for eventual performance in design but are astonishingly,

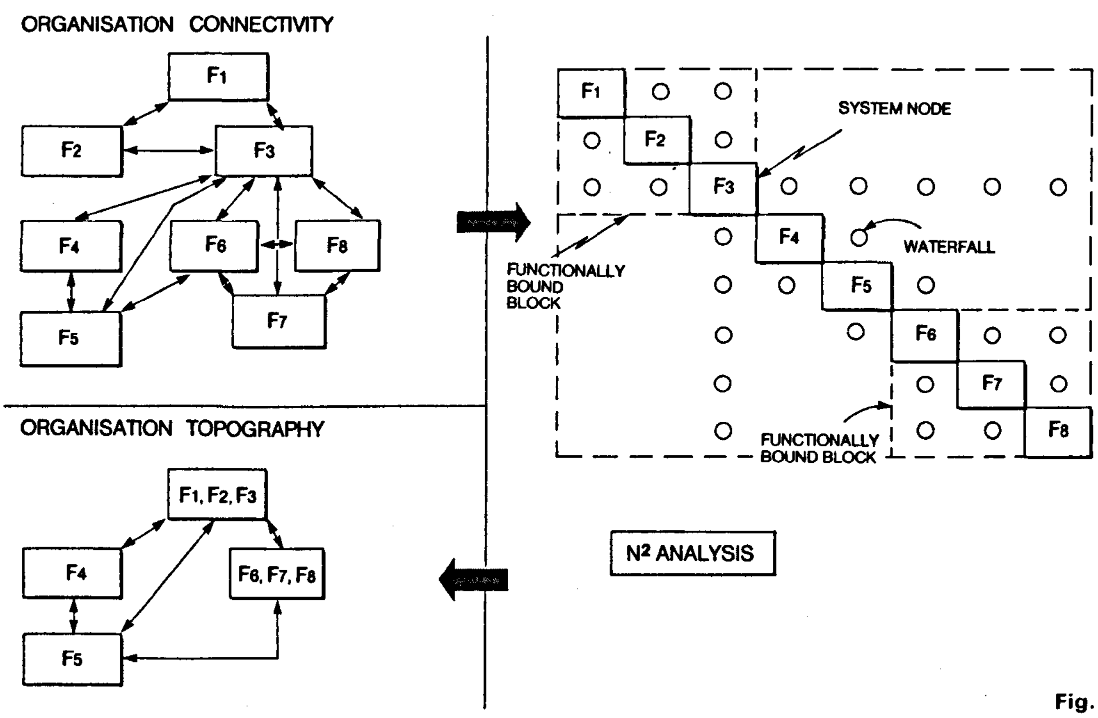


Fig. 10 *Architectures and interfaces*

often cursorily dismissed as 'obvious' or unimportant. Architecture is concerned with the identification and linking of nodes, nodal replication for survival and mobile operations, information replication for resilience, and the mapping of functionally bound subsystems on to physical architectures.

Top left of Fig. 11 shows a related set of entities. These could be people, organisations, software modules, distributed processors, The right-hand side of Fig. 11 shows the interfaces between these entities are represented in an N-squared chart [1] so-called because it comprises N rows by N columns giving N-squared rectangles. The leading diagonal represents the N entities, leaving N-squared-N interface rectangles. Each of these lies on the X-Y coordinates of two entities. Upper right rectangles represent 'down' interfaces in the hierarchy, while lower right rectangles represent 'up' interfaces.

The N-squared chart reveals three distinctive interface patterns in the organisation hierarchy. F_1, F_2 and F_3 have all mutual up and down interfaces represented; F_1, F_2 and F_3 comprise a tightly bound functional group, which would be a candidate for physical cohesion in design. F_6, F_7 and F_8 form a similar functional/physical group. F_3 is at the centre of an interface cross; it has interfaces to all entities and is a unique system node. Failure of F_3 isolates both halves of the organisation. F_3 is a candidate for reliability improvement, replication and/or protection. The third N-squared feature is the waterfall chain of command which shows orders 'rolling' down the organisation.

Finally, the topography at bottom left shows that the original hierarchy is essentially simple and highlights its underlying fundamental characteristics. Using the N-squared technique, particularly for large systems, permits the most complex architectures to be analysed, and their interfaces to be identified and controlled.

11 Interoperability

The advent of layered communication/information protocols promises to enhance performance and flexibility as well as interoperability. End-user compatibility may emerge as a new problem once the protocols are in place.

System design seeks to promote smooth continual flow. For production lines, the flow would be components modules and facilities contributing to the end product. For IDA systems information flow is the factor which transcends and unifies communication, processors, displays and operator interfaces.

Interoperability in IDA systems is essentially between human operators in the wider system community. A number of protocols is being developed which support this dialogue, the chief of these being (for IDA systems) the International Standards Organisation open-systems-interconnection (ISO OSI) seven-layer protocol shown in Fig. 11 [2]. This protocol consists of layers which buffer and isolate interoperability functions so that each layer interacts only with those immediately above and below it. The potential for evolution and growth is greatly enhanced by this 'damage limitation' layered protocol approach. The three lowest layers are concerned with the establishment of a networked communication facility. The transport layer provides an 'envelope' inside which information is sent over the network. 'Session' binds and unbinds groups of communicants. 'Presentation' formats, translates and presents information. 'Application' manages both the communication and information facilities in support of the end-user's particular activity.

As Fig. 11 shows, there are several layers needed above the seven-layer protocol to ensure that the human operators truly understand each other. The major obstacle to interoperability is not inability to understand — that can be detected and remedied; the principal risk arises from a mistaken belief that full understanding exists and that there is, therefore, no cause for concern.

It is not generally realised that the OSI concept can be very widely applied; it is not restricted to fixed-site computer-based transaction-processing systems but can readily extend to tactical, mobile real-time systems too.

12 Test integration and simulation

Test and integration require that the developing system be immersed in an environment to represent the extremes that might be met in practice by the end system, so that all system algorithms, capacities and performance can be established, tuned and proved. Simulation of scenarios, of nonavailable end-system components, and of system response will all be needed.

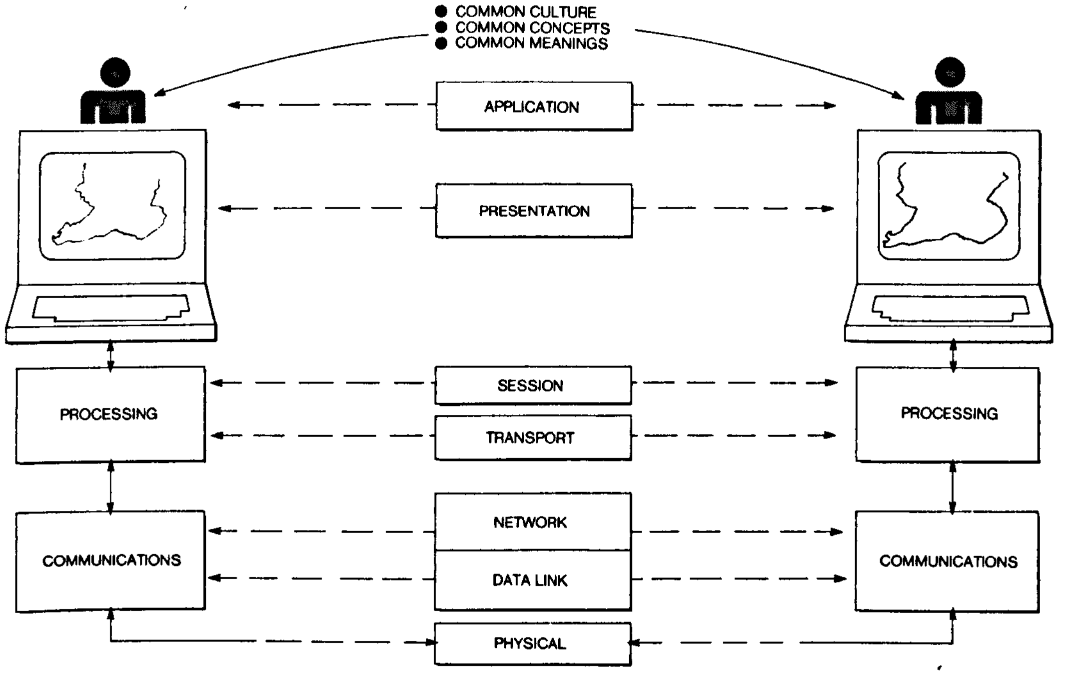


Fig. 11 *Interoperability*

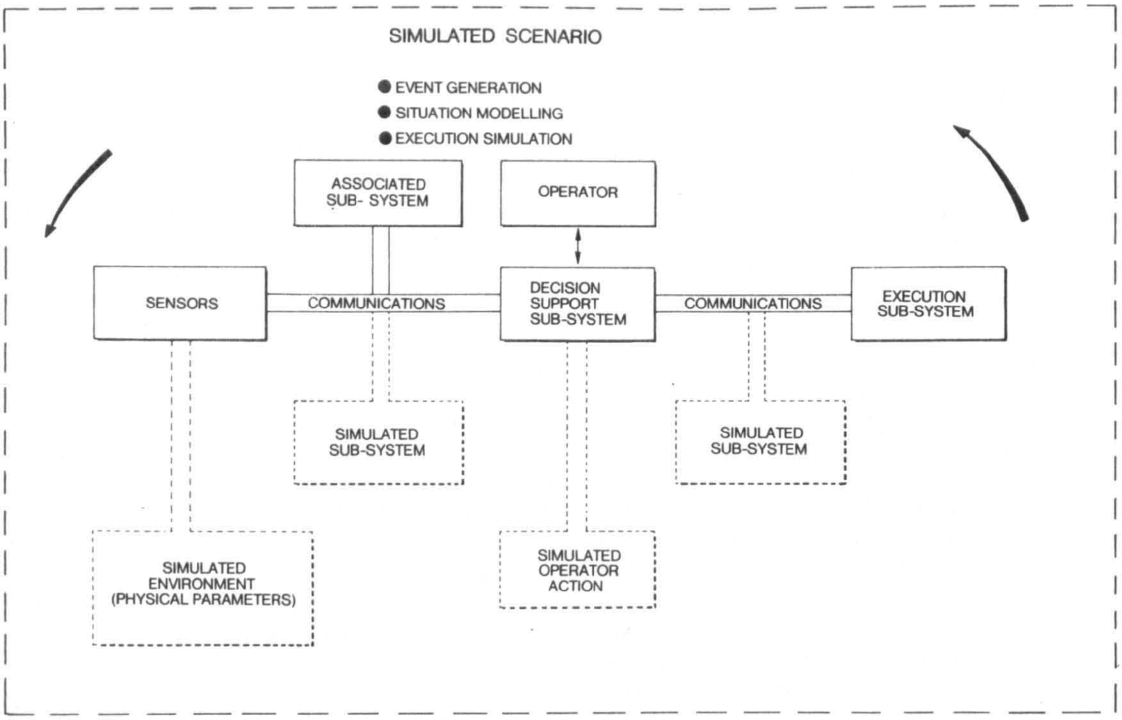


Fig. 12 *Test, integration and simulation*

The process of system creation will bring together hardware and software for test and integration. In any system of significance, the full testing of the system, man–machine interface, communications, software algorithms, etc. will necessitate the construction of a rig to represent parts of and interfaces to the wider system.

Progressive rig proving will require equipments and facilities to be interconnected, stimulated and analysed. There are three distinct elements to this process. Some equipments may not be easily obtained and their behaviour must be simulated. External events which activate system functions must also be artificially represented (e.g. radar returns to represent a situation which might only be met under extreme operational conditions). Finally, the response of the end-system which the rig represents must be simulated (e.g. the response of the eventual operator to the radar returns may change vehicle orientation and, hence, instruments and displays on the rig must be made to change too).

The process of simulating the local and wider environment surrounding the system under test can necessitate a set of simulation facilities which is more elegant in its way than the end-system itself. The development of naval and air real-time systems, in particular, requires elaborate simulation, which includes natural and man-made interference, garble and loss of information, simulated defects and the combination of situations and conditions most likely to 'break' the system.

13 System management organisation

Preceding Sections concentrated on the systems engineering aspects of system creation; the mangement of systems engineering is vitally important too, and will be highlighted in following Sections.

Unlike product companies which are physically partitioned in support of production, systems companies may be functionally partitioned around a project base. Progressive development of a systems company may lead through matrix management to the formation of a consultant's group.

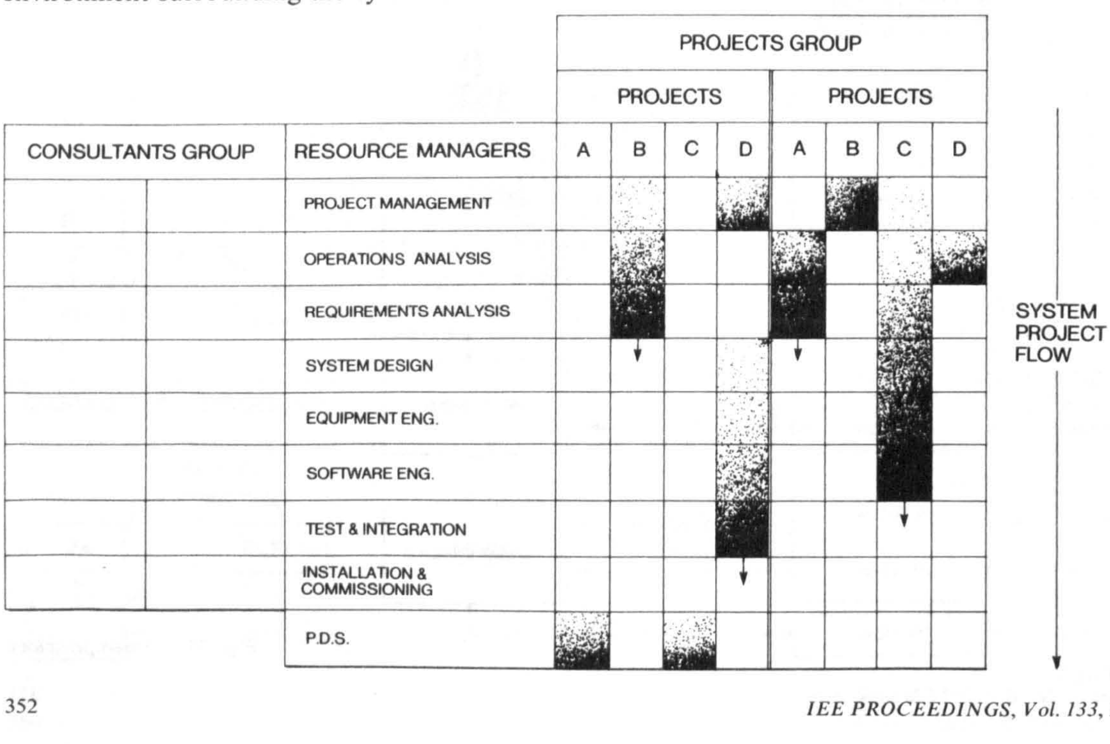


Fig. 13 *System management organisation*

Systems companies and product companies will generally be organised quite differently. For the product company, the production facility often represents a significant, and immobile investment in real estate and mechanical equipments. For the systems company, emphasis is on system-level design and software development, test and integration. As a result, product companies tend to be physically partitioned with goods inwards, stores, development laboratories, production areas, etc. Systems companies tend to be functionally partitioned around projects and groups of projects.

Fig. 13 shows a typical systems company organisation. On the right-hand side of the Figure a number of projects can be seen (*a*, *b*, *c*, etc.) at various stages of their progression through the company. The projects are grouped into two areas, probably according to customer category.

As systems companies grow, the projects mutually compete for resources. Fixed resources, such as the test and integration areas and software development facilities, may require separate managers. A matrix management organisation can grow as shown under the heading 'resource managers', with all project resources centrally managed.

As projects wax and wane some resources will temporarily become surplus. A good resource manager may 'sell' these surplus personnel resources of expertise as consultants outside the company, and the growth and eventual independence of a consultants group can be expected.

14 System project manning

System creation is presently labour-intensive, particularly in the areas of system design and software development. The various system creation activities tend to overlap in time, and it is to be expected that a considerable degree of work will continue during transition, post delivery.

The passage of a system creation project through a company has been presented, so far, as comprising sequential phases; in practice that may be an oversimplification.

Fig. 14 shows a typical set of manpower profiles for a systems project, with milestones along the top. The derivation of the requirement and the design can be seen, leading to equipment and, particularly, software engineering activity. The need for a design improvement is mooted, and it has consequences in equipment, software and test and integration areas. User training is shown during the development; end-users can usefully be incorporated into test and integration, both for their training and to smooth the path of customer acceptance of the system.

Noteworthy points include:

(*a*) the continuance of system design throughout the project

(*b*) the labour intensity of software development

(*c*) the overlap of the various functional activities

(*d*) the major requirements of the project postdelivery, particularly for inservice support facilities.

Attempts to cost major systems generally underestimate costs, often because the estimators believe that the system will be developed and completed according to a fixed schedule, and by a due date. In practice, this would be quite exceptional. The process of transition is not entirely within the control of the system-creation manager, who will nonetheless be responsible for tuning the system to meet the users' expectations and needs.

15 Transition

Transition is a greatly underrated phase of systems creation. Causes of slow or ineffective transition can be identified and ameliorated but essentially instant switchover from old to new system is an unreasonable expectation.

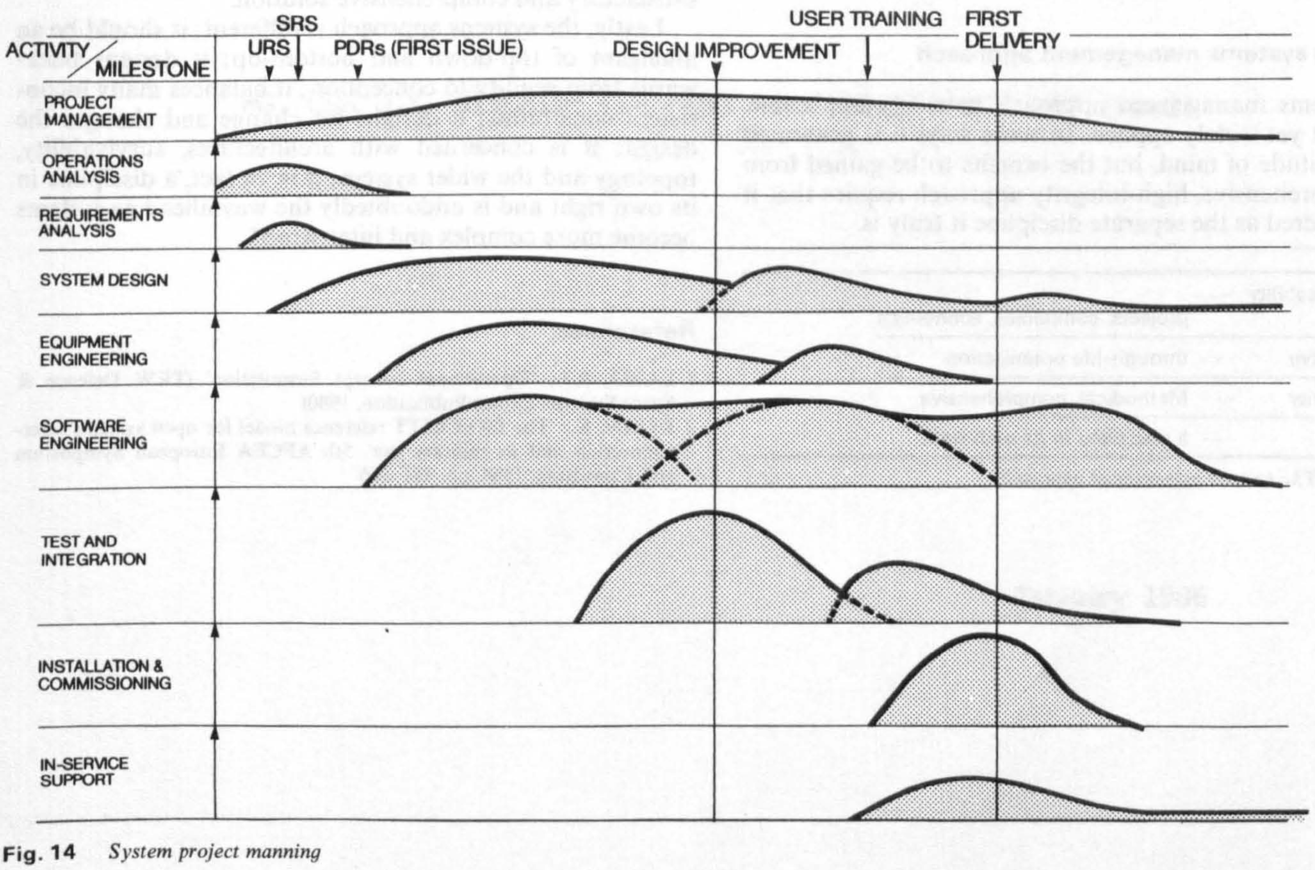


Fig. 14 *System project manning*

Transition factors	Probable cause(s)	Remedy
Slow transition	• Poor requirements analysis	• Discrete transition subsystem
Inoperative interfaces to 'wider' system	• Interface evolution	• Close control of interfaces through life
User resistance	• New system 'unattractive' • Advantages not evident	• User-friendly design • Co-operative implementation
Customer dissatisfaction	• Requirement has changed • False mental image	• Employ system flexibility • Early prototypes
Poor hardware reliability	• Infant mortalities/ • Poor design	• Design for reliability • Modular portable software
Poor software reliability	• Inadequate testing • Inadequate user trials	• Thorough tests & trials
Late support facilities	• Subordinate to primary system	• Emphasise cost of delay

Fig. 15 *Transition*

Transition has been highlighted in the preceding Sections; it is undoubtedly the most underrated phase of system creation. We expect to nurture a newborn infant, but a newborn system must be instantly capable.

Fig. 15 shows a number of reasons why instant capability will not occur. The potential remedies have a constant theme:

(*a*) continual attention to wider system interfaces
(*b*) attention to human factors at all levels
(*c*) co-operative user/creator design and evolution
(*d*) thorough test and trials prior to transition
(*e*) flexibility.

But, in essence, it is unreasonable to expect perfection. Instead, perhaps the systems creation process should indeed plan for a period of inservice nurture, with old and new systems running side by side and a gradual controlled transfer occurring from one to the other; such approaches have been tried with great success, but the desire for an instant 'switch' is hard to resist.

16 The systems management approach

The systems management approach is widely applicable, but is not yet widely applied. In some ways it is generated by an attitude of mind, but the benefits to be gained from this comprehensive, high-integrity approach require that it be considered as the separate discipline it truly is.

Wide applicability	--- projects, companies, economics
Cost effective	--- through-life optimisation
High integrity	--- Methodical, comprehensive
Unique	--- a discipline in its own right

Fig. 16 *The systems management approach*

The systems approach has almost universal applicability, but is not widely applied at present outside of the defence industry. Its applicability extends to development, production, administration, organisation and detailed design, although this volume has been limited principally to large IDA systems and management methodologies for large systems creation. The systems approach aims to be cost effective throughout the system life, but the pattern of expenditure differs from the bottom-up approach of selecting off-the-shelf, which is simple, tangible and generally unsatisfactory in realisation. It has also to be said that the systems approach can be misapplied, with generally expensive results; 'top-down' design can only be truly effective if it recognises the pragmatic limits of the 'bottom' towards which it is aiming.

Systems creation has to be based on integrity; it is a complex co-operative, systematic endeavour which requires continual customer involvement, plus a large element of forward thinking and planning to produce a satisfactory and comprehensive solution.

Lastly, the systems approach is different; it should be an amalgam of top-down and bottom-up; it designs backwards from senility to conception; it balances many inconsistent constraints; it designs for change and changes the design; it is concerned with architectures, survivability, topology and the wider system; it is, in fact, a discipline in its own right and is undoubtedly the way ahead as systems become more complex and intertwined.

References

1 LANO, R.J.: 'Operational concept formulation' (TRW Defence & Space Systems Group Publication, 1980)
2 KUHN, K.: 'The ISO/CCITT reference model for open systems interconnection and its military use'. 5th AFCEA European Symposium and Exposition, 1984, pp. 181–186

MOSAIC CONCEPTS FOR THE FUTURE

DEPLOYMENT OF AIR POWER IN

EUROPEAN NATO

BY

WG. CDR D.K. HITCHINS MSc CENG MIEE MRAeS AMBIM RAF (Retd)

January 1986

INDEX

MOSAIC - CONCEPTS FOR THE FUTURE DEPLOYMENT OF AIR POWER IN EUROPEAN NATO

1.0 INTRODUCTION

Air power in Europe is conventionally deployed in support of the anticipated Land Battle. Since World War 2 (WWII), the many roles of Air Power have developed somewhat independently; NADGE (Nato Air Defence Ground Environment) for example is the extant Air Defence (AD) element which draws the nations' AD forces together.

NATO is actively analysing the requirements for, and shape of, the future Air Command and Control System (ACCS) which is required to ameliorate deficiencies in present Command and Control (C2) of air power, particularly in the Central Region. These deficiencies are concerned with such fundamental issues as offensive/defensive integration, survivability, mobility and fratricide reduction.

Even as these early ACCS considerations are getting underway, technological advances are driving both the threat spectrum and threat counters to evolve at an ever-increasing rate. There is, therefore, a prima facie case for examining the underlying deployment of air power in these dynamic circumstances; otherwise, it may be that a future ACCS will be faced with the task of accommodating deployment deficiences rather than providing the essential Force Multiplier.

It is timely, then, to consider alternatives to current force deployment in European NATO, to assess those alternatives and to consider the impact that they might have on air power C2. This paper introduces 3 interlocking concepts which together present a radical option to current air power deployment conventions, which is consistent with threat and technology trends and which at the same time addresses some present day shortcomings. These so-called MOSAIC Concepts emphasise survivability and "moveability" as the key issues; integration of offence and defence is seen as a fundamental part of survivability, since neither perfect defence nor perfect offence is a viable military stance.

2.0 THE CHANGING EUROPEAN MILITARY SCENE

At the highest level Western Europe could find itself in one of the following states: Peace, Tension, Theatre Conventional War, Global Conventional War, Theatre Nuclear War or Global Nuclear War. Transition from one state to another could be triggered by a number of factors; some are shown in Figure 1, in which the six states are represented by boxes, and the causes of transition between states are at the commencement of arrows. For example, Menace causes transition from Peace to Tension; Invasion escalates Tension into Theatre Conventional War, and so on. Of particular interest is the basis for transition from Theatre Conventional, to Theatre Nuclear War; it is shown on Figure 1 as Rollback.

Rollback, orderly withdrawal by NATO ground forces in the face of superior odds, is designed to provide sufficient time for a political solution to be set in place. Should Rollback be far too fast, or the political solution fail to materialise, then it is possible that NATO would be obliged to use nuclear weapons to stem the tide. For Rollback to be too fast implies that key land features or installations would be forfeit. Amongst those key installations may be included fixed air power facilities: radars, Control and Reporting Centres (CRC's), Sector Operations Centres (SOC's) runways, fuel and ammunition dumps. It therefore follows that forward siting of fixed or immobile air assets will contribute to the earlier need during Rollback to resort to non-conventional weapons and hence to a lowering of the nuclear threshold.

With a swing in the tide of conflict, NATO forces might wish to pursue retreating forces into their own territory. Immobile air assets would limit pursuit effectiveness in this instance, by becoming progressively further behind the line of battle. Fixed air assets clearly have limitations when supporting a mobile war. On the other hand, the growth of such fixed assets has not seemed unreasonable over a period of thirty years of peace; static facilities have been developed, utilised, superseded and discarded with no recourse to mobility other than perhaps the occasional exercise.

The advent of new systems into the military inventories of both sides renders any complacency ill-advised. Improvements in tactical weapons delivery accuracies and the use of spaceborne sensors and weapons are two factors of relevance to the fabric of future Air Power in European NATO, and both President Reagan's Strategic Defence Initiative and Chairman Gorbachev's move to eliminate nuclear weapons may result in fundamental changes in future NATO strategy. The only certainty seems likely to be one of change; for Air Power this implies flexibility, mobility and survivability as being key issues.

3.0 AIR POWER IN CONTEXT

The concept of indivisibility of air power is writ large upon the minds of many air forces. Originally, the concept argued against naval-air, army-air, and independent air forces such as those of the UK during WWI. Indivisibility also refers to the concept that offence, defence and support cannot sensibly be considered in mutual isolation. The original concept is not acknowledged worldwide; navies and armies often have their own air forces, and there are cogent arguments in favour of such specialist and available air support. The second concept is more relevant to the discussion and will be expounded below.

Figure 2 shows Air Operations as comprising three major arms; offence, defence and support. Some of the activities and characteristics of each arm are also shown. Examination of the characteristics shows:

- Air Defence shields vulnerable Offensive and Support Air as well as itself.
- Air Defence is relatively short range and presently tied predominantly to fixed Air Defence Ground Environment (ADGE) facilities.

- o Hence overall flexibility and high speed mobility are hampered by the relative immobility of Defensive Air, which is only partly overcome by the use of point defence weapons such as RAPIER and JAVELIN during bridgehead operations.

There are, therefore, grounds for re-examining closer integration between the elements of air power. If the objective is to enhance the overall effectiveness of air power then, as Figure 3 shows, a systems approach is needed. The figure shows the same three major air power divisions: (1) Offensive (2) Defensive and (3) Support Operations. Optimising each of these divisions individually need not result in an optimally effective air arm overall; an overall objective must be established for optimisation of Air Power as an entity.

Integrating all the elements of air power would be impracticable in most circumstances, however. Figure 4 shows four classifications of operations, Offensive, Defensive, Maritime and Air Transport, with major activities peculiar to each class. Maritime Operations, with its Long Range Maritime Patrols (LRMP), Anti-Submarine Warfare (ASW) and Tactical Aid in Support of Maritime Operations (TASMO) is clearly set aside from the others in its roles, its tasks and its sphere of operations. Closer examination reveals that each of the others is also independent, and requires not only its own aircraft and weapons, but discrete procedures too; each attracts different personalities and characteristics amongst the air crew. It can also be seen that function grouping:

- o Concentrates and develops specialist skills.
- o Affords a simple organisation.
- o Enhances motivation and completion.
- o Allows simple measures of effectiveness by function.
- o Divides Capital funding into defensive and offensive.

The last item proffers a two-edged sword. Since the second world war, several periods of detente have arisen; during such periods it is politically more acceptable to spend money on defence than on offence. Spending on defensive facilities thus becomes continual and defence can be developed and optimised in isolation from its spasmodically-funded offensive counterpart. Taken to its limit this process lowers survivability both of air power and the land/sea battle it supports, by placing undue reliance on a defensive potential which cannot sustain itself in the face of continual attack.

The interrelation between the various elements of air power is illustrated in Figure 5, which is a System Dynamics view of air operations conducted in the face of an advancing enemy ground threat which is threatening to overrun forward air assets on the ground. In the so-called Influence Diagram technique, a 'plus' or 'minus' sign at the arrowhead signifies that source of the arrow will enhance or diminish the focal point in question. Thus Successful Air Defence Engagements (top centre) reduce (minus) the Advancing Enemy Ground Threat (bottom centre) by improving conditions for Air Superiority Operations. In the figure, enemy Command and Control (C2), radar and bases are at the right. The very simplicity of this abstract representation of air operations highlights the importance of:

- Balanced offensive and defensive operations.
- Common and dynamic C2 where offensive and defensive effort and resources can be co-ordinated.
- Mobility for C2 and Radar units, bases and depots.
- Fratricide Reduction.

Clearly any future deployment solution should recognise the close interrelation between areas of air power, while tempering the move to full integration at execution/operations level with the need to preserve specialisations.

4.0 THE THREAT FACING AIR COMMAND AND CONTROL

Advances in C2 are threat driven. The threat in the future air/land battle is developing swiftly; tactical threat developments can be categorised simply under three headings.

- Launch vehicle
- Delivery vehicle
- Warhead, weapon and payload

Launch vehicles include tactical ballistic missiles (TBMs) and cruise missiles which may be expected from any elevation. In the time frame of ACCS we must expect spaceborne weapons too; these could take many forms and arrive from any direction.

'Conventional' warheads will include chemical and possibly biological payloads as well as high explosive devices, and the variety of nuclear warheads is well catalogued.

Two very significant threat trends concern the detectability of threats; voluntary-emission control (V-EMCON) and Stealth. V-EMCON reduces or eliminates the detectable voluntary emission including not only communications, but also Doppler radars, radio altimeters and radiating DMEs (Distance Measuring Equipments). Stealth seeks to reduce radar cross-sectional area, and encourages V-EMCON, and the use of terrain to screen essential transmissions such as those from Terrain Following Radars (TFRs). For entirely pre-planned missions even TFR and radio altimeter emissions may be unnecessary in the foreseeable future.

There are clear pressures on the future ACCS and its associated subsystems, including sensors, communications and weapons if the aims of ACCS are to be realised:

- Increased vulnerability to new all round threats.
- Need to operate in a degraded mode after sustaining damage, possibly from nuclear chemical or biological weapons, and in the future from beam energy weapons.
- Reduction in attack warning time, leading to increased risk of fratricide, and to reduce defensive weapon effective range.

This last point is expanded at Figure 6 which has two parts. At the left is a graph showing target height against an approximate number of ground radars needed to provide contiguous radar cover in a 300 km-square as shown in plan at the right. As target height reduces, the number of radars rises towards infinity, classically pursuing the law of diminishing returns. Advances in terrain following aircraft performance are continuing to reduce operational heights, and terrain screening can reduce pick up range even more dramatically. The plan at the right of Figure 6 shows a notional detection range of 25 km as requiring 36 radars to cover the 300 km square. The resulting 50 km squares form columns and correspond to about the expected daily WP rate of ground force advance. Thus, in the given simple scenario, either six ground radars (one column) would be overrun per day, or (if mobile) they would move swiftly to the rear and redeploy. Simple although Figure 6 may be, it amply illustrates two important points; the usefulness of ground radars is being restricted by enemy tactics, and those radars near the FEBA/FLOT should be mobile.

Threat evolution is driven by the same technology which will provide the counters, although a time delay may arise between development of threat and of counter. Figure 7 illustrates the balance in C3 trends, with the air threat at left. Remaining columns show selected Sensor, Communication and C2 improvements fuelled by recent technological improvements. Taken together these improvements accumulate increased overall effectiveness. Command and Control is of and by military personnel, however; major contributions to improved effectiveness will come from improvements in organisation, co-ordination and procedures which best employ and protect limited resources.

5.0 AIR COMMAND AND CONTROL DEFICIENCIES

Deficiencies in the current command and control of air power in European NATO have been identified, and may be summarised as follows:

- Lack of timely, secure and adequate information leading to an inability to derive timely and high quality command decisions.
- Lack of co-ordination between all elements of air power, but especially between offensive and defensive air.
- Vulnerability of sensors, communications, processing, HQs and C2 personnel to physical and electronic attack.
- An unacceptable fratricide rate due principally to the lack of high-integrity identification and weapons control.

In the light of threat developments, it can also be seen that the very nature of our air power deployment in European NATO is suspect, principally because:

- Facilities (sensors, communications, airfields) are almost entirely static.
- Air defences assume that air threats come from the East and are oriented accordingly.
- Current hierarchical C2 organisations contain nodes and are not designed to tolerate damage.

Following sections introduce MOSAIC which specifically aims to improve Survivability and Mobility but which also addresses the central issue of Fratricide.

6.0 INTRODUCTION TO MOSAIC

6.1 Air Power Deployment Options

There are several ways in which air power may be tactically deployed. Figure 8 illustrates the terms and concepts employed here. Many forces employ layered air defence, in which offensive air power operates behind the relative safety or discrete defensive layers, comprising, for example, belts of Surface to Air Missiles (SAM), long range interceptors, and short range interceptors. Layered air defence is perhaps at its most practical in naval use where the battle group presents a concentrated target. In Central Region, the targets to be defended are widely dispersed and a 360 degrees azimuth defensive layering would not be practicable due to size and cost. The layered system therefore covers only an arc, facing into the ground threat axis.

Laned defence (top right) presents lanes in the direction of the threat axis and is of use principally where there is no significant mix of weapon systems. Long range interceptors, for example might operate furthest from home, while shorter range interceptors operated nearer the safe area from where offensive and support air must operate. Laned defence can be integrated with layered defence by making layers contain lanes.

Area and overlapping point defences do not presume a threat axis. Area defence covers vital points as sets within an area, while overlapping point, as the name implies, concentrates on point defence. In terms of concentration of force, point defence must be attractive when resources are limited and threat direction unknown, but of course it requires definition of what points are considered vital, which may prove to be a far-from-trivial problem. MOSAIC concepts employ an area and point defence philosophy to protect vulnerable air assets on the ground from any direction.

6.2 The MOSAIC Concepts

MOSAIC considers the evolving threat and the current C^3I shortfalls and takes these various influences to their natural conclusion in reaching a design for ACCS. MOSAIC is an acronym, representing: Moveable Semi-Autonomous Integrated Cells.

Each of the words is separately important to the concepts, and the whole acronym is also appropriate. The aim of the MOSAIC concepts, of which there are three, is to greatly enhance survivability by selecting organisational structures which will continue to operate effectively in degraded mode. To achieve this, the structures must be:

- Moveable - either mobile or transportable - so that they may avoid targetting and overrun.
- Semi-Autonomous - so that loss of Superior control can be accommodated by co-operation laterally with other groups in the structure and, in extremis, total isolation can be accommodated for a period.

- Integrated - in two senses; first to contain offensive, defensive and support elements in each group and so constitute a complete, highly survivable, fighting unit; second, so that the groups, although having the potential to operate semi-autonomously or even autonomously in degraded mode, normally act as one under superior control.
- Cellular - so that individual cells which, by definition are moveable, may be grouped together - interfaced - swiftly, easily and in various ways to reconfigure the superstructure.

6.3 The Ground Environment MOSAIC (GEM)

Figure 9 illustrates the GEM concept which addresses offensive / defensive integration and survivability directly by providing self contained cells and by recognising that neither attack nor defence is, of itself, a tenable military role. Each MOSAIC, containing its indigenous air defence, offensive air and air support, can be redeployed. At Figure 9A ten cells form a layered MOSAIC which corresponds to conventional layered air defence except that offensive air and support air are not sheltering in a safe area, but form part of the layer. Figure 9B moves three cells to accommodate changes in ground threat axis, and 9C shows area and overlapping point configurations, remembering that since each cell has its own 360° air defence, the resulting MOSAIC also offers omni-directional air defence. Figures 9D and 9E show the flexibility needed to fall back in the face of advancing ground forces and to configure killing zones if such were appropriate.

The concept of cell mobility to reconfigure tactical deployment clearly has potential for providing second echelon support to front echelons experiencing undue pressure. Survivability too is potentially very greatly enhanced; any one cell remaining after an attack has the ability to move and to keep fighting. Sheltering of offensive air behind a defensive screen has been superseded for two reasons.

- The threat axis for air attack will not be known in the future, and certainly need not coincide with the main ground thrust so that point/area defence of offensive assets is the only viable option.
- A rapid WP attack which broke through a conventional, layered defensive screen would find offensive assets both concentrated and vulnerable.

(Other aspects of GEM and the two remaining concepts will be expanded after all three concepts have been introduced).

6.4 The Sensor MOSAIC (SEM)

The SEM concept is illustrated at Figure 10, where it is compared with current philosophies at top left. The SEM concept concentrates on freeing air C2 from its ground fetters and on integrating target data from airborne sensors.

Current C2 for air power is strongly ground centred in (largely) static facilities. Ground sensors are linked together to form an integrated sensor picture, to which it is hoped to add the radar pictures from NAEW, AWACS or perhaps aerostats (tethered balloons) or airships. This ground centred concept must be doubtful for the following reasons:

- The ground based sensors will see less and less with the advent of V-EMCON, Stealth, TBM and Cruise Missiles.
- Static ground sensors and their equally static intercommunication links will be early targets.
- NAEW radar and the links with ground C2 will be prime targets for electronic and physical attack (with ARM for example).

The SEM concept recognises the introduction of track-while-scan radars and ECM Resistant Communications Systems (ERCS) such as JTIDS into airborne service. Figure 10 also shows that the excellent lookdown capability of these radars forms a patchwork, or MOSAIC of coverage, with some valuable system qualities when compared with ground based, static sensors:

- There is a number of discrete sensors, each switching on and off as the crews pursue their operational roles.
- The density of such sensors increases near the air battle.
- Distances between sensors is relatively small near the battle, reducing the effects of jamming, especially on the ERCS.
- The sensor population is dynamically stable as fighters go off-station and are replaced.
- The sensors follow the air battle, unlike their ground based counterparts.

For the SEM concept to be realised, the various sensor inputs must be brought together. The most likely candidate to undertake the integration is the NAEW/AWACS which, because it may be out of contact with the ground environment, assumes almost implicitly the role of Air Action Group (AAG) co-ordinator. Taking this notion to its logical conclusion, we see that the Sensor MOSAIC concept can free air power from the ground shackles which currently restrain it and limit its essential flexibility. The AAG is analogous to the naval Surface Action Group and in like manner it may have offensive as well as defensive elements, may move with the action, and may acquire temporary autonomy when conditions dictate.

6.5 The C3 MOSAIC

The Ground Environment MOSAIC and the Sensor MOSAIC may enhance flexibility, mobility and survivability, but they clearly present a challenging C2 task and the corresponding communications facilities will require careful analysis.

These aspects are the subject of the third concept, the C3 MOSAIC which is most directly relevant to current ACCS endeavours. The C3 MOSAIC has the essential task of binding and controlling the moveable GEMs and the AAGs.

Figure 11 introduces the concept, showing a three-tiered structure with cell C2 controlling its own defence, offence and support and reporting to its respective Cell Group C2. These Group facilities control a varying number and population of cells during mobile operations, and are themselves capable of moving their defence, offence and support facilities. Above Cell Group level there may be either a further grouping or direct reference to the MSC (Major Subordinate Commander).

Certain salient features become apparent from Figure 11:

- o The functional groups within each level have substantially similar divisions into operations, intelligence, plans, logistics and engineering;
- o Within a base, cell or cell group control there will be continual interchange between the functional groups; this implies tight functional binding.
- o Between cells and between levels, however, communication is specifically coupled by function; operations will speak externally to operations (as shown), intelligence to intelligence, but not intelligence to engineering or operations to logistics. (This and the previous features are common to all sophisticated C2 organisations).
- o The sensor MOSAIC requires a number of line of sight air/ground communication gateways to accommodate AAG mobility.
- o The communications structure allows one cell group to control another's cells, and also allows a cell to control an adjacent cells base facilities should the need arise (battle damage, mobile operations etc..)
- o In providing this cross coupled communication system, the emphasis moves away from vertical communication in support of a simple command hierarchy towards a 'rectangular network' of vertical and lateral communications, with the communications channels group by function (operations, intelligence, etc..)

7.0 EXPANDING THE MOSAIC CONCEPTS

The conceptual issues raised by MOSAIC are fundamental and if fully embraced would impinge on almost every aspect of air power. Before assessing MOSAIC, it may be useful to consider some of the key implications of the three interlocking MOSAICs:

- o Moveability of cells, cells controls, weapon systems, resources and support infrastructure.
- o Interopability between cells and their controls.
- o Communications and interfaces to permit and enable moveability and interopability even post nuclear exchange if necessary.

7.1 Moveability

Moveability (i.e. mobility and/or transportability) of ground-based sensors is not a new concept, but it does have design implications. The need for enhanced performance, for example from antennae, must be reconciled with ease of breakdown and transportation. Similarly the size of CRC/SOCs in terms of personnel facilities, power supplies, hardening, air filtration, etc., must be limited by moveability dictates. These problems are generally understood. Less well understood perhaps are the implications of moving complete squadrons of aircraft with their supporting infrastructure, air traffic control, aircraft protection, recovery facilities, fuel, weapons, avionics maintenance, etc.. There are however ample precedents in the several ways employed by armies and, particularly, by navies.

Navies carry their total air power infrastructure on board ship, including runways, surveillance, air traffic control radar, recovery and support. Their aircraft are adapted to the purpose but, as aircraft such as the Sea Harrier, F14, F18 and E2C Hawkeye illustrate, the process of adaptation need not inhibit performance. In a MOSAIC-ACCS, aircraft might be similarly adapted.

One approach to air power moveability, therefore, is to emulate the navy approach, at least in part. The transporting of runways is not out of the question, using special track laying vehicles and adapted aircraft, but there are many extant airfields which could be used, and autobahns have already been considered for some types of aircraft. There are also VSTOL aircraft designs for the future which could ease the problem further. Alternatively, the armies of the world have also been mobile since warfare began, and their approach is not dissimilar to MOSAIC in many respects. The use of step-up Headquarters, so that one HQ may operate statically whilst the second is on the move, in leapfrog fashion, is tried and trusted, and could be considered for some air force applications, although it does carry penalties of increased manning.

Looking at navy and army moveability in this way, it becomes apparent that MOSAIC seeks to bring air power moveability into line with the capabilities already present in navy and army air power, and in so doing impinges on the design of ground sensors, HQs and aircraft. Paradoxically, mobility may reduce manning levels to accommodate transport requirements. Air Power composition seems likely to rebalance to meet the moveability requirement: more short/medium range STOL transports would be necessary to move cells rapidly during Rollback or advance, when road/rail links might be choked or severed. MOSAIC also encourages the development of multi-role fighter aircraft and aircrew; at present such aircraft change between roles with relative difficulty, and the aircrew often find the change even more difficult. In MOSAIC, the emphasis on combined offence and defence within each cell would be sensibly accompanied by a rationalisation of spares and weapons to reduce the logistics transportation burden, and by an ability of each fighter type/aircrew to perform a second role. This follows logically from the fundamental objective of survivability; if, in extremis, the cell must be able to function, then too as crews and aircraft are lost, remaining crews and aircraft must have the flexibility to attack, defend and redeploy almost at will.

7.2 Interoperability

Interoperability has many meanings in practice in addition to the interoperation between facilities in different nations. The MOSAIC concepts highlight a current problem in interoperability which is illustrated by Figure 12. Ideally, and subject to security protocols, Operator 'A' should be able to access not only his own parochial information but also that of his counterpart Operator 'B' in another cell, perhaps in another region. One operator could be in an aircraft, another in a mobile HQ.

It might be possible in principle to oblige both operators to store and display information identically, but in practice, due to language, operational and personal differences, each operator will tend to prefer his own display formats. Thus Operator A in viewing a file from Operator B's database will expect the display to appear as Operator A's other files appear. For example, an aircraft readiness tote from a remote region should be formatted identically with a local aircraft readiness tote.

As Figure 12 shows, this objective requires many factors to be in place:

- o An ACCS-wide data dictionary.
- o ACCS-wide communications protocols, including file transfer protocols, such as the ISO Open Systems Interconnection (OSI) Seven layer protocol.
- o Agreed information requirements for operators on an operational function-by-function basis.
- o Common or compatible display standards.
- o A common understanding of the words and symbols transferred between systems.

Even then, difficulties can be foreseen, since operators 'A' and 'B' may wish to discuss the different displays of identical information which they respectively view. This problem is likely to be resolved only by common training and practice.

The need for standard protocols and has been mentioned and is as obvious as it is difficult to achieve. One concept which would bind the various communications bearers together in underpinning a common theme is that of an ACCS-wide Tactical Data Interchange Language, which may be called TADIL-ACCS. We already have TADIL A, TADIL C and TADIL J in association with NATO LINKS 4, 11 and 16. If the valuable TADIL concept is carried to its logical limit, then the problems of interoperability and moveability would be considerably eased. The TADIL-ACCS concept is not closely coupled with the communications bearers; figure 13 illustrated the point, with a TADIL ACCS track message being formulated by an airborne radar and being borne over a variety of bearers via NAEW, air/ground gateways and ground to ground links before initiating a track display.

The TADIL-ACCS philosophy thus engenders:

- o TADIL-ACCS based subsystem designs in aircraft and ships such that messages are formulated at source (button press) in TADIL-ACCS.
- o TADIL-ACCS based processing and display systems on the ground.
- o TADIL-ACCS bearer systems, such as MIDS, LINK 11, LINK 1, etc..

For full mobility in the MOSAIC sense, TADIL-ACCS addresses only part of the problem, the real-time operational task. There will be a need for a parallel data interchange for Intelligence, Logistics, Engineering and Plans as well as for operations; each function will require an Information Interchange Language. There maybe, therefore, a need for LOGDIL for logistics data; ENGDIL for engineering data and INTDIL for intelligence data to provide interoperability at all C2 infrastructure levels. With such a set of languages and with compatible databases for each cell, the prospect of moving a cell and "plugging-in" at a new site becomes realisable.

7.3 Communications

Communications to support the MOSAIC concept must be battle damage tolerant and yet have sufficient capacity and flexibility to accommodate Cell moveability and the full bandwidth of communication in peace time.

At present, voice communications, while moving towards digital modulation, are nonetheless diverse as Figure 14 illustrates, and the consequent need for translation between systems is very real now, and would be exacerbated by MOSAIC.

Communications survivability can be tackled in a variety of ways, including:

- o Hardening
- o Non-nodal organisation
- o Multi-nodal organisation

For MOSAIC, hardening must be reconciled with mobility. Non-nodal communications such as the TDMA variety of JTIDS would prove essential to MOSAIC for air-to-air and air-to-ground communications survivability. The situation on the ground is the most interesting however. At present, the commercial pressures for communications, are causing PSTN to burgeon and to adopt fibre optics for cost ease of maintenance and capacity reasons. This closely woven communications backcloth will, in the near future, offer a ready made solution for ground to ground bearers, provided of course that the military can either adopt or adapt the CCITT protocols. Last, but not least, HF radio communications would be a vital link in the communications chain, especially post nuclear exchange. In particular, ground wave may be the only effective communication in emergency, and as such would be core to the MOSAIC concept for survivability.

To summarise, Figure 15 shows the communications interface complexity of MOSAIC and emphasises the need for systematic control on an ACCS-wide basis. As the figure shows, there are sets of interconnected functional blocks which appear as 'tiles' in a MOSAIC.

- o PSC and Cell Controls form a command group.
- o Cell Control, Cells (and the SOC/CRCs inside the cells) form a CIS Functional Block for near-real time control.
- o SOC/CRC and NAEW form a real-time control.
- o Finally NAEW, fighters and tankers form the Air Action Group (AAF).

The figure also shows two NAEWs to avoid one NAEW becoming a node, and to allow data relay when the AAG is out of communication range. Replication of NAEW will be an essential plank of survivability, implying that future NAEW will need to exist in greater numbers and cost less than at present; the USNs E2C sets an example, with its small size and cost, at least relative to AWACS/NAEW. Although Figure 15 is MOSAIC based, it is noteworthy that a very similar diagram could be drawn for any ACCS; the interface problem is significant in any design, MOSAIC simply highlights the needs.

8.0 IMPLEMENTING MOSAIC

8.1 Military and Political Obstacles

European NATO is a patchwork of different nations, organisations and infrastructure mapped on to widely differing terrains from Norway to Greece. No one part of the organisation is perhaps any more complex than the funding and acquisition of equipments on a NATO/national basis. Thus MOSAIC faces two immediate implementation hurdles; it seeks to move cells according to overall military needs rather than political foundaries and those cells are unlikely to be equipped with the same aircraft, radars and communication.

MOSAIC also prejudices the current organisation of power within individual national air forces. For MOSAIC to be implemented requires a greater degree of delegated authority than is presently accepted within NATO air forces in order to allow for MOSAIC Cell and AAG potential autonomy. (This mirrors the relative authorities of, say, an air base commander and a ship's captain; the latter generally has much greater local authority consistent with his relative independence).

Land forces in Central Region have their organic Air Defence SAMs. Naval forces in the Baltic and Mediterranean also form part of the AD fabric. For MOSAIC in particular, with its dynamic mobility, these land and naval elements of air power present a special need for coordinated battle management.

Last, there are certain features of air power employment which lend themselves to central rather than distributed control and management. These include: reconnaissance, interdiction, recovery, search and rescue, electronic warfare, in-flight refuelling, contingency planning, resource management, joint service coordination and many others. The degree to which delegation of these presently unified facilities can, or need be, effected is uncertain. Instead it may be that some functions should stay centralised and be kept "above" the MOSAIC tactical level.

8.2 Army, Navy and National Cells

MOSAIC lends itself to the introduction of national cells, with defined operational capabilities and interfaces, but with few specific requirements within the cell. So, we can envisage a number of established Norwegian, Greek, Italian, UK and US cells, say, capable of being integrated into a MOSAIC layer using standard interfaces to communicate laterally and hierarchically. Further, the connection of land and naval MOSAIC cells can be envisaged, and this concept reveals one of MOSAIC's underlying strengths; such land force and naval cells already exist in current practice, are already mobile and already integrate offence, defence and air transport support.

8.3 Applicability

These simple extensions of the MOSAIC concepts can overcome many of the basic military and political obstacles to funding, preserving national interests within NATO and coordinating land, sea and air power. A full embracing of MOSAIC in place of our current evolutionary approach might be ill-advised however. MOSAIC is unlikely to find universal applicability across European NATO, with its widely differing threats, terrains and political objectives. Rather MOSAIC can be seen as most relevant in areas where mobility of the ground battle will be greater, notably in Central Region and its bordering areas. In other areas, a move towards some mobility might introduce a mixture of fixed and MOSAIC like features.

8.4 Transition and Cost

Since MOSAIC is different from current systems, it is reasonable to suspect that transition to MOSAIC would alter timescales and cost vis-a-vis the present plans.

The MOSAIC concepts outlined in this paper describe a complete, eventual system which might not be in place for several decades. Having set the end target, it is possible to backtrack to our present baseline and to define transition stages.

For example, MOSAIC would contain few static radars other than perhaps OTH-B and satellite trackers. Since most radars would be mobile, we can establish mobile radars, CRCs and SOCs as constituting part of Phase 1. Similarly new aircraft designs such as the European Fighter would be affected by the need to operate from prepared strips, or to use arrester-gear on landing. Such requirements might also form part of Phase 1.

The establishment of a communications master plan would constitute part of Phase 1, with the definition of TADIL-ACCS, the definition of protocols, interfaces and gateways, of data dictionaries, display standards and information requirements and the location of nodes for plug-in points during deployed operations.

A mobility plan would interlock with the Communication Plan, but would also address resource dumps, runway availability, maintenance facilities, power supplies and the administration inherent in mobile operations, and already well known to navy and army personnel in their respective spheres of operation.

The Air Action Group concept, which mirrors its naval and army counterparts, is not an essential feature; it requires considerable study by experts to establish credibility. The underlying Sensor MOSAIC, however, is a technical enterprise offering Air Picture Survivability which does not depend for its value on the AAG. Until recently, the SEM would not have been practical; now it is, given the appropriate airborne communications and NAEW processing. Specification of requirements could form an early part of Phase 1, since the SEM can operate in isolation from the other MOSAICs.

Cost is difficult to estimate but could conceivably be reduced in a MOSAIC-ACCS. On the one hand, mobile facilities are of themselves often simpler, on the other hand more such facilities may be needed and the communications infrastructure must be provided. It is reasonable to suppose that, overall, more money will be spent at the execution level, and less will be spent on the higher levels of command and control due to the integration of offence, defence and support. Thus the burden of expenditure will shift towards the "cutting edge" of air power.

9.0 MOSAIC STRENGTHS AND WEAKNESSES

So far MOSAIC has been introduced and discussed in relatively absolute terms; no comparisons have been undertaken with alternatives, and indeed no evaluation process has been considered. One structured approach to comparing options is shown in Figure 16 which, inter alia, identifies the principles of Air Weapons Employment, and option assessment, both of which will be enlarged upon below. At the highest level, then, it may be asked whether MOSAIC is consistent with the Principles of Air Weapons Employment; that question is addressed at Figure 17 in part, from where it can be seen that the three MOSAIC concepts find their strengths chiefly in:

- Flexibility
- Concentration of Force
- Unity of Effort
- Survivability
- Maintenance of the aim

Potential MOSAIC weaknesses are also apparent, and are addressed below.

9.1 Economy of Effort

MOSAIC could require some increase in manpower in particular cases owing to the need for mobility. If support personnel contribute to Economy of Weapons Employment which in one sense they do, then MOSAIC would not score in those particular cases. The term usually applies, however, to accuracy of weapon delivery and sufficiency of damage; in this context MOSAIC is no worse or better than current systems.

9.2 C2 at Highest Practicable Level

The C3 MOSAIC is based on the assumption that operation may be necessary in degraded mode. MOSAIC would therefore draw a compromise betwen C2 at highest level and sufficient delegation to allow effective SUCOC - Succession of Command. The key issue centres on the word "practicable" and requires further assessment.

9.3 Comparative Assessment

Since the basis of MOSAIC is to restructure the deployment and grouping of the elements of air power, it is essential to compare the MOSAIC approach with that which would be achieved by evolution from the present systems, and indeed to compare MOSAIC with other concepts, as previously suggested by Figure 16.

To effect a sensible comparison it is necessary to establish a set of criteria against which to judge and a methodology for trading advantage under one criterion, with disadvantage under another. Figure 18 divides assessment into two areas: performance and survivability (of performance). Performance is concerned both with the effectiveness of forces and the ease of command and control. Survival too is subdivided into Survival of facilities and personnel on the one hand, and the ability to fight on the other.

Each of the broad headings of Figure 18 is capable of considerable elaboration. Certain key issues are discussed below.

9.4 MOSAIC Fratricide Reduction

Fratricide presents one of the most intractable problems facing the employment of air power. The MOSAIC concepts offer surprising and unexpected potential to reduce fratricide levels significantly, for a number of reasons:

- Own offensive forces would be "mixed-in" with defensive forces, and would not therefore traverse a complete layered air defence when going to, and returning from, a target. On average, and viewed simplistically, they would overfly only half of the friendly defensive positions from which they are at risk, offering "at a stroke" potential for a major fratricide reduction.
- The close coordination of offence and defence essential for fratricide reduction is made possible at cell and cell group level in MOSAIC.
- The lateral communications and the possible inclusion of land and naval forces in the MOSAIC concept lend themselves to effective development of the Indirect Sub-System (ISS) of NIS, the NATO Identification System, which will contribute directly to fratricide reduction.

MOSAIC offers sufficient scope for fratricide reduction that for this reason alone it may be worth pursuing MOSAIC concepts further.

9.5 Availability

Operational availability of weapon systems and sensors is likely to be affected by movement of bases, and generally the effects may be adverse. This aspect pales to insignificance, however, when set against the rise in survivability offered by MOSAIC which, in any conflict, would rapidly over compensate for loss in availability, and in the employment of less sophisticated ground systems which tend to be more reliable.

9.6 Force Management

Because MOSAIC presents many more options, and because it tends to operate with smaller groups in many cells, force management will be more complex than for static conventional systems. It also has the potential to be immeasurably more effective by rapidly redeploying mobile units - the essence of Force Multiplication.

9.7 Airspace Management

Centralisation of military air movement plans and 'actuals' within MOSAIC Cell Groups will assist with Airspace Management, which will nonetheless remain a largely intractable problem owing to the separate interest of civil and military air traffic.

9.8 Avoidance of Detection

Moveability as the keynote of the GEM is concerned with concealement of Battle Order and avoidance of detection in the sense that cells should be able to set up, operate, break down and move in the time it would take an enemy to detect, locate, target and attack. The use of PSTN communications, passive sensors, LPI radars and LPI communications are all contributors to Avoidance of Detection.

9.9 Self Defence

The MOSAIC concept engenders complete fighting units as cells, including self defence, which is therefore one of the key benefits of MOSAIC.

9.10 Damage Tolerance

Within each MOSAIC cell, damage tolerance may be prejudiced by lessening the degree of hardening consistent with the need for moveability. However, the multi-nodal nature of the MOSAIC cell pattern provides overall structural damage tolerance in another way, since the separate cells can operate, in the final analysis, autonomously. MOSAIC thus enhances Damage Tolerance of the Air Power overall.

9.11 SUCOC/COLOC

The C3 MOSAIC provides a natural ease for both SUCOC and COLOC (Change of Location of Command), since the lateral communication facilities will provide the basis for alternative command locations and the multi-layer cellular structure presumes SUCOC as a means of degraded-mode operation.

9.12 SOPs

Standard Operating Procedures are the intertia which will keep the flywheel of operations rotating when the command and control driver is absent. For MOSAIC, SOPs would be doubly important to promote semi-autonomous and autonomous operation as the C2 fabric became progressively disrupted by battle damage.

10.0 MOSAIC IN ACTION

Before drawing to a close, it is interesting to look into the future and envisage an engagement in which MOSAIC concepts are in action under a future threat. Figure 19 shows a collection of systems, some real, some conceptual. At bottom centre is a Ground Enviroment MOSAIC (GEM) being threatened by a ground force at right and by a 360° air attack. The GEM contains all of the essentials to attack, defend and redeploy autonomously if need be. One GEM cell has been vacated at right as part of Rollback and is reforming at left. Immediately overhead are NATO Early Warning/Air Command Aircraft, acting as focal points for Air Action Groups, long-range radar sources and communications relays, at left a NATO Emergency Air Command Post (NEACP) aircraft is postulated, containing a senior battle commander to coordinate, inter alia, MOSAIC activities.

In space are a set of satellites for navigatiion, communications, nuclear flash detection (IONDS is the US Integrated Operational Nuclear Detection System), surveillance, intelligence, missile IR-flash launch detection (IMEWS is the US Integrated Missile Early Warning System) and missile/warhead engagement. Fighting Mirror is one of the SDI concepts for such engagements but is modified in this representation. In the full SDI concept a ground based excimer laser directs energy at a mirror in geosynchronous orbit, which reflects that energy to a second, lower orbit mirror and thence on to a moving ballistic missile or warhead target, delivering up to 50 Megajoules per square metre and vaporising the metallic target covering. In Figure 19, only one mirror is shown for simplicity.

At mid-left, in addition to the excimer laser are satellite trackers, C2 HQ employing US military satellites and two long range radars, COBRA DANE and PAVE PAWS, the latter being a phased array warning radar said to be able to detect a grapefruit-sized target at 1000 miles. PAVE PAWS is a so-called JANUS radar because it can comprise two back-to-back phased array antennae.

The scenario of Figure 19 is interesting in several respects,

- Only Fighting Mirror is truly futuristic; with the exception of MOSAIC, which is not technologically innovative, all the other systems are operational now.
- The space sensors cover, interalia, areas of direct interest to a future ACCS, MOSAIC or otherwise.
- Although the space sensors and weapons may have a strategic label, the speed and accuracy of responses implicit in their implementations renders them of great tactical importance too.

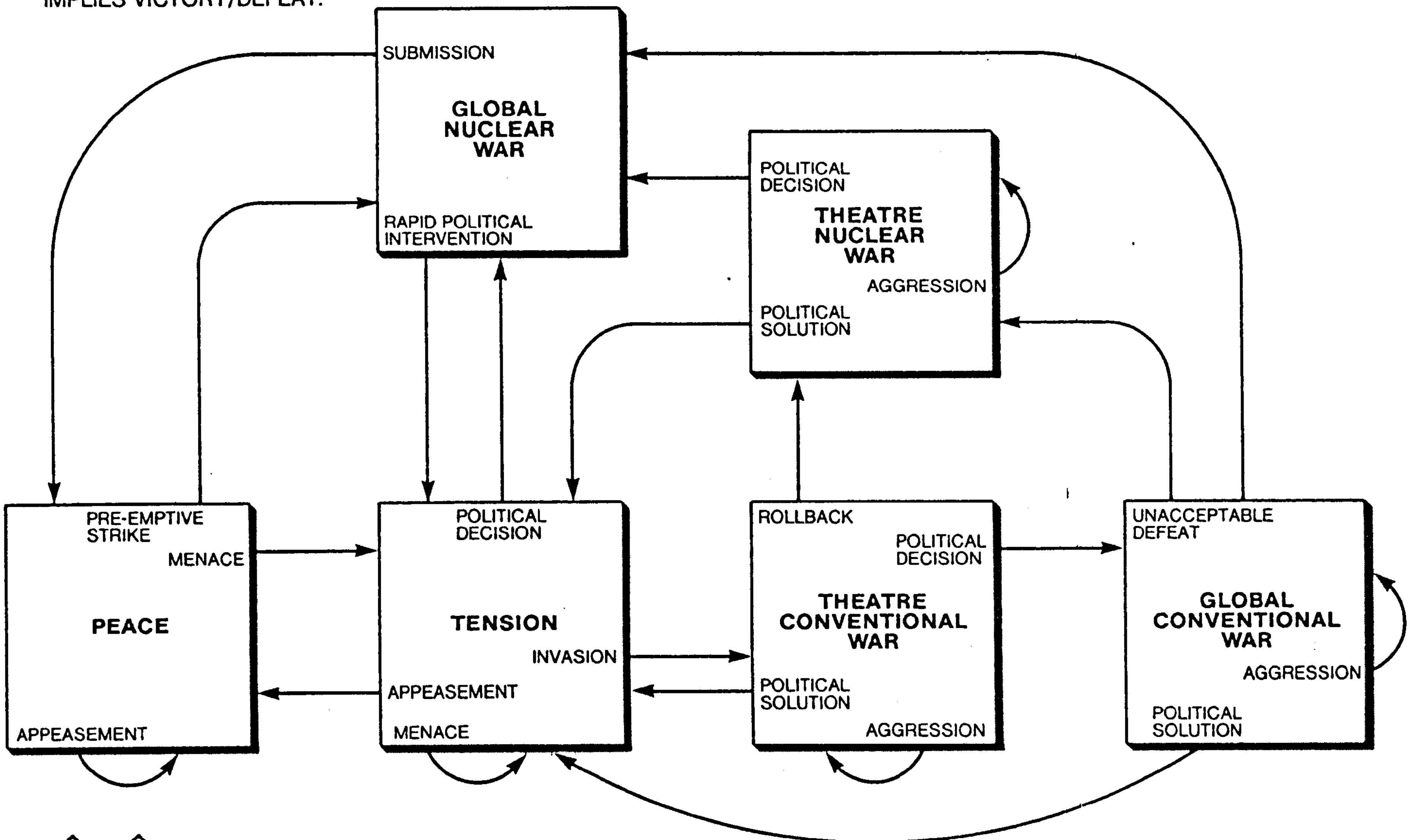
11.0 CONCLUSIONS

The presentation and assessment of MOSAIC have:

- Provided a conceptual ACCS design solution which is consistent with the evolving threat and political factors, yet which addresses some current Air C3I deficiencies.
- Highlighted the implications of full mobility, and in so doing illustrated the implications of any significant degree of mobility in any future ACCS concept.
- Introduced new concepts which arise out of MOSAIC, but which may have much wider application:
  - -the Air Action Group
  - -TADIL ACCS and the other interchange languages.
- Identified the potential for very significant reductions in fratricide using MOSAIC.
- Presented a set of high level criteria by which MOSAIC and other ACCS concepts may be judged.
- Assessed MOSAIC against the Principle of Air Weapons Employment and found it to be substantial.
- Indicated that transition to MOSAIC is practicable and offers a highly survivable, effective and flexible ACCS.

One major value which can be achieved from analyses such as MOSAIC is that, in driving system designs towards their logical limits, existing concepts need to be reassessed and new concepts may be introduced which, like TADIL-ACS and SDI interfacing, are of value to any future ACCS design.

NB. POLITICAL SOLUTION GENERALLY
IMPLIES VICTORY/DEFEAT.



MOSAIC

FINITE STATE DIAGRAM OF EUROPEAN POTENTIAL CONFLICT TRANSITION/RESOLUTION

AIR OPERATIONS

OFFENCE

- FAR RANGING
- FAST
- FLEXIBLE
- VULNERABLE ON GROUND
- VULNERABLE IN AIR

DEFENCE

- RELATIVELY SHORTER RANGE
- VERY FAST
- VULNERABLE TO JAMMING
- VULNERABLE ON GROUND
- VULNERABLE IN AIR
- SPECIALISED
- DEFENDS OFFENCE, DEFENCE & SUPPORT
- PRESENTLY 'TIED' TO ADGE

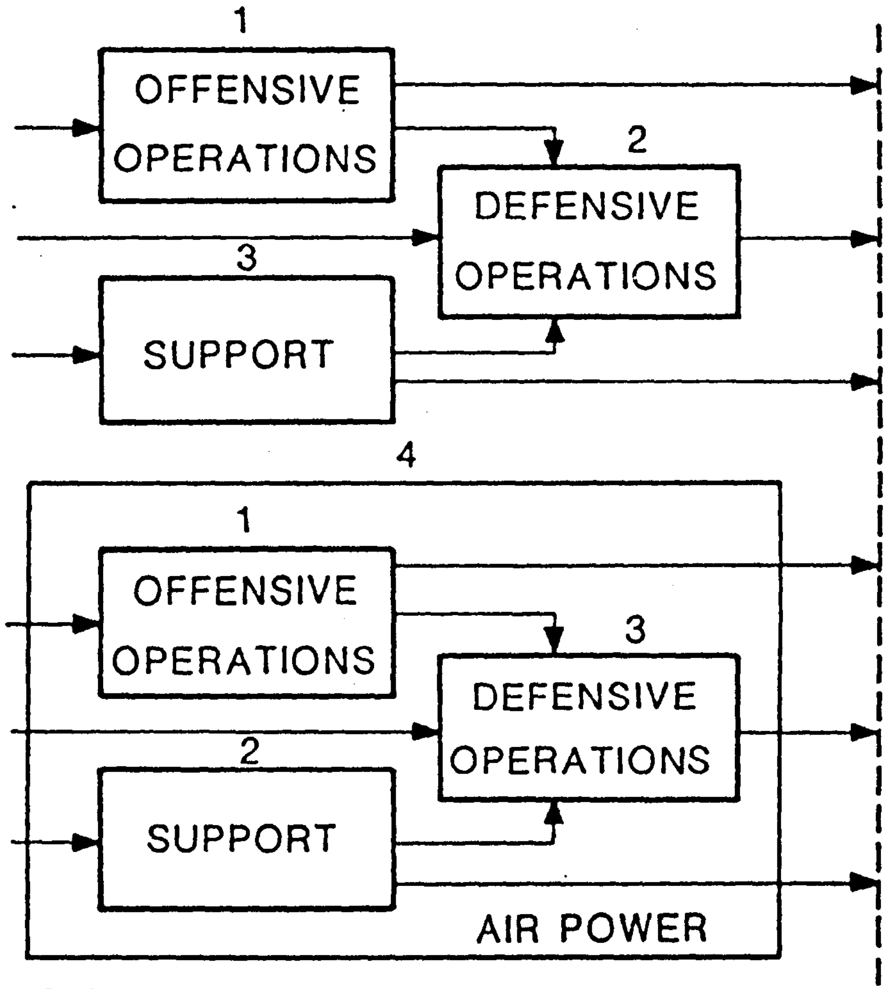
SUPPORT

- FAR RANGING
- FAST
- FLEXIBLE
- SUPPORTS OFFENCE, DEFENCE & SUPPORT
- VULNERABLE ON GROUND
- VULNERABLE AIRBORNE

MOSAIC

THE CONCEPT OF INDIVISIBLE AIR POWER

SYSTEM DESIGN AXIOM AND COROLLARY



AXIOM

"A SYSTEM COMPRISED OF ELEMENTS WHICH HAVE BEEN SEPARATELY OPTIMISED, WILL NOT ITSELF BE OPTIMISED"

COROLLARY

"AN OPTIMISED SYSTEM WILL CONTAIN ELEMENTS WHICH ARE NOT, OF THEMSELVES, OPTIMAL"

REASONS:

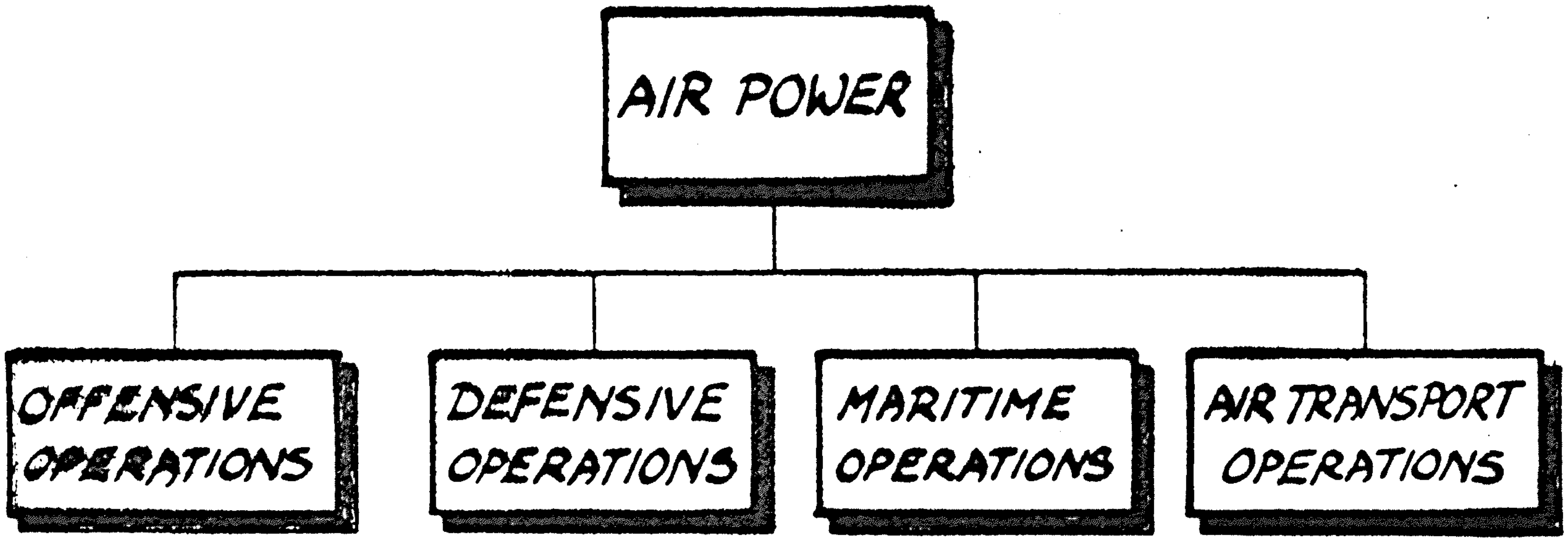
(A) 4 IS A NEW FUNCTION. ITS OPTIMISATION HAS NOT BEEN ATTEMPTED.

(B) THE INTERFACES BETWEEN THE ELEMENTS 1, 2 AND 3 HAVE NOT BEEN CONSIDERED IN THEIR CONTEXT OF 4.

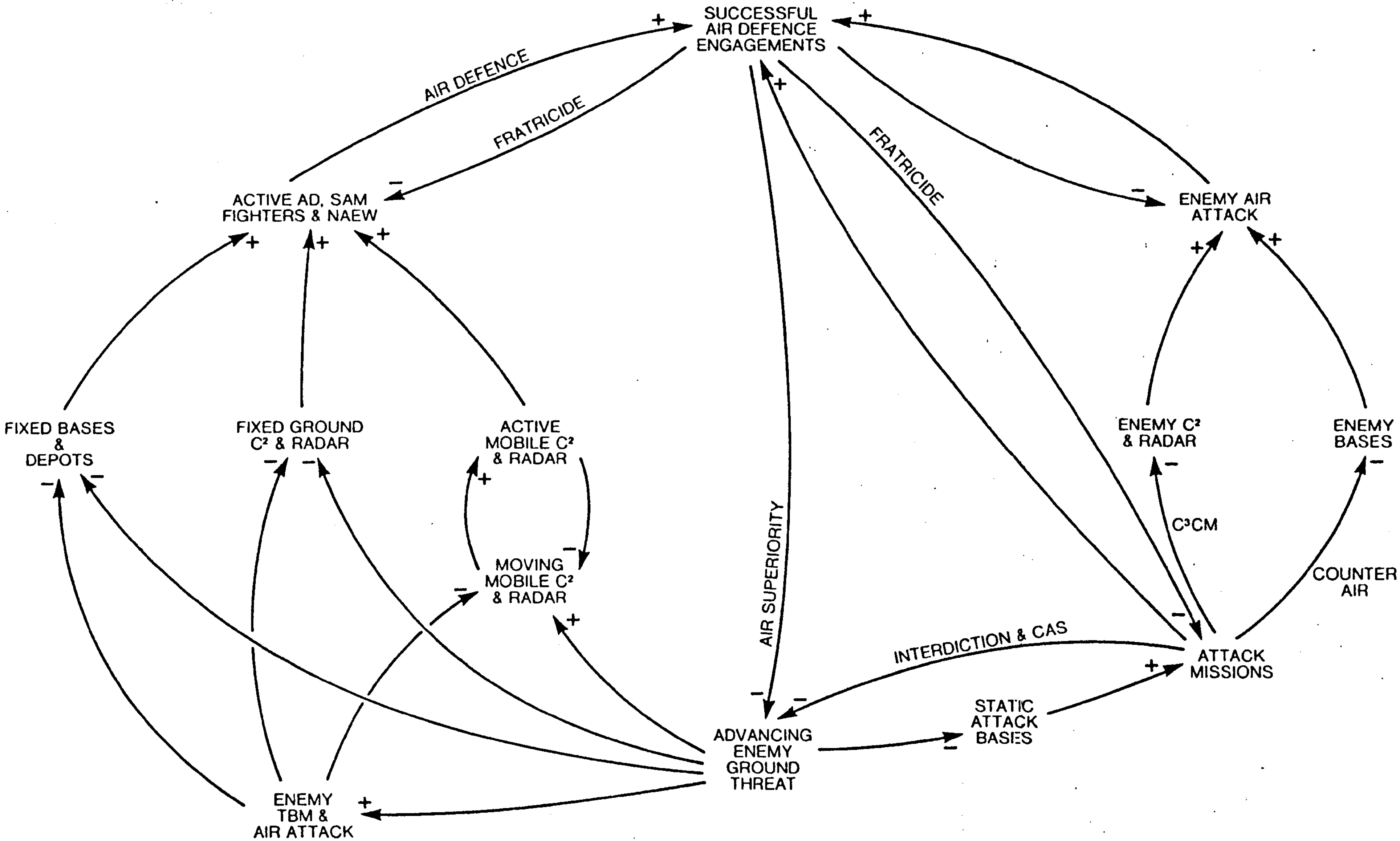
CONCLUSION:

IMPROVING THE EFFECTIVENESS OF AIR POWER REQUIRES A TOTAL SYSTEM APPROACH, RATHER THAN A FUNCTION-BY-FUNCTION DEVELOPMENT

AIR POWER – IDEALISED ORGANISATION CHART

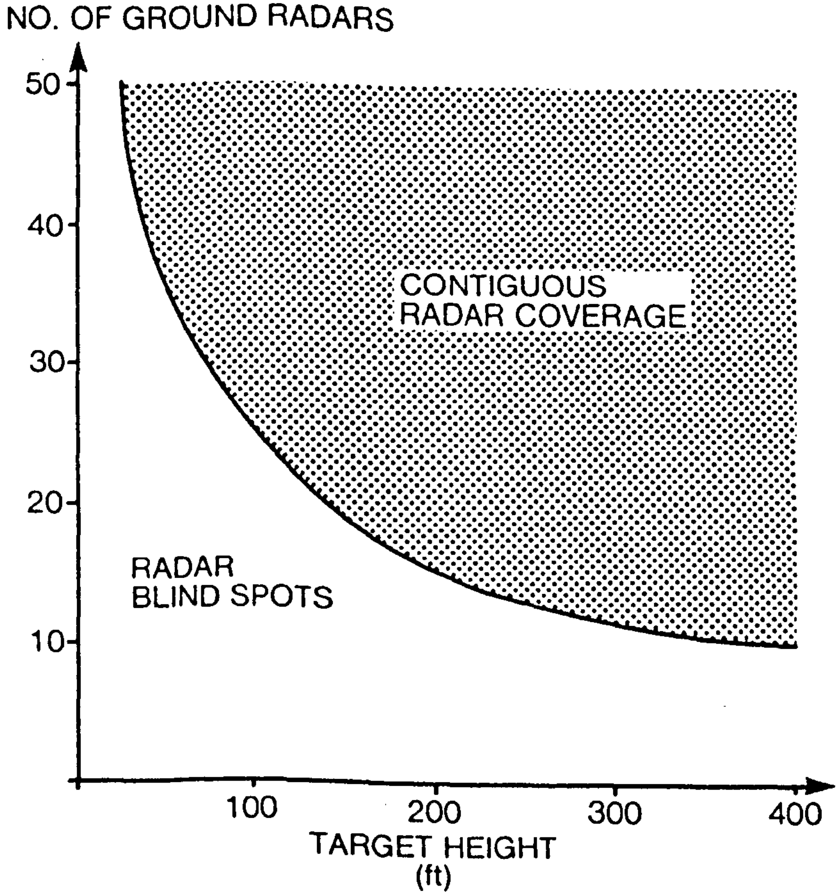


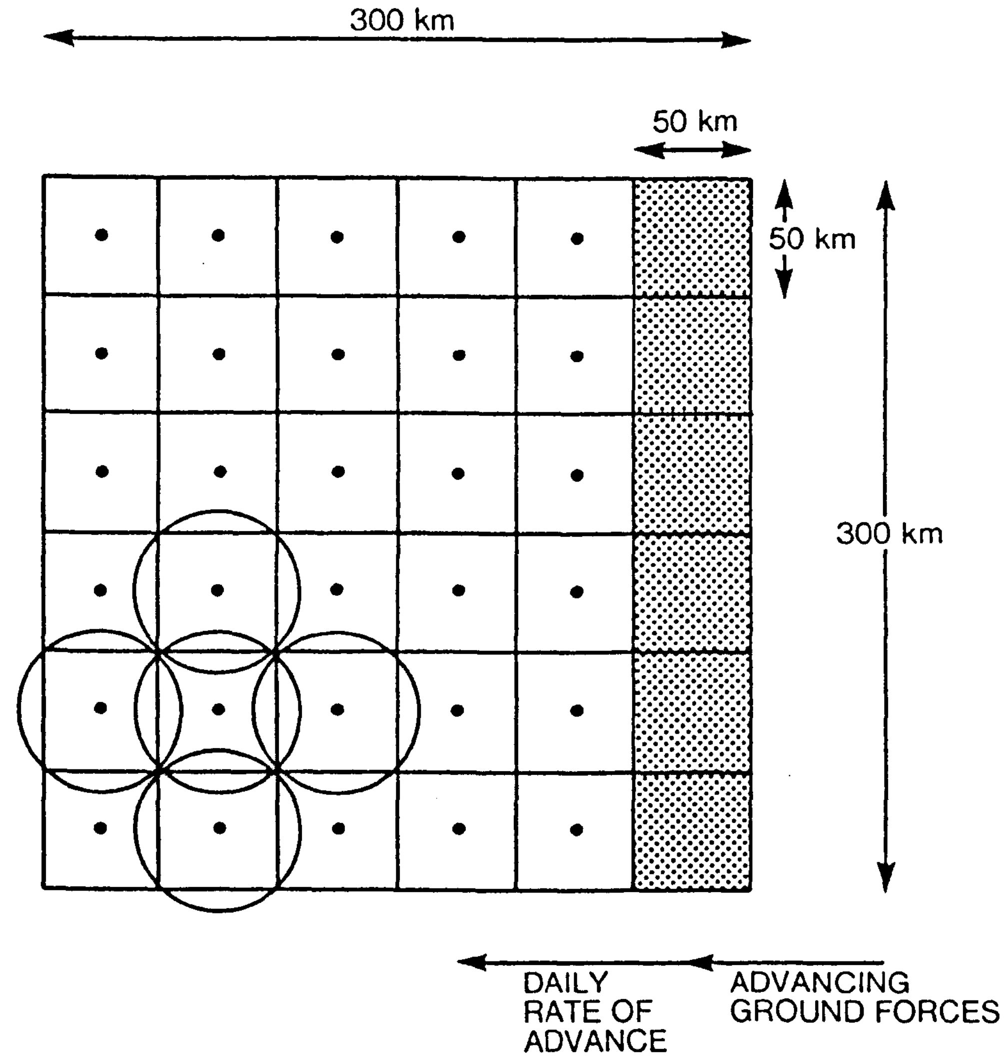
• IDS	• FIGHTER INTERCEPT	• LRMP	• TACTICAL TRANSPORT
• CLOSE AIR	• SAM-POINT - AREA	• ASW	• REFOR & RESUPPLY
• COUNTER AIR	• AEW/AWACS	• TASMO	• MOBILE FORCES
• AIR SUPERIORITY	• ADGE	• AIR DEFENCE	



MOSAIC

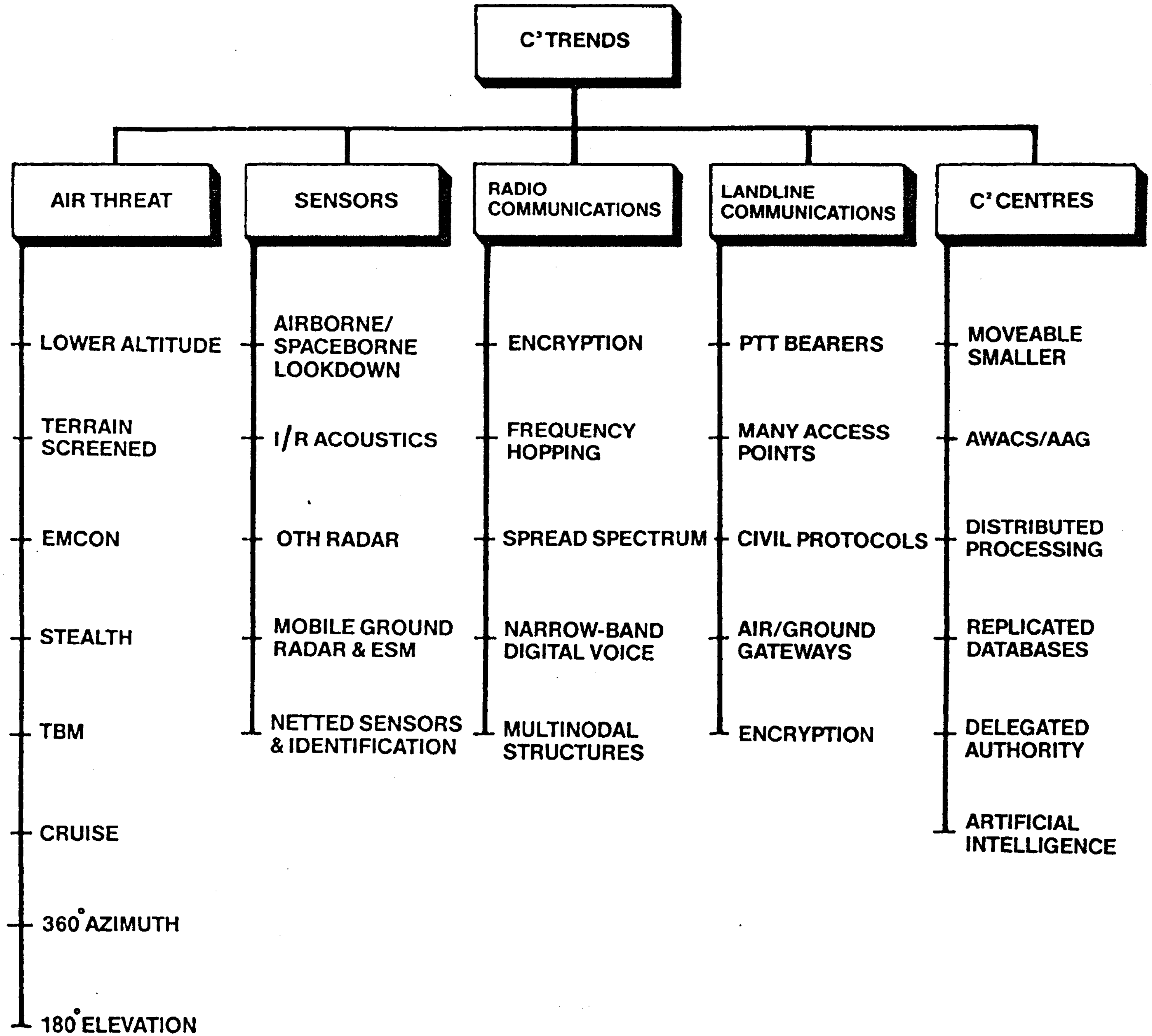
ADVANCING GROUND THREAT AND THE AIR BATTLE

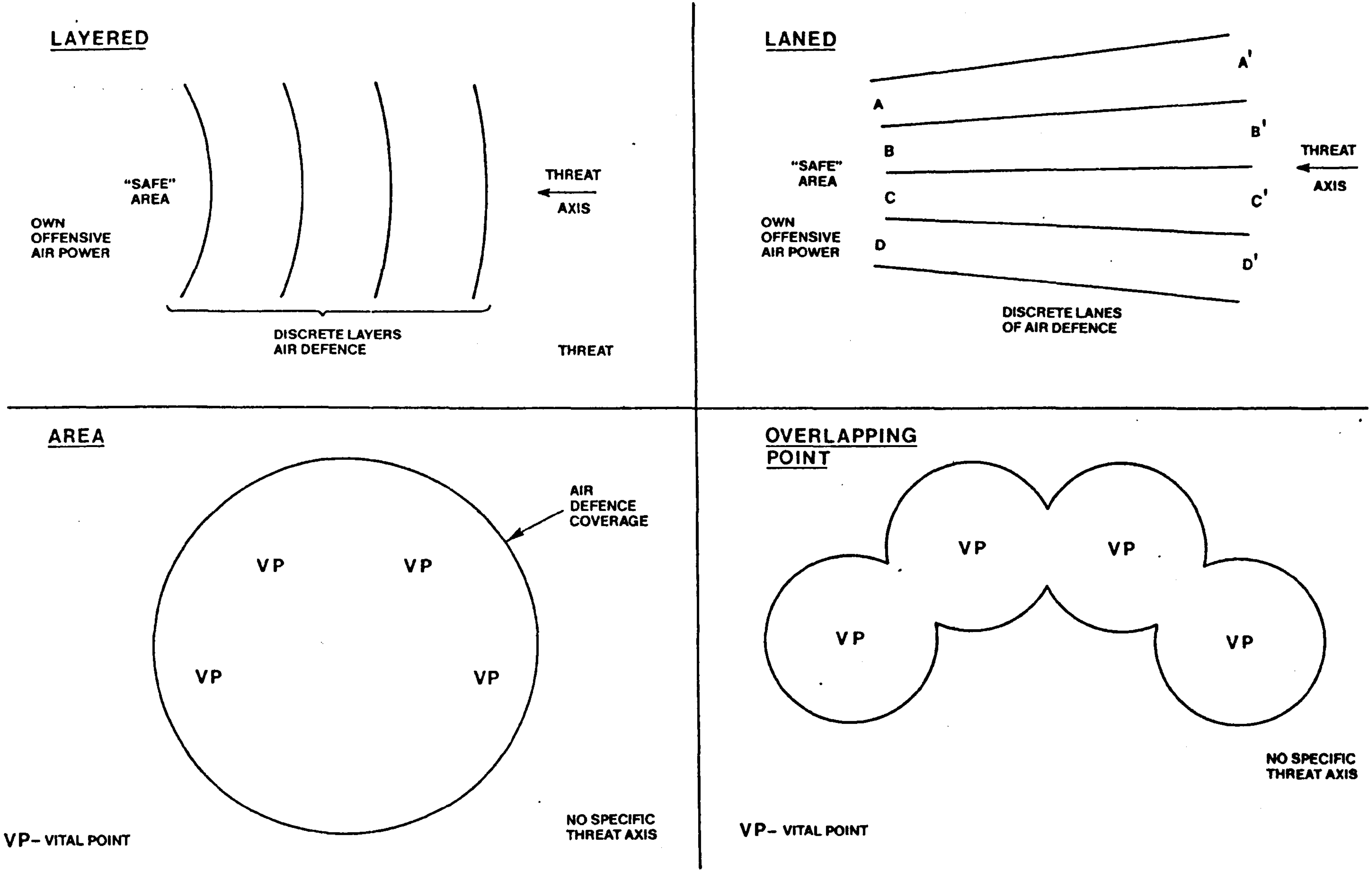




MOSAIC

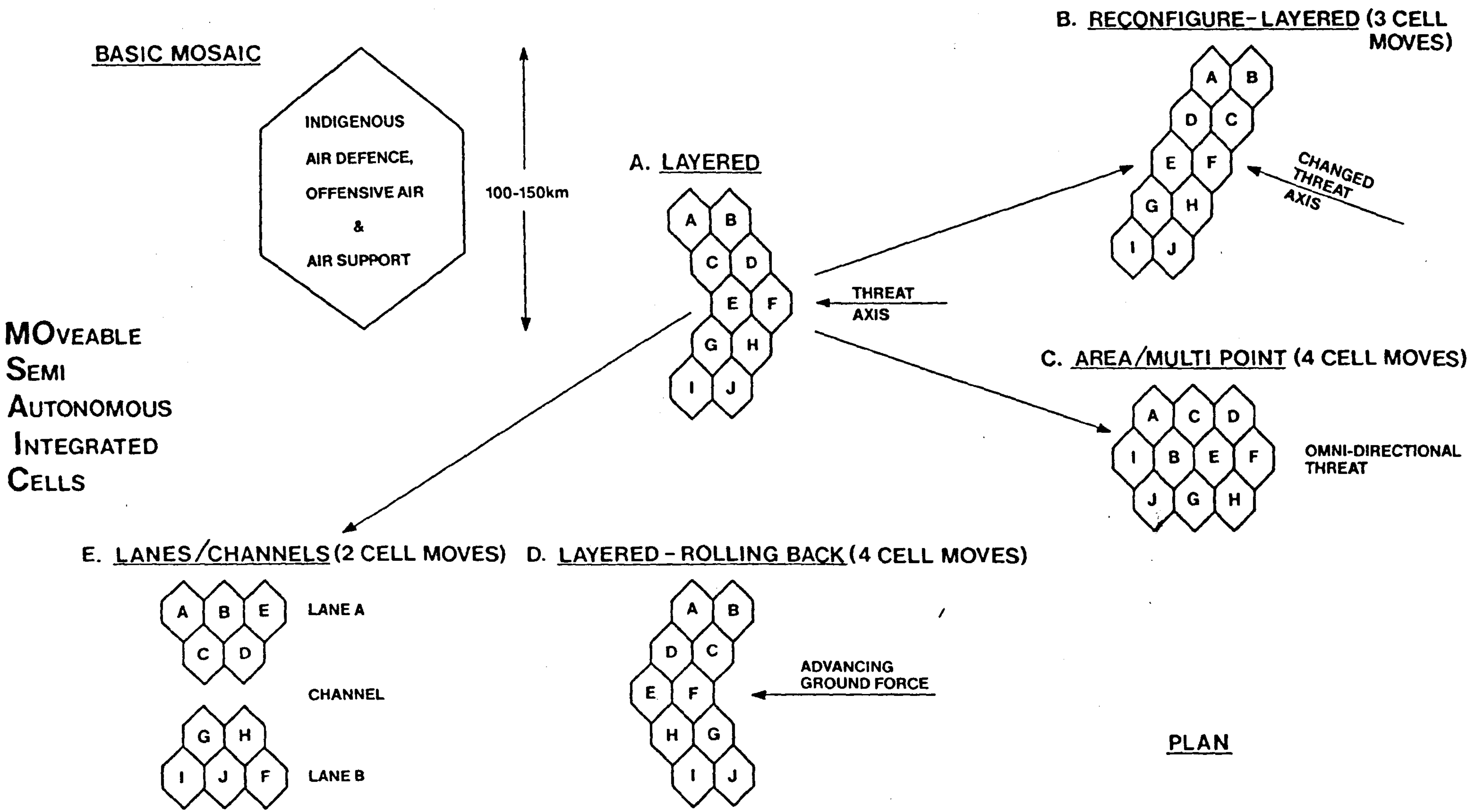
MINIMUM GROUND RADAR DENSITY FOR CONTINUOUS TRACKING OF A 50' TARGET (SMOOTH EARTH)


C^3 TRENDS
AIR THREAT
LOWER ALTITUDE
TERRAIN SCREENED
EMCON
STEALTH
TBM
CRUISE
360° AZIMUTH
180° ELEVATION
SENSORS
AIRBORNE/ SPACEBORNE LOOKDOWN
I/R ACOUSTICS
OTH RADAR
MOBILE GROUND RADAR & ESM
NETTED SENSORS & IDENTIFICATION
RADIO COMMUNICATIONS
ENCRYPTION
FREQUENCY HOPPING
SPREAD SPECTRUM
NARROW-BAND DIGITAL VOICE
MULTINODAL STRUCTURES
LANDLINE COMMUNICATIONS
PTT BEARERS
MANY ACCESS POINTS
CIVIL PROTOCOLS
AIR/GROUND GATEWAYS
ENCRYPTION
C^2 CENTRES
MOVEABLE SMALLER
AWACS/AAG
DISTRIBUTED PROCESSING
REPLICATED DATABASES
DELEGATED AUTHORITY
ARTIFICIAL INTELLIGENCE

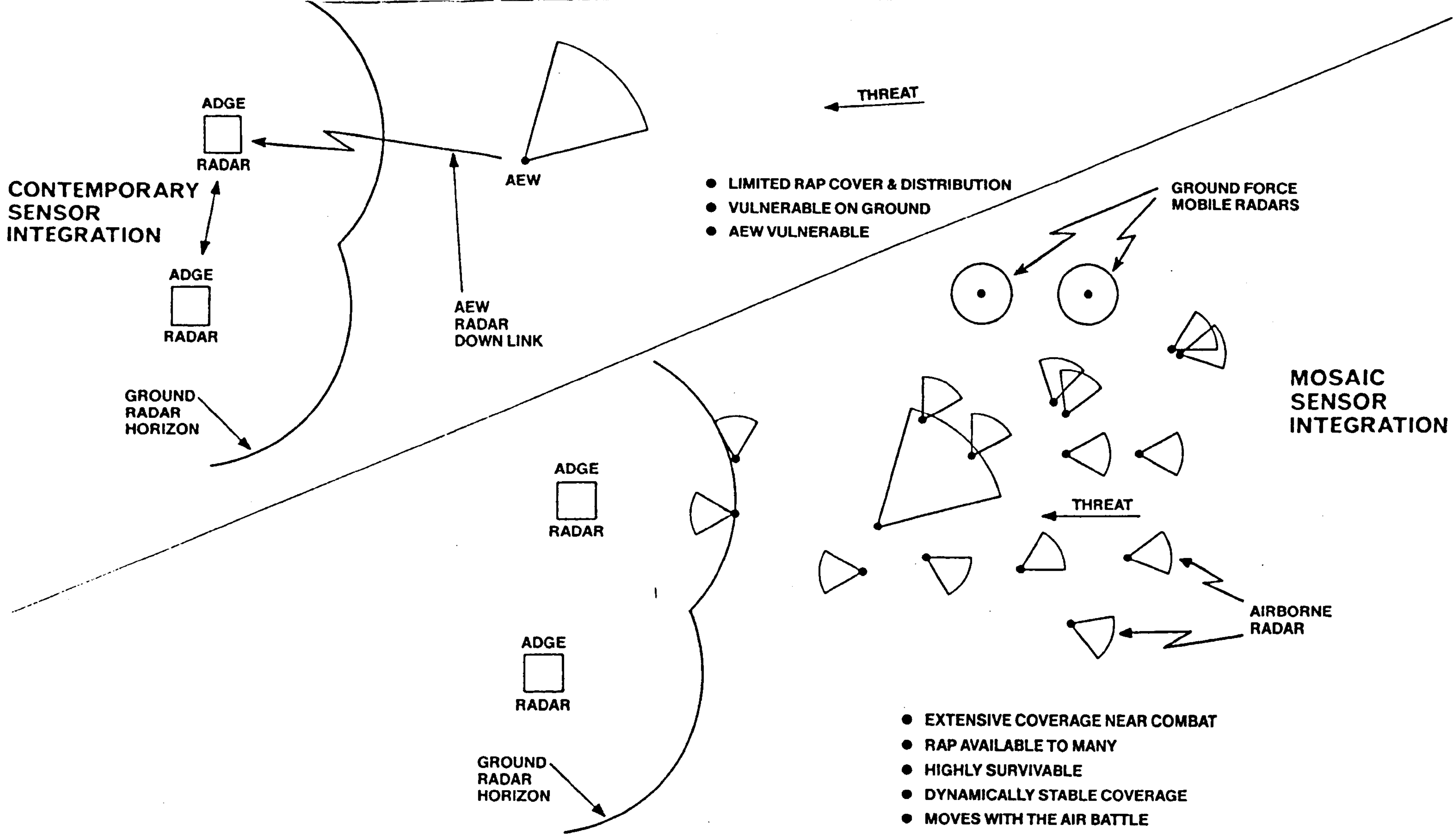


MOSAIC

AIR POWER DEPLOYMENT OPTIONS

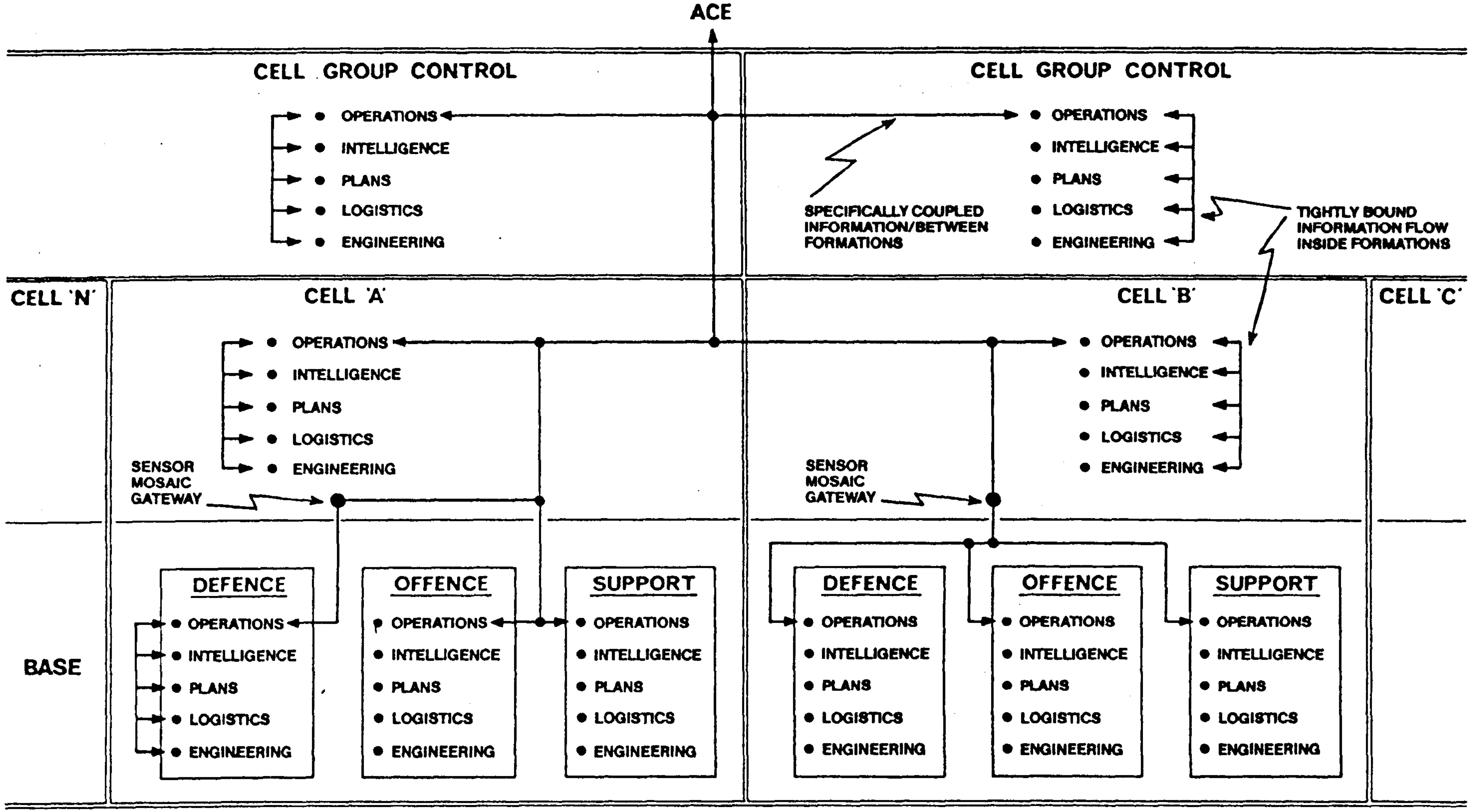


THE GROUND ENVIRONMENT



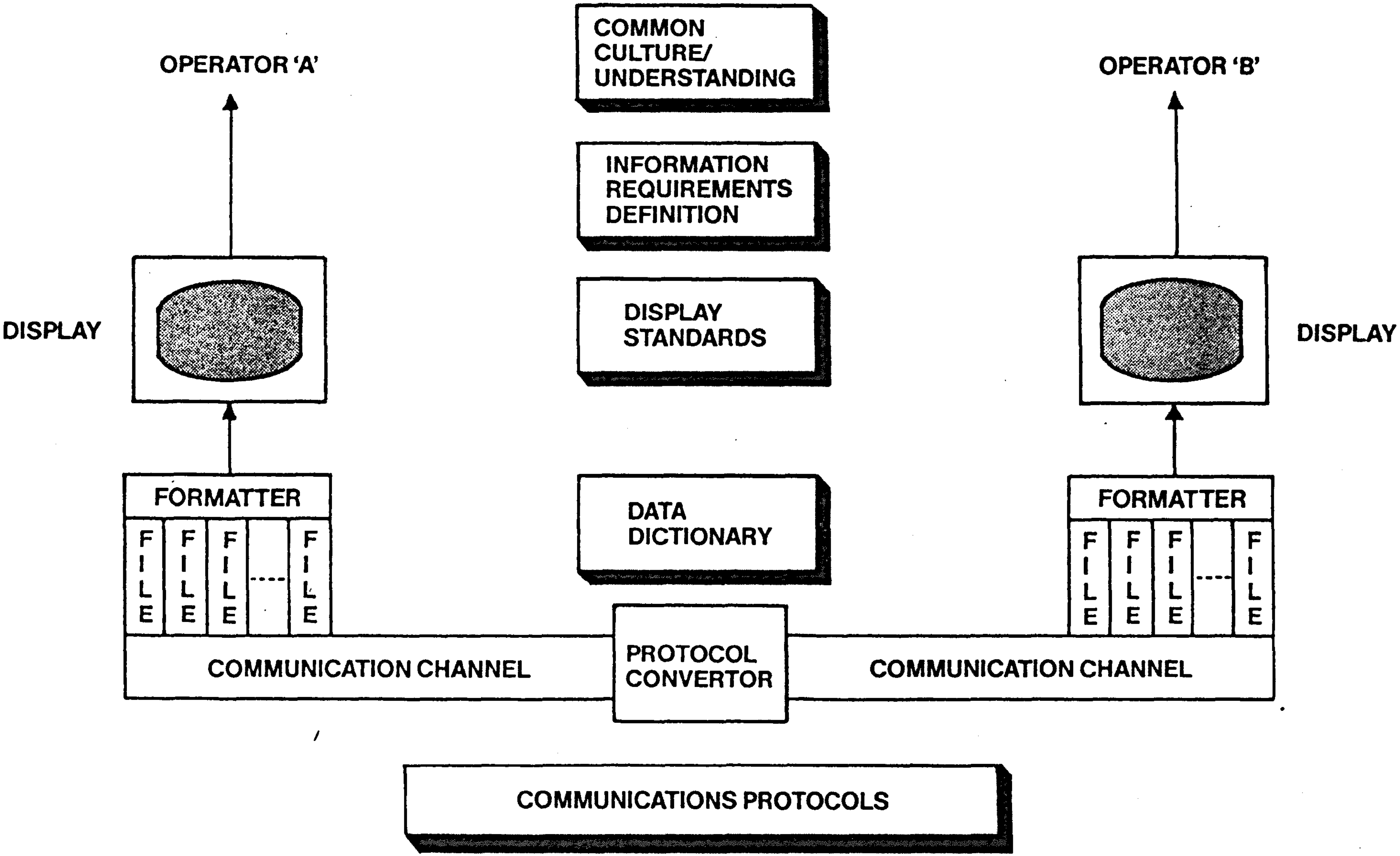
MOSAIC

THE SENSOR MOSAIC



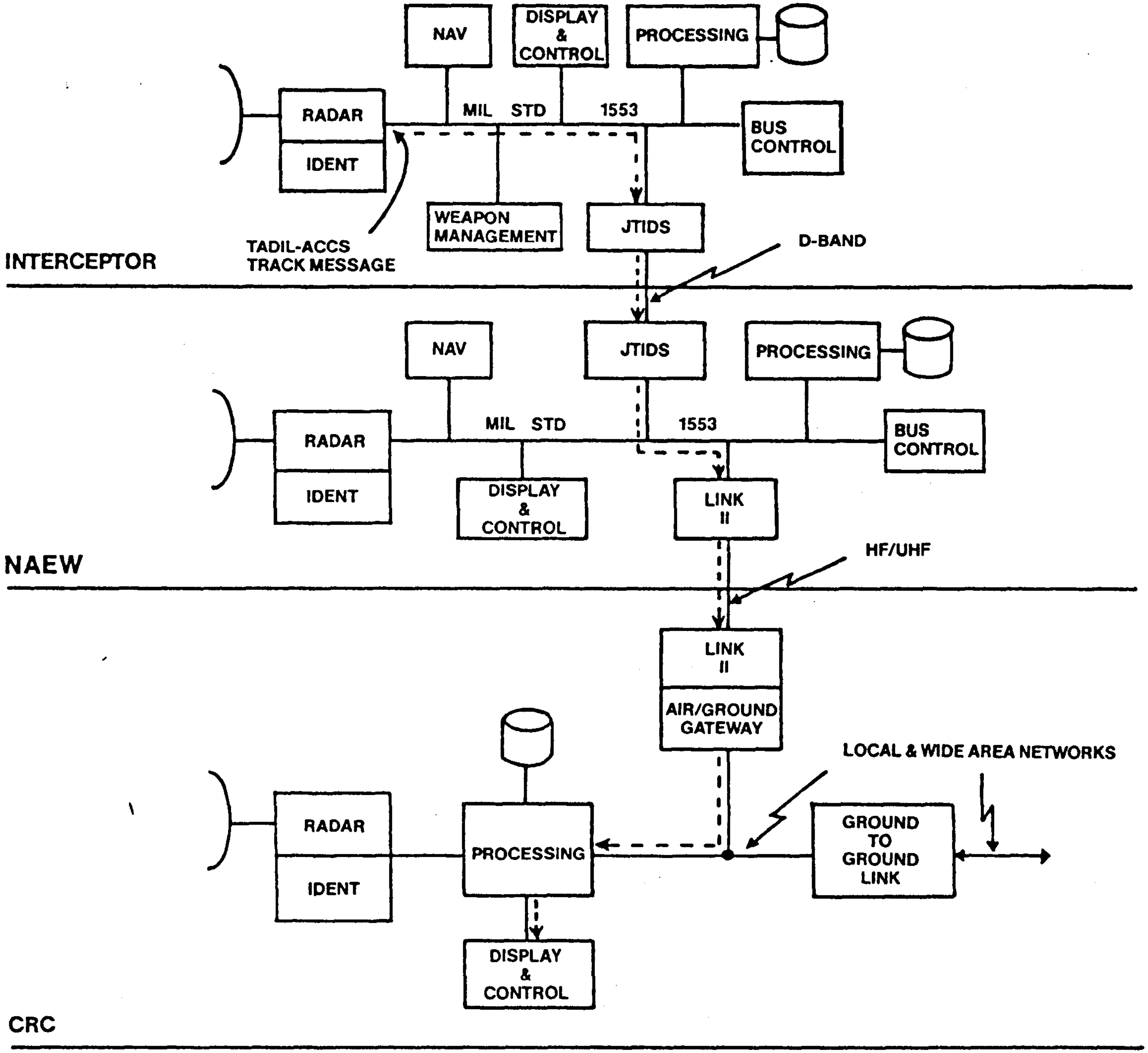
MOSAIC

THE C^3 MOSAIC



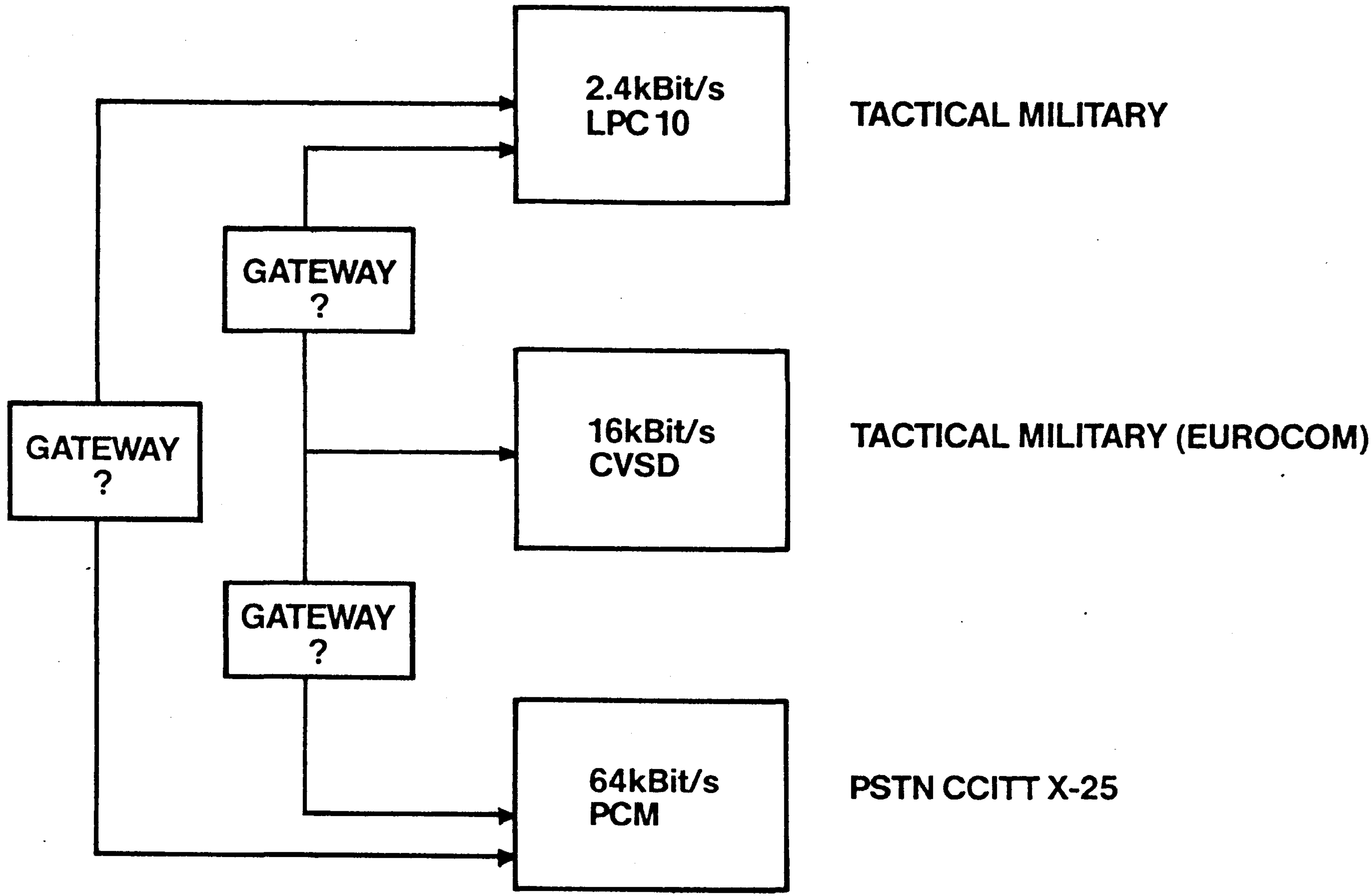
MOSAIC

INTEROPERABILITY



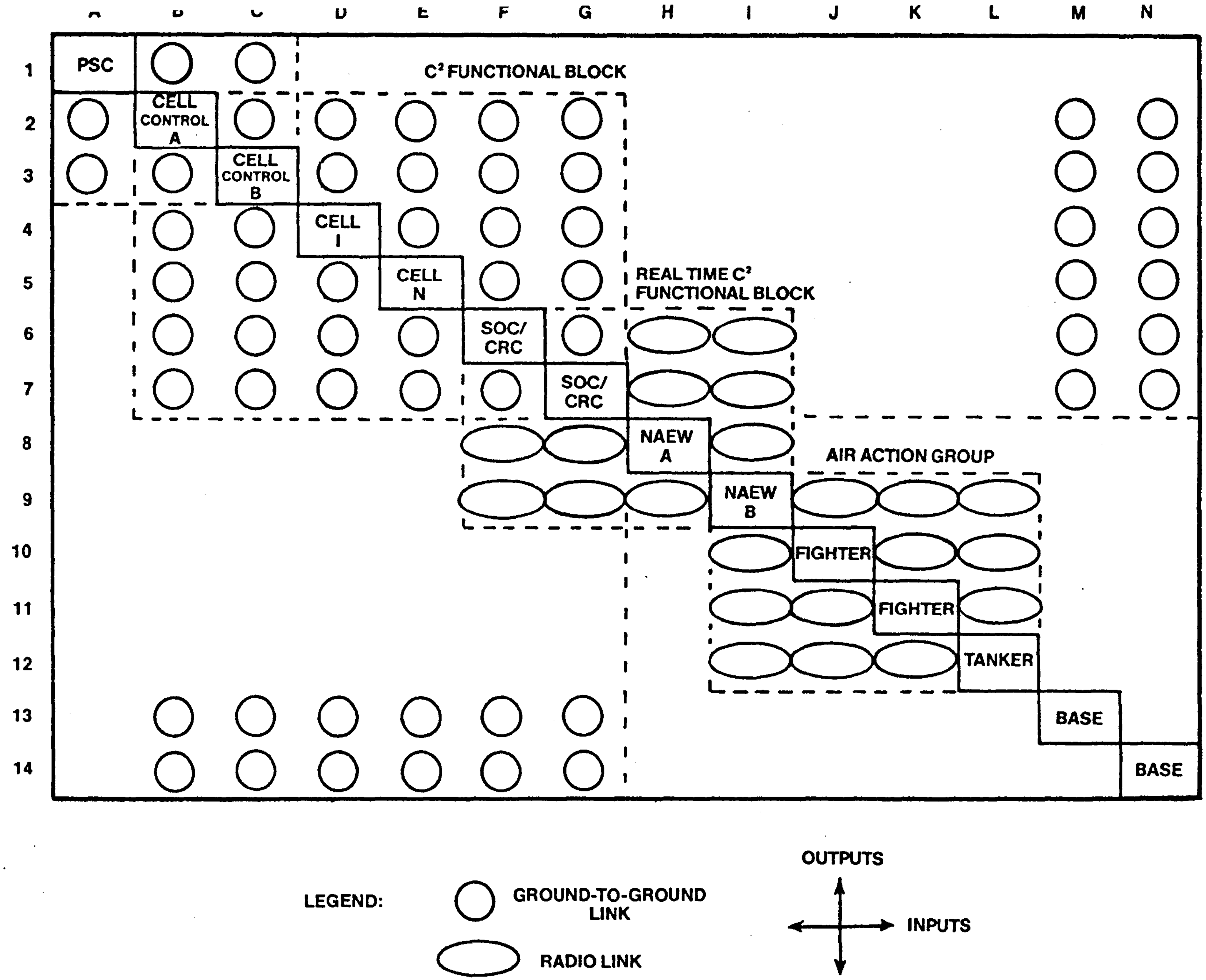
LEGEND: - - - TADIL-ACCS TRACK MESSAGE ROUTING (EXAMPLE ONLY)

TADIL-ACCS



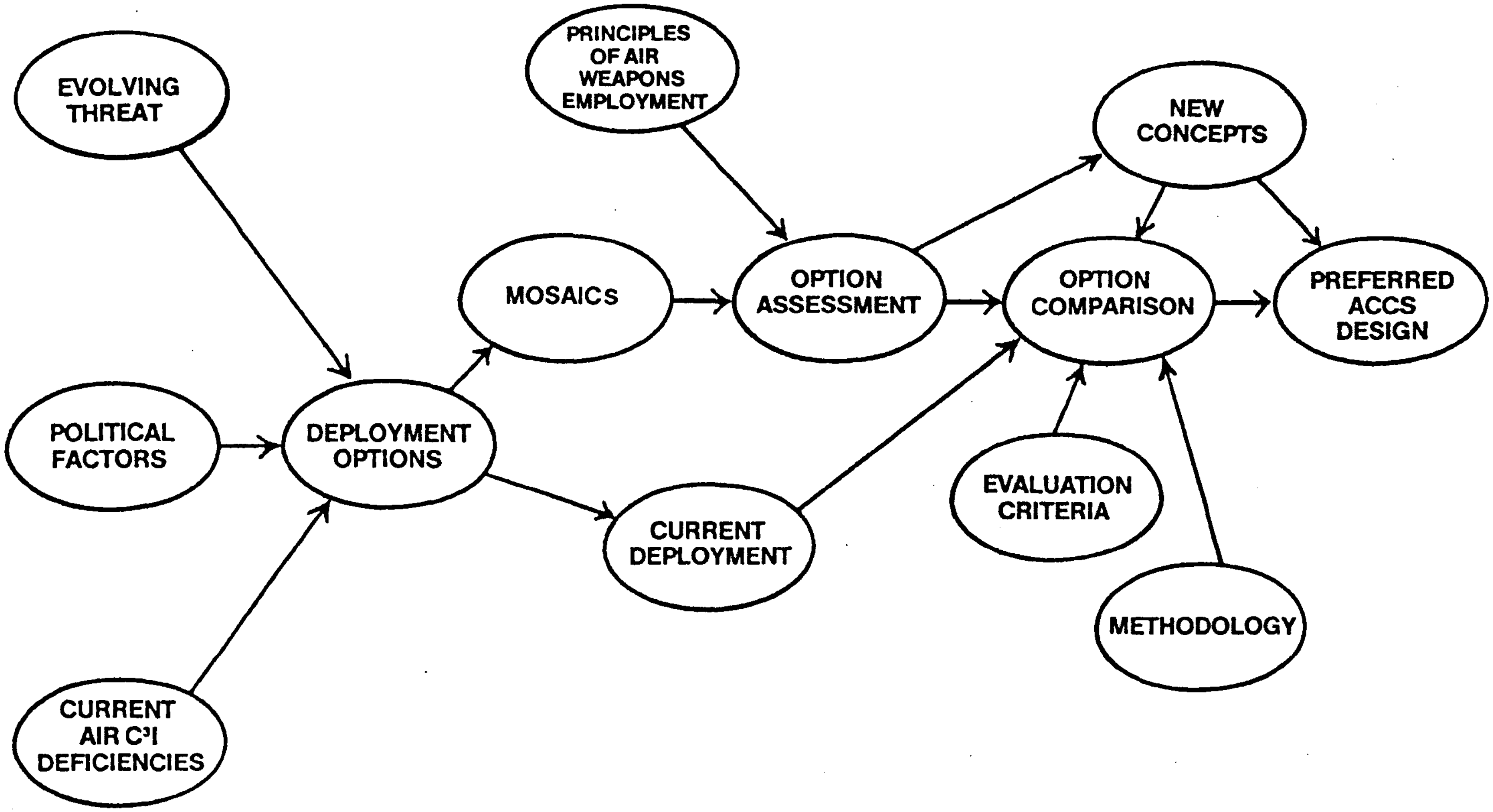
MOSAIC

VOICE COMMUNICATIONS GATEWAYS



MOSAIC

SIMPLIFIED MOSAIC INTERFACE STRUCTURE **Figure**



EVOLVING THREAT
PRINCIPLES OF AIR WEAPONS EMPLOYMENT
NEW CONCEPTS
MOSAICS
OPTION ASSESSMENT
OPTION COMPARISON
PREFERRED ACCS DESIGN
POLITICAL FACTORS
DEPLOYMENT OPTIONS
CURRENT DEPLOYMENT
EVALUATION CRITERIA
METHODOLOGY
CURRENT AIR C³I DEFICIENCIES

MOSAIC

Principle of Air Weapons Employment	MOSAIC			COMMENTS
	GROUND ENVIRONMENT MOSAIC	SENSOR MOSAIC	C^3 MOSAIC	
1. Attainment of a favourable air situation	●	●	N/A	MOSAIC contributes by surviving
2. Flexibility	●●	●●	●●	Principal objective
3. Concentration of Force	●●	●●	●	Principal objective
4. Economy of Effort	?	?	?	Mobility means more personnel,
5. Unity of Effort	●●	●●	●●	combines offence, defence & support
6. C^2 at highest practicable level	●	●●	?	C^3 Mosaic presumes degraded operation as a potential need
7. Offensive action	N/A	N/A	N/A	
8. Survivability	●●	●●	●●	Principal objective
9. Surprise	●	●	N/A	Reconfigurability and the AAG contribute to surprise
10. Pre planning	N/A	N/A	N/A	
11. Maintenance of the aim	●●	●●	●●	Principle objective

MOSAIC AND THE PRINCIPLES OF WEAPONS EMPLOYMENT

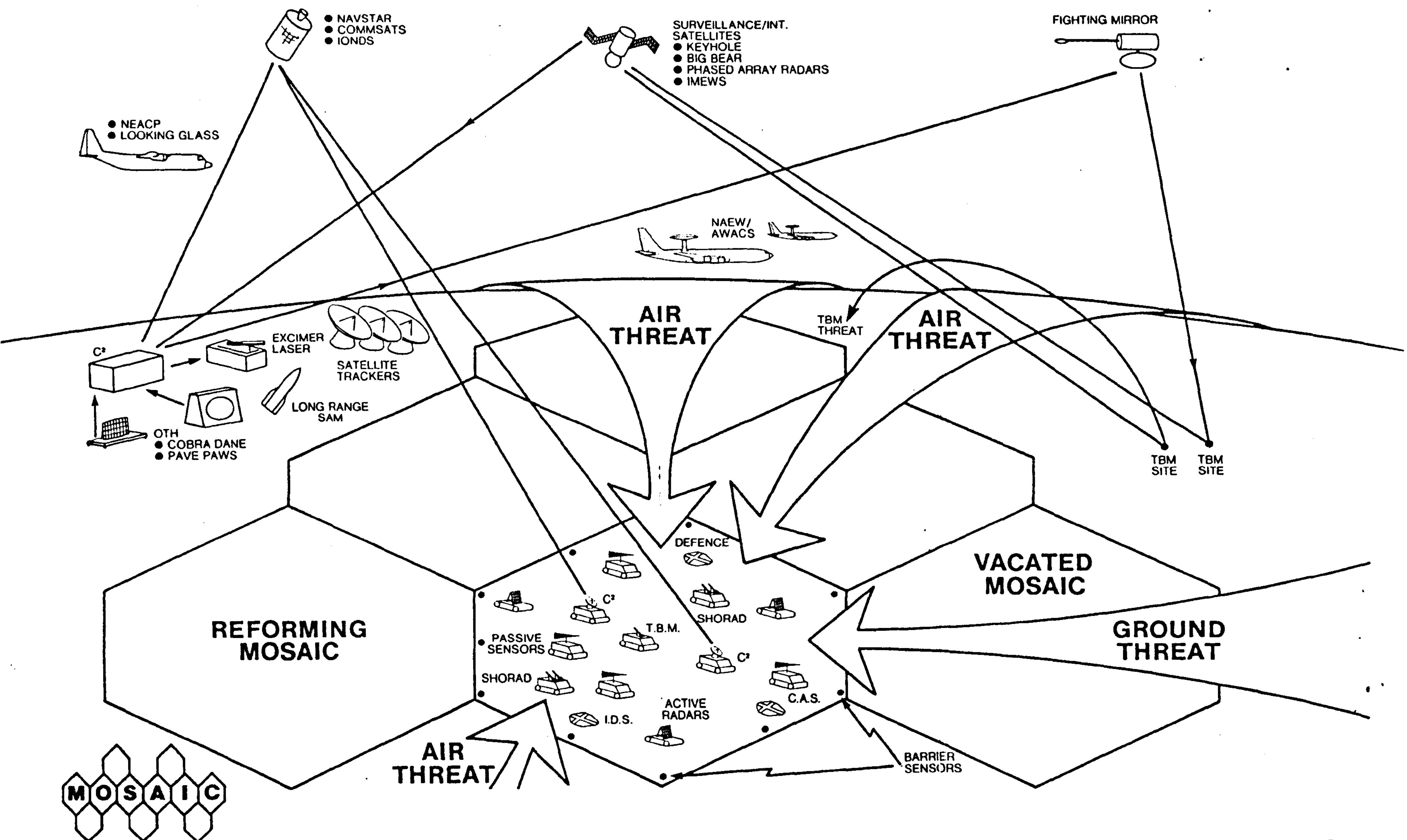
MOSAIC

ASSESSMENT OF DEPLOYMENT OPTIONS

- PERFORMANCE
  - EFFECTIVENESS
    - OFFENCE
    - DEFENCE
    - FRATRICIDE
    - AVAILABILITY
  - EASE OF COMMAND & CONTROL
    - BATTLE MANAGEMENT
    - FORCE MANAGEMENT
    - AIRSPACE MANAGEMENT
- SURVIVAL (OF)
  - TACTICAL: FACILITIES/ PERSONNEL
    - AVOIDANCE OF DETECTION
    - SELF DEFENCE
    - DAMAGE TOLERANCE
  - STRATEGIC: ABILITY TO FIGHT
    - SUCOC/COLOC
    - SOPs
    - AUTONOMOUS OFFENSIVE/DEFENSIVE CELLS

MOSAIC

ASSESSMENT OF DEPLOYMENT OPTIONS


● NAVSTAR
● COMMSATS
● IONDS
SURVEILLANCE/INT. SATELLITES
● KEYHOLE
● BIG BEAR
● PHASED ARRAY RADARS
● IMEWS
FIGHTING MIRROR
● NEACP
● LOOKING GLASS
NAEW/ AWACS
AIR THREAT
TBM THREAT
AIR THREAT
C^2
EXCIMER LASER
SATELLITE TRACKERS
LONG RANGE SAM
OTH
● COBRA DANE
● PAVE PAWS
TBM SITE
TBM SITE
DEFENCE
C^2
SHORAD
T.B.M.
PASSIVE SENSORS
C^2
SHORAD
ACTIVE RADARS
C.A.S.
I.D.S.
VACATED MOSAIC
REFORMING MOSAIC
GROUND THREAT
AIR THREAT
BARRIER SENSORS
MOSAIC

(19)

Ref.No: M7101/DH/1

SYSTEM SURVIVABILITY SCIENCE

Prepared for EASAMS

by

PROFESSOR DEREK HITCHINS

of

CITY UNIVERSITY

Ref. No: M7101/DH/1

SYSTEM SURVIVABILITY SCIENCE

Prepared for
EASAMS

By
PROFESSOR DEREK HITCHINS
of
CITY UNIVERSITY

1 THE BASIC SURVIVABILITY EQUATION

The approach is to formulate a stochastic equation based on contingent probabilities. This leads to a separate threat and counter-threat (survival) categorisation.

- Interaction between enemy perception and stance vs own vulnerabilities
- Process must separate features into enemy threat and own survival capability
- Probability approach gives basis for national separation/categorisation
- Shows multi-faceted survivability, related to enemy and to own performance.

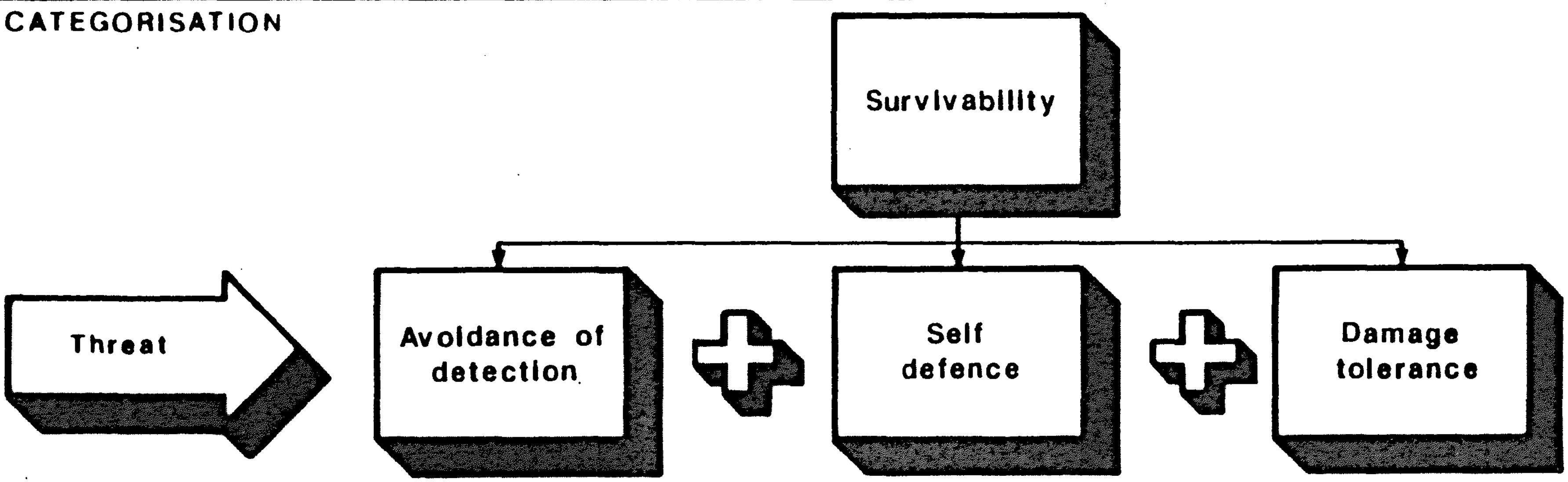
Basic Survivability Equation - Contingent probabilities reveals a coherent structure of threat and counter-threat features.

PROBABILITY OF DAMAGE (1-D) = PROBABILITY OF DETECTION (P_{DT}) * PROBABILITY OF ATTACK (P_A) * PROBABILITY OF BREAKTHROUGH (P_B) * PROBABILITY OF DEBILITATION (P_{DB})

PROBABILITY OF DETECTION (P_{DT})	PROBABILITY OF ATTACK (P_A)	PROBABILITY OF BREAKTHROUGH (P_B)	PROBABILITY OF DEBILITATION (P_{DB})
• Stealth • Deception • Camouflage • EMCON/LPI • Mobility/ Evasion	• Threat resources • VP Perceived worth • Military stance • Intelligence/ LPI • Weapon effects/ Target density	• Hardness • Security • Self-defence weapons • Tactics • Mobility	• Nodes • Reversions • Delegation • VP Density • Functional replication • Control/ repair

EQUATION

CATEGORISATION



THE BASIC SURVIVABILITY EQUATION

2 AVAILABILITY/SURVIVABILITY PARADIGM

A basic mathematics can be developed which mirrors the established Availability mathematics. The new mathematics assumes that the interval between successive attacks is random. As with Reliability, this assumption needs to be tested in particular cases.

- Availability mathematics based on m (Mean Time Between Failures) and MDT(F) (Mean Down Time due to a Fault)
- For Survivability introduce "d" (Mean Time Between Damage Inflicting Attacks) and MDT(D) (Mean Down Time due to inflicted Damage)
- Then the equivalent of Reliability is Durability,
  the equivalent of Maintainability is Repairability
- Approach founded on assumption that the interval between attacks is random - reasonable in complex scenarios.

Survivability Mathematics - Availability mathematics provide a model for survivability calculations.

AVAILABILITY PARAMETERS

- Reliability (R) = EXP (−t/m)

  m = Mean time between failures

  t = Elapsed time

- Availability (A) = $\dfrac{m}{m + MDT\ (F)}$

  MDT (F) = Mean down time due to a fault

- Maintainability : ease of fault repair

MDT (F) = Fault detection and location time + Fault repair time + Recovery time

SURVIVABILITY PARAMETERS

- Durability (D) = EXP (−t/d)

  d = Mean time between damage-inflicting attacks

  t = Elapsed time

- Survivability (S) = $\dfrac{d}{d + MDT\ (D)}$

  MDT (D) = Mean down time due to inflicted damage

- 'Repairability': ease of damage repair

MDT (D) = Damage detection and location time + Damage repair time + Recovery time

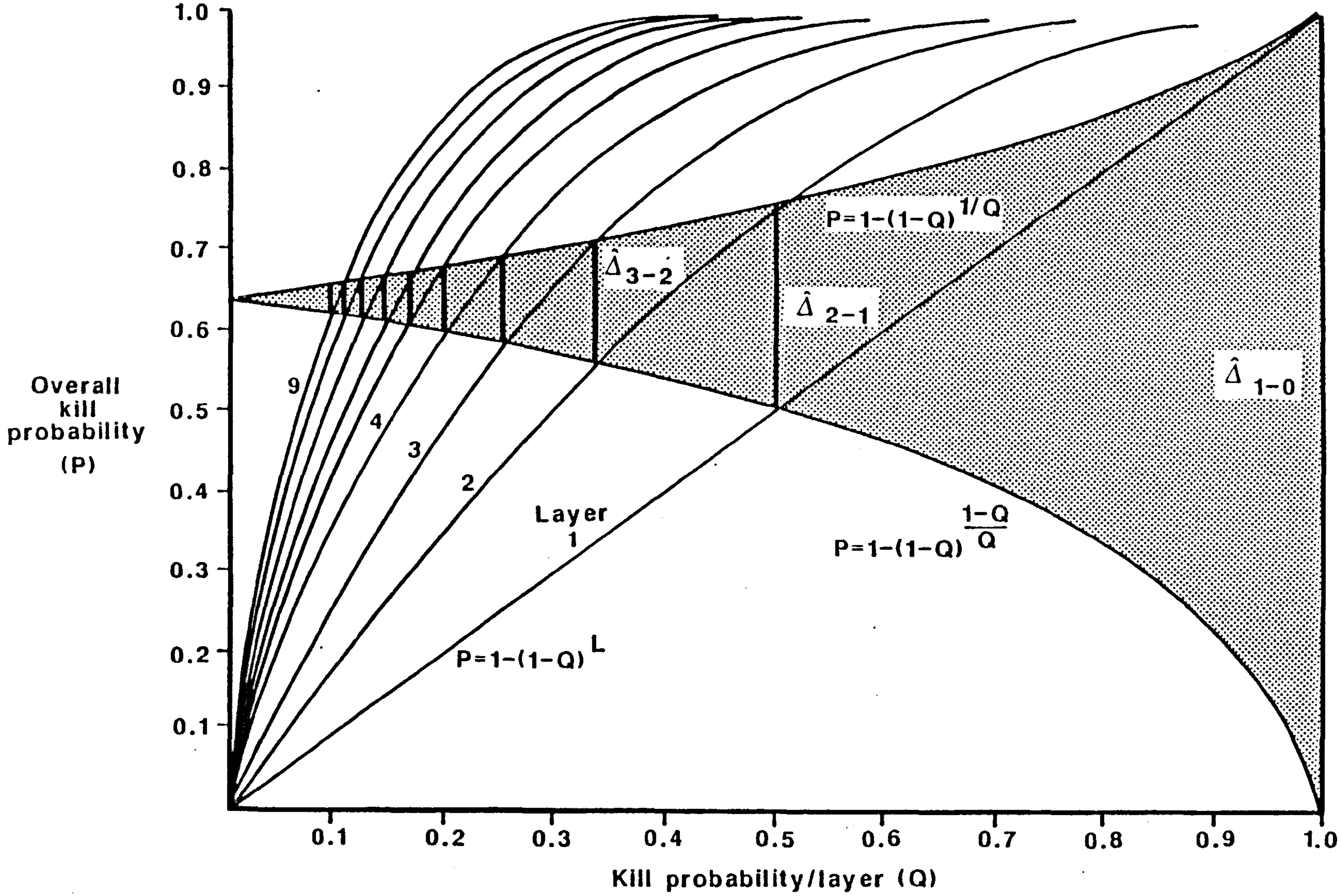
AVAILABILITY/SURVIVABILITY PARADIGM

EASAMS

3 LAYERED DEFENCE MATHEMATICS

The use of sequential layers to improve intercept performance is practiced. Probabilistic considerations show regions where loss of a layer due to attack would make greatest inroads into system performance and hence survivability. Notional single layer performance suggests between three and six layers as sensible.

- No one defensive barrier can be 100% effective - so we consider multiple barriers which targets pass through sequentially
- The graph shows lines relating overall kill probability to single layer kill probability (Q) for multiple layers (L) - law of diminishing returns as more layers added
- Overlay shows maximum increment/decrement loci for either adding another layer or losing one in combat - decrements reduce as number of layers increases
- Practical performance parameters for individual layers quite uncertain - anywhere between 5% and 95%. Choose 50% mean (insensitive area - 35% - 75% gives substantially same answer)
- Conclusions:
  - One layer inadequate
  - Two layers highly vulnerable to losing one layer
  - No value beyond six layers
  - Therefore $3 < L < 6$.



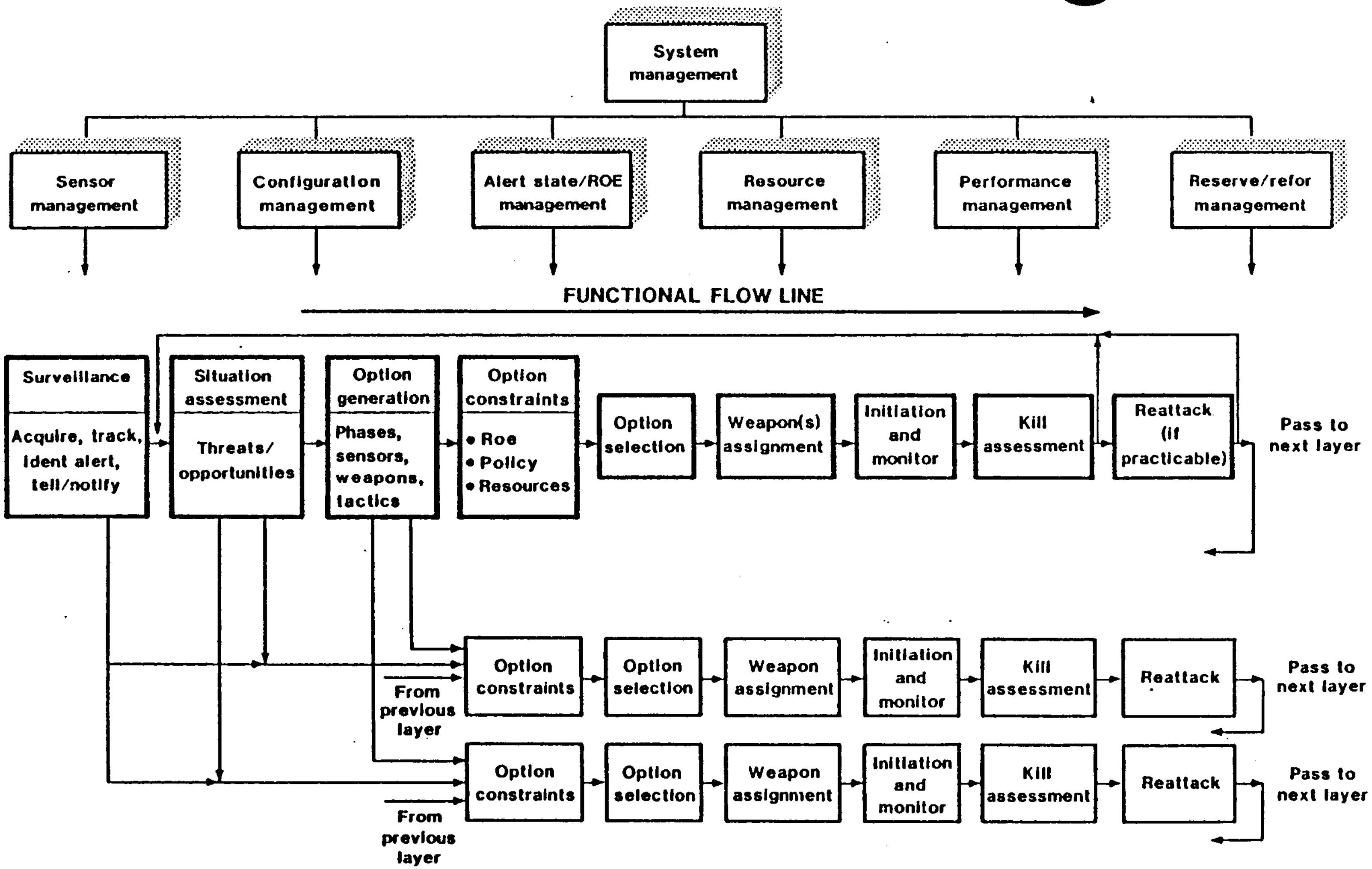
LAYERED DEFENCE THEORY AND THE MAXIMUM INCREMENT/DECREMENT ZONE

FUNCTIONAL SURVIVABILITY

Timelines generate sequential functions coupled through information exchange. Survival equates to maintenance of the sequence and its information flow. System management aids both performance and hence survivability. Functional-to-physical mapping is the crucial step in design for survivability and the key element is 'functional redundancy'.

- Each layer follows the same sequence of events from examining intercept options through to attack and reattack - if possible
- Each function must survive in a layer, for a layer's performance to survive
- Managing one intercept system or one layer or one battle invokes the transverse features, i.e. sensors, configuration, resources, Rules of Engagement (ROE), performance, etc.
- The key to survivability is mapping these multi-layered functions to a physical architecture that can sustain repeated damage without loss of functional sequence - hence replicated functions, or 'functional redundancy'
- Timelines suggest minimum number of layers consistent with performance. Hence from probablistic considerations, three/four layers are optimal and cost-effective.

Functional Survivability - Functions are time sequential and survival equates to maintenance of information flow between functions.



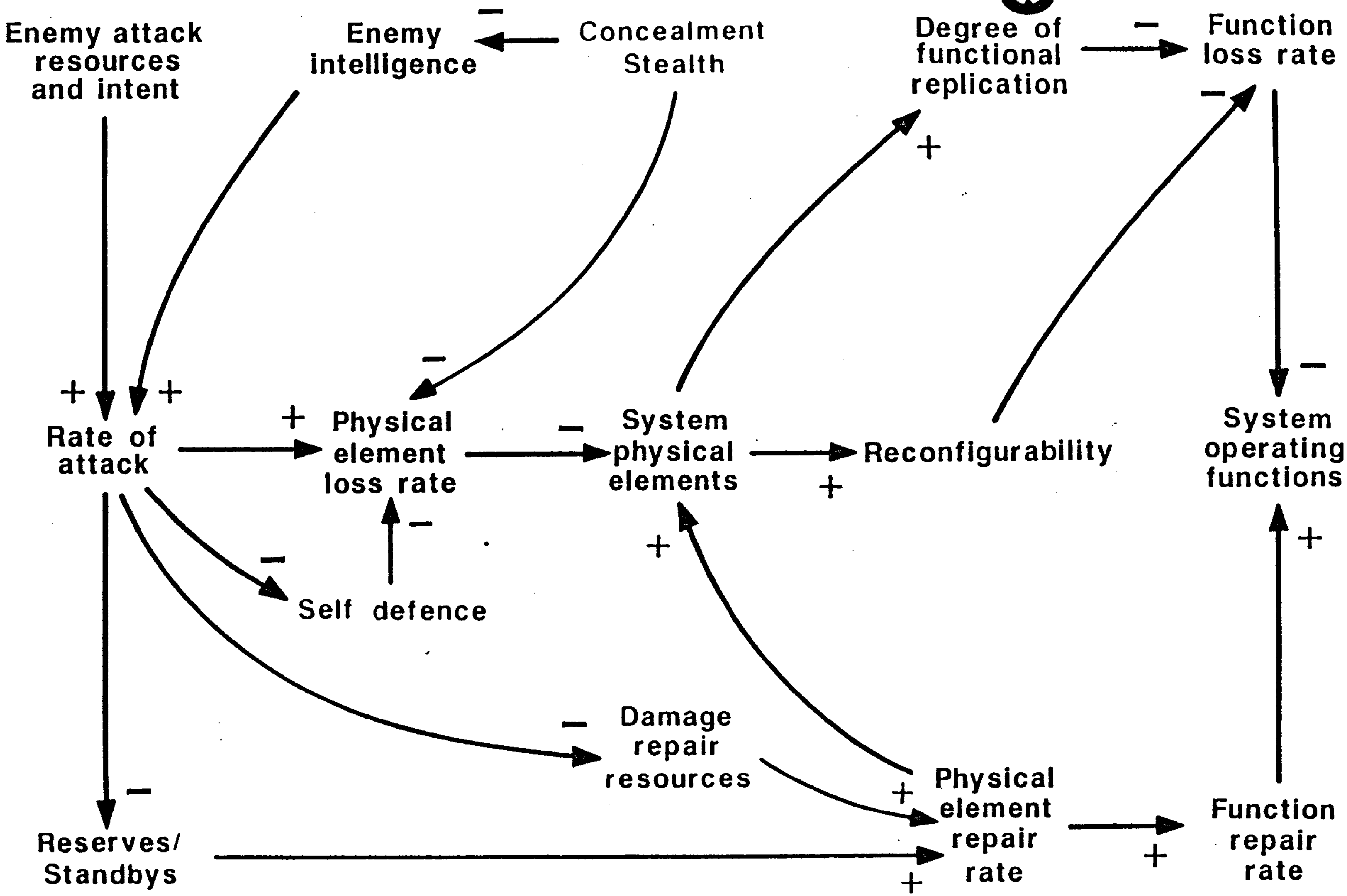
- **FUNCTIONAL SURVIVABILITY** – Physical facilities will be overrun
- **OPERATOR CONTROL** – Full automation is unacceptable
- **FUNCTIONAL-TO-PHYSICAL MAPPING** is key BMC3 design task

FUNCTIONS AND FUNCTIONAL SURVIVABILITY

5 BASIC SURVIVABILITY MODEL

The interaction between the components of survivability is complex and is unquestionably a subject for modelling and simulation. As the simple influence diagram shows, such a model/simulation should be dynamic and comprises three parts: a physical/functional map, a physical survivability component and a threat component.

- Simplest model is system dynamics influence diagram
- Objective is to maintain system operating functions. These are provided within physical elements, but there need not, indeed should not, be a one-to-one mapping
- A high degree of functional replication, coupled with reconfigurability can reduce the rate at which functions are lost - and damage repair can reduce the effect further or even reverse it
- Enemy interaction can occur at many points. Avoidance of detection seeks to deny intelligence, including targetting data. Self defence, itself vulnerable to attack, protects physical elements, and damage control and repair may be subject to enemy attack, too.



BASIC SURVIVABILITY MODEL

6 SYSTEM SURVIVABILITY MEASURES OF PERFORMANCE/MEASURES OF EFFECTIVENESS

Measures of Performance (MOPs) and Measures of Effectiveness (MOEs) for survivability are not generally well formulated, but practitioners seek quantitative indices. The overall 'battle survivability' measure can be broken down into three component ratios of merit, with physical-to-functional loss rate being key.

- Overall survivability is measured by the ability of the system to perform its function in the face of attack. Since repair is possible during/after attack, the situation should be viewed dynamically
- The overall ratio can be broken down into (at least) three component ratios
- Architecturally, the most interesting is the ratio of Vital Point Losses to Function Losses. This ratio contains the essence of design for survivability and it can be seen in isolation from specific details of enemy capability.

Architecture Survivability - The overall index relates area attack to loss of system function. The ratio of Vital Point Loss to System Function Loss is the key architectural design driver for survivability.

Overall system survivability measure =

$$\frac{d(AA)/dt}{d(FL)/dt}$$

Where :

$d(AA)/dt$ = Rate of area attack

$d(FL)/dt$ = Rate of system function loss

Then :

$$\frac{d(AA)}{d(FL)} = \underset{(a)}{\frac{d(AA)}{d(VPA)}} \times \underset{(b)}{\frac{d(VPA)}{d(VPL)}} \times \underset{(c)}{\frac{d(VPL)}{d(FL)}}$$

Where :

VPA = Vital Point Attack

VPL = Vital Point Loss

Ratio (a) Measures avoidance of detection

Ratio (b) Measures self defence

Ratio (c) Measures system damage tolerance

RATIO (c) IS THE PRINCIPAL DESIGN DRIVER FOR ARCHITECTURAL SURVIVABILITY

"ARCHITECTURAL SURVIVABILITY IS ENHANCED WHEN THE IMPACT OF PHYSICAL DAMAGE ON SYSTEM FUNCTIONAL CAPABILITY IS MINIMISED"

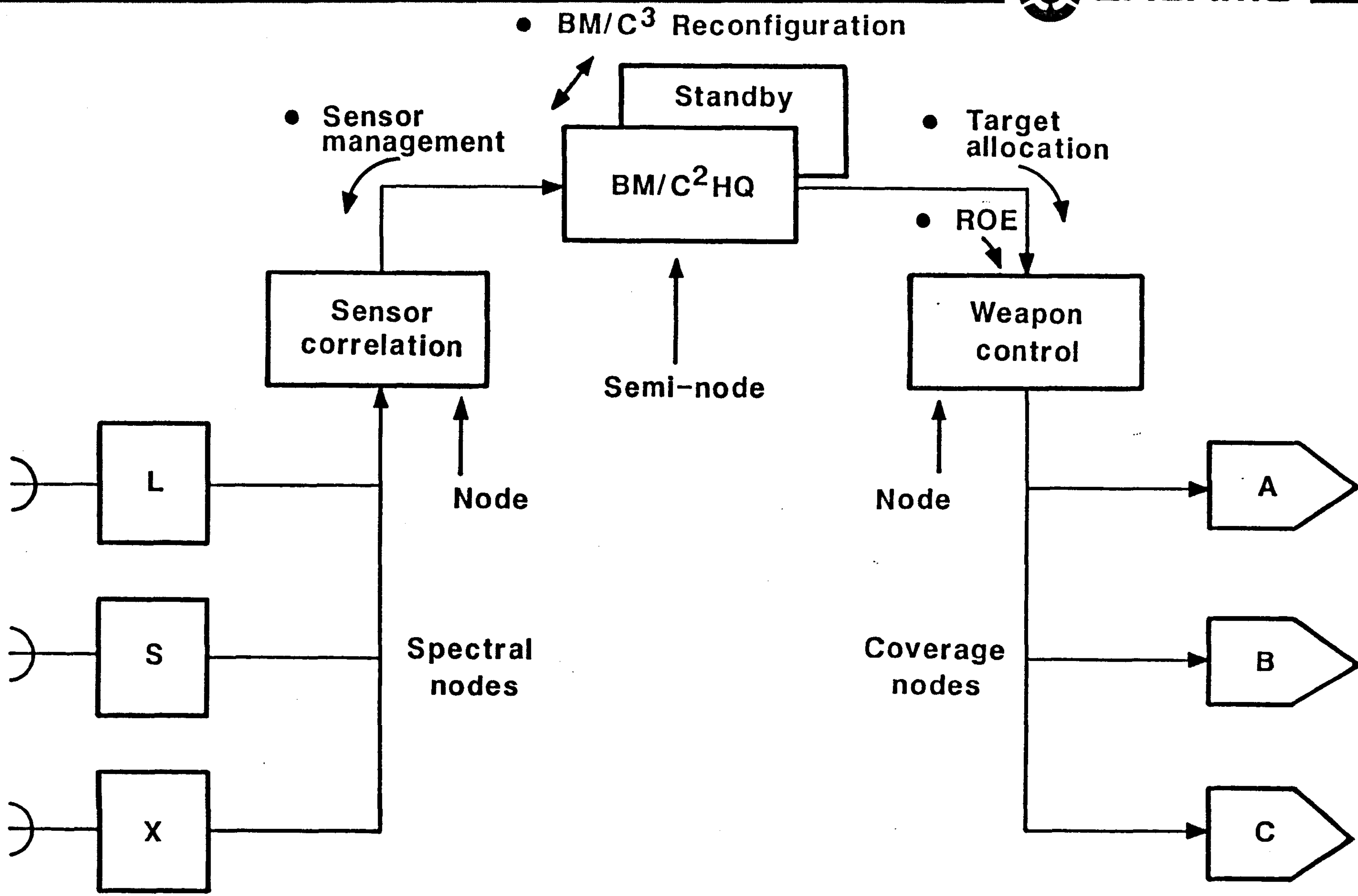
ARCHITECTURE SURVIVABILITY

7 MONOPOLISTIC ARCHITECTURES

Contemporary Battle Management, Command, Control and Communications (BM/C^3) system designs are low survivability. Cost and reluctance to delegate may be causal. Often their nodality is believed to have been overcome by physical replication, but organisational, spatial and spectral nodes may still abound.

- Contemporary Defensive Systems (CDS) pay lip-service to survivability - expensive, complex, etc.
- CDSs also have inherent functional nodes, even where physical nodes may have been replicated
- The figure shows physical nodes - sensor correlation/fusion, weapon control - and functional nodes - reconfiguration, target allocation and sensor management
- Physical replication is valueless on its own, e.g. BM/C^2 standby must have full functional capability.

Note: Sensor spectral nodes - enemy visible in only one band is observed via a sensing node.



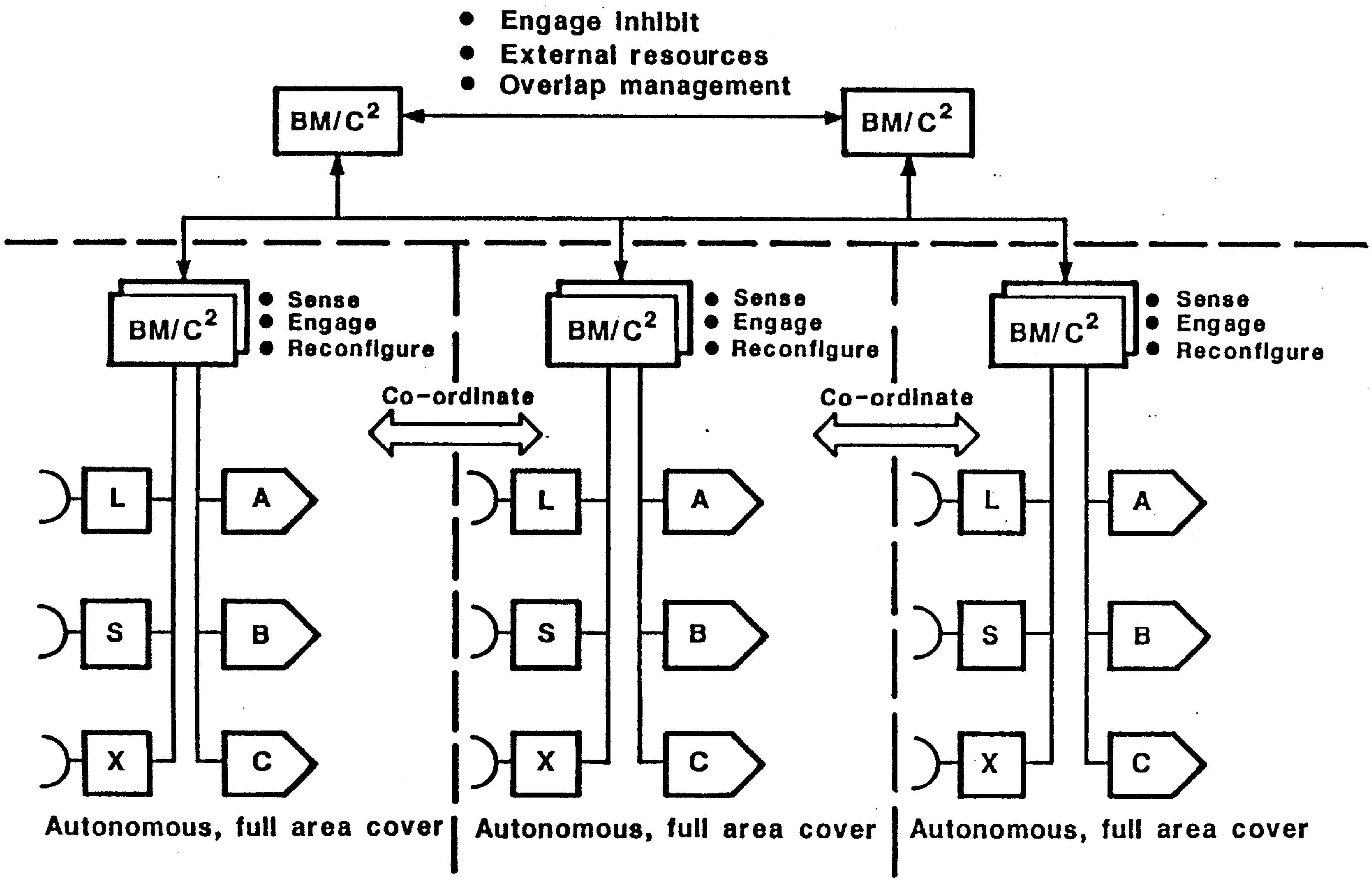
FUNCTIONAL MONOPOLY

FUNCTIONAL DELEGATION

Survivability of the system requires bottom-up design. Redundant autonomous units are provided with superior BM/C^2 only to inhibit their autonomy. In extremis, each fighting unit can operate alone (ultimate system survivability).

- In the figure, three autonomous, self-contained and complete fighting systems are shown. Each could provide some capability throughout the whole sphere of action
- Superior BM/C^2 provides resources, co-ordination and the 'engage inhibit' or 'do not fight' control. In its absence, all fighting systems will continue to fight
- Superior BM/C^2 is replicated (it has few functions and is relatively inexpensive) but should its functions be neutralised, they are replicated again directly between the fighting systems.

Bottom Line. Survivability means spatial, spectral, functional and engagement redundancy and overlap. Above all, it starts at the bottom with autonomous units and provides superior control only to inhibit their autonomy.



FUNCTIONAL DELEGATION

Systems creativity

D.K. Hitchins, MSc, CEng, FIEE

Indexing terms: Engineering administration and management, Project and production engineering, Management

Abstract: The paper identifies the important area of creativity in the design process as least amenable to definition, being comprised of seminal ideas which must fall into a fertile ground of expert domain knowledge coupled with a progressive management environment. Creativity is presented as an essential element in design which can be taught and encouraged, however, and the paper presents creative design 'seeds' for decision-based systems, flow-line manufacture and maintenance processes and management information systems. Concept evolution techniques are presented as a structured decomposition methodology. The first example of creative design presented in the paper is of the design of a notional systems engineering company which is unusual to most readers in that it contains no manufacturing or production, and preconceived ideas of company organisation are likely, therefore, to be misleading. The application of state-variable analysis to the cash flow of a start-up defence company follows, and the application of systems dynamic to intercompany competition is presented. Creativity is not restricted to initial design; flow-line analysis using queuing theory for intensive flying of training aircraft is demonstrated as a means of resequencing activities to reduce mean throughput times. And finite state/transition analysis is used on the grand scale to examine the potential for European conflict. Finally, the author presents the principles of creativity and suggests that, far from being peculiar to engineering, one finds parallels in music, composition and the arts.

Introduction

The spark which represents a creative idea or concept often passes unnoticed, and dies unrecognised. Industry needs such ideas; they need to be encouraged, nurtured and used as the seeds of our future products, services, structures, organisations and capabilities. Among engineers, however, the idea of creativity might conjure up some artistic concepts which they would consider alien to their craft.

Systems creativity seeks to present the process of conceiving, designing, developing and implementing creatively in an engineering environment. Identifying and nurturing the spark necessitates the right organisation, techniques, tools and tenets, the principal elements of which are presented by example. This paper is the sequel to Reference 1.

Paper 6055A (M3, M4), first received 23rd June and in final form 7th September 1987

The author is Business Development Director at EASAMS Ltd., Lyon Way, Frimley Road, Camberley, Surrey GU16 5EX, United Kingdom

1 Ingredients of creativity

Creativity is a blend of many things: youth and experience, discipline and freedom, knowledge and serendipity. These elements, coupled with energy and ability, and set in the right environment, are the basic ingredients of creativity. Luck? Luck is opportunity meeting preparation.

Creativity is a difficult concept with which to grapple, especially in an engineering environment. Certain ingredients are often seen in retrospect to have combined to result in a seminal idea, design or solution to a problem. Fig. 1 shows these ingredients, and, while some may

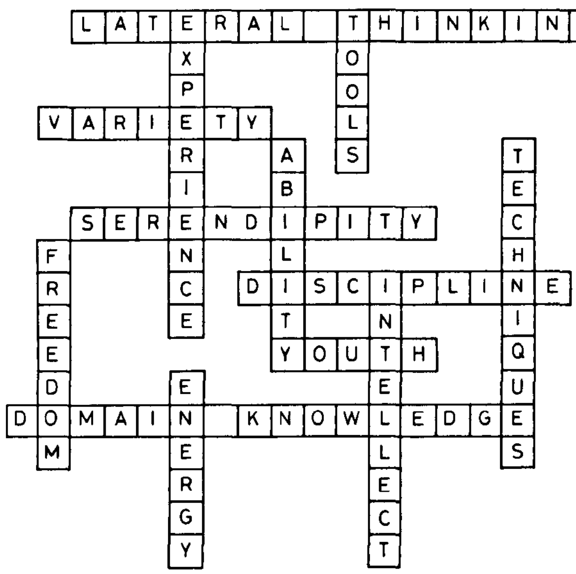


Fig. 1 *Ingredients of creativity*
Creativity requires a blend of many constituents, set in a disciplined framework

appear to be in conflict, in practice they go together as follows:

(*a*) *Experience and youth:* A blend of youthful energy and ideas is a valuable basis for generating creative solutions, provided it is tempered by sound experience. Experience alone will rarely be innovative. Mixed teams give the best results.

(*b*) *Discipline and freedom:* There is a time for both; freedom of thought and expression, positively encouraged and set in a formally disciplined procedure, will identify and find solutions to problems quickly. Brainstorming is a form of disciplined freedom.

(c) *Domain knowledge and serendipity:* Domain knowledge implies a good understanding of the sphere of endeavour in which the solution to a requirement must be created. Serendipity implies chance or luck. And both have a valuable role to play. The good fortune which stumbles across an idea must recognise the idea and view it in the context of its own domain.

The remaining ingredients of the nonexhaustive list are also important. Ability coupled with intellect is essential. Some are creators, some are improvers, some are critics; all have their role. But the creators will invariably be able and energetic. To be effective in a sophisticated engineering environment, these creators need force multipliers; tools and techniques allow creators to be more productive and co-operative.

2 In pursuit of creativity

Creativity can be incorporated into a systematic procedural framework, the so-called top-down systems approach. Optimising the preferred solution is often left out of this otherwise well-established process.

Fig. 2 shows an almost classic, systematic approach to finding a solution to the problem at hand. A top-down approach requires a complete understanding of the requirement and the objectives. So often this solution-transparent requirement analysis is ruined by premature assumptions: 'we'll use our standard product . . .', 'what we did last time was . . .', 'sell the hardware first, we will sort out maintenance later . . .'.

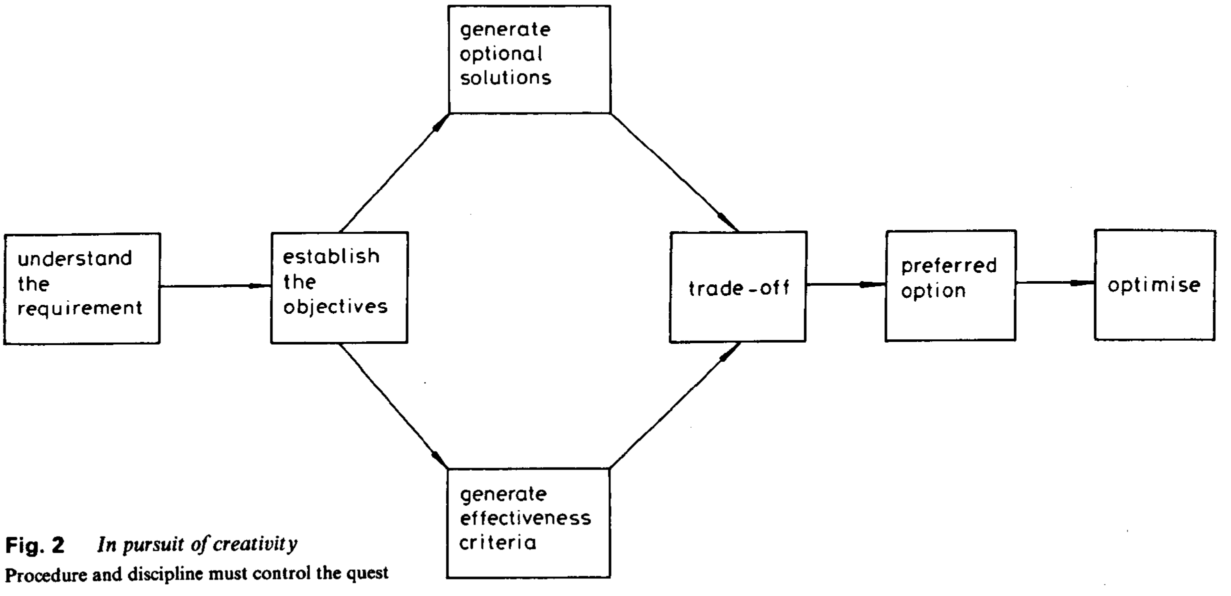


Fig. 2 *In pursuit of creativity*
Procedure and discipline must control the quest

The generation of optional solutions to the requirement should occur ideally in isolation from the generation of effectiveness criteria, and strict discipline is needed during trade-off to avoid 'massaging' effectiveness criteria to suit preconceived notions about the best answer. Preconceptions are the enemy of creative innovation. Trade-off techniques have already been presented in Reference 1.

The final box in the Figure is the one most often overlooked. It would be rare indeed for the preferred option to excell against all criteria; indeed, such a result implies either that the problem was trivial or that the optional solutions were inadequate. Following topics will show by example how this systematic problem-solving approach is applied in practice and how optimisation is achieved.

3 Creative seeds

Most creative ideas in engineering require a seed or kernel of an idea to form effectively. The decision cycle and the flow line are two common kernels for structure and organisation systems.

Real solutions in engineering can very often be formed about a conceptual seed or kernel. The two most common, the decision cycle and the flow line, are shown in Fig. 3A.

(a) *The decision cycle:* Decisions follow a natural progression in many real situations, starting with an assessment of the situation at the top of Fig. 3A. The ubiquitous decision cycle is at the heart of management, command and control, driving a car, running a football team, structuring an expert system etc. A company board, for example, shares the decision cycle elements; economic threats are identified by the financial director, while constraints may be presented by the operations director, and so on. Thus, structured decision-making can be constructed around the decision cycle. The decision cycle can also be straightened out, in which case it becomes a flow line.

(b) *The flow line:* In process control of information flow, products or concept evolution, a central line of flow is invariably present. Fig. 3A shows the central line in a computer integrated manufacturing (CIM) flow line (after an idea from DEC), in which components are bought, received, inspected, assembled, stocked, distributed and sold. All elements not on the central path have as their purpose the maintenance and enhancement of flow. In the example, sales information is used to plan future production, and marketing identifies the need for new products. By working outwards from the flow line it is possible to organise commercial, personnel, accounting and all other company functions in support of the central theme, so providing unity of effort.

The vigour of these simple, even simplistic, concepts can be remarkable and this is due, in no small part, to the ease with which everyone in a group can relate to the decision cycle or the flow line. Retention of simplicity is more difficult (but a valuable goal) as the experts and

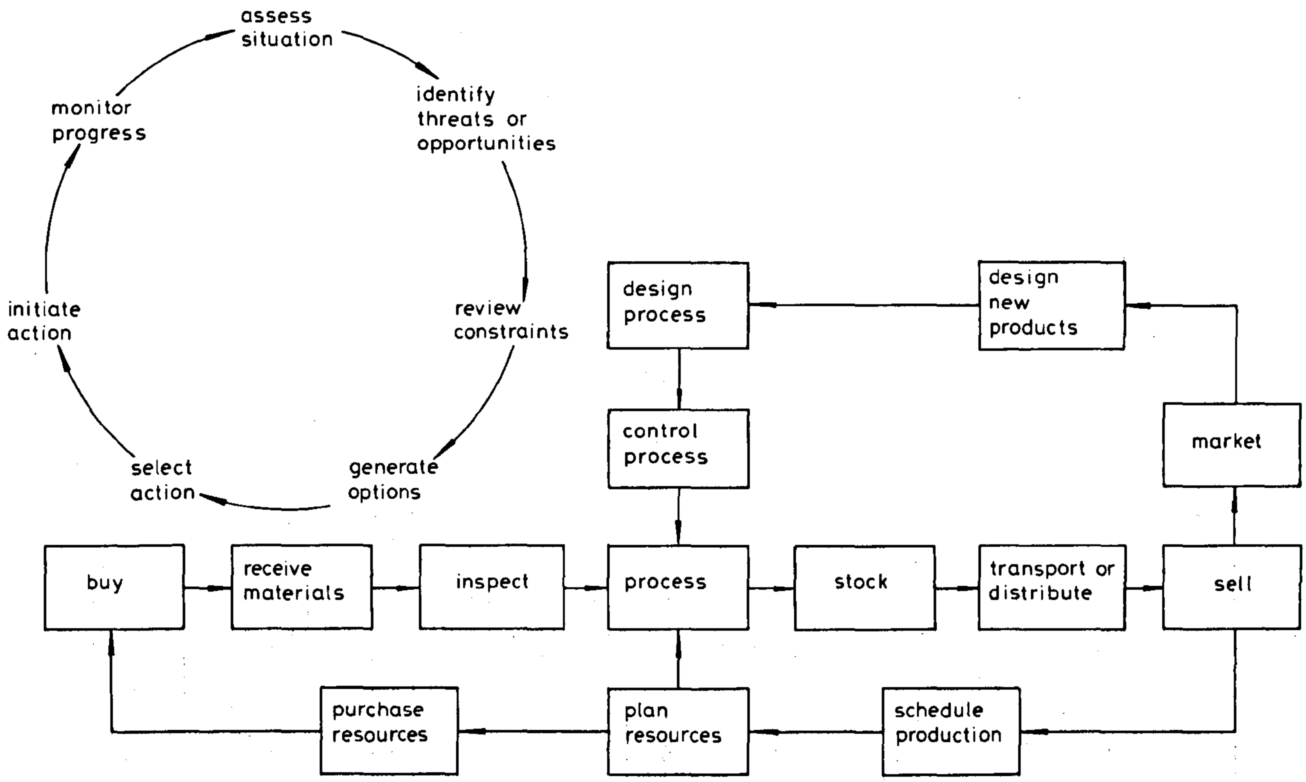


Fig. 3A *Creative seeds*
Decisions and the flow of activities and processes are inherent in all creations, be they artefacts or organisations

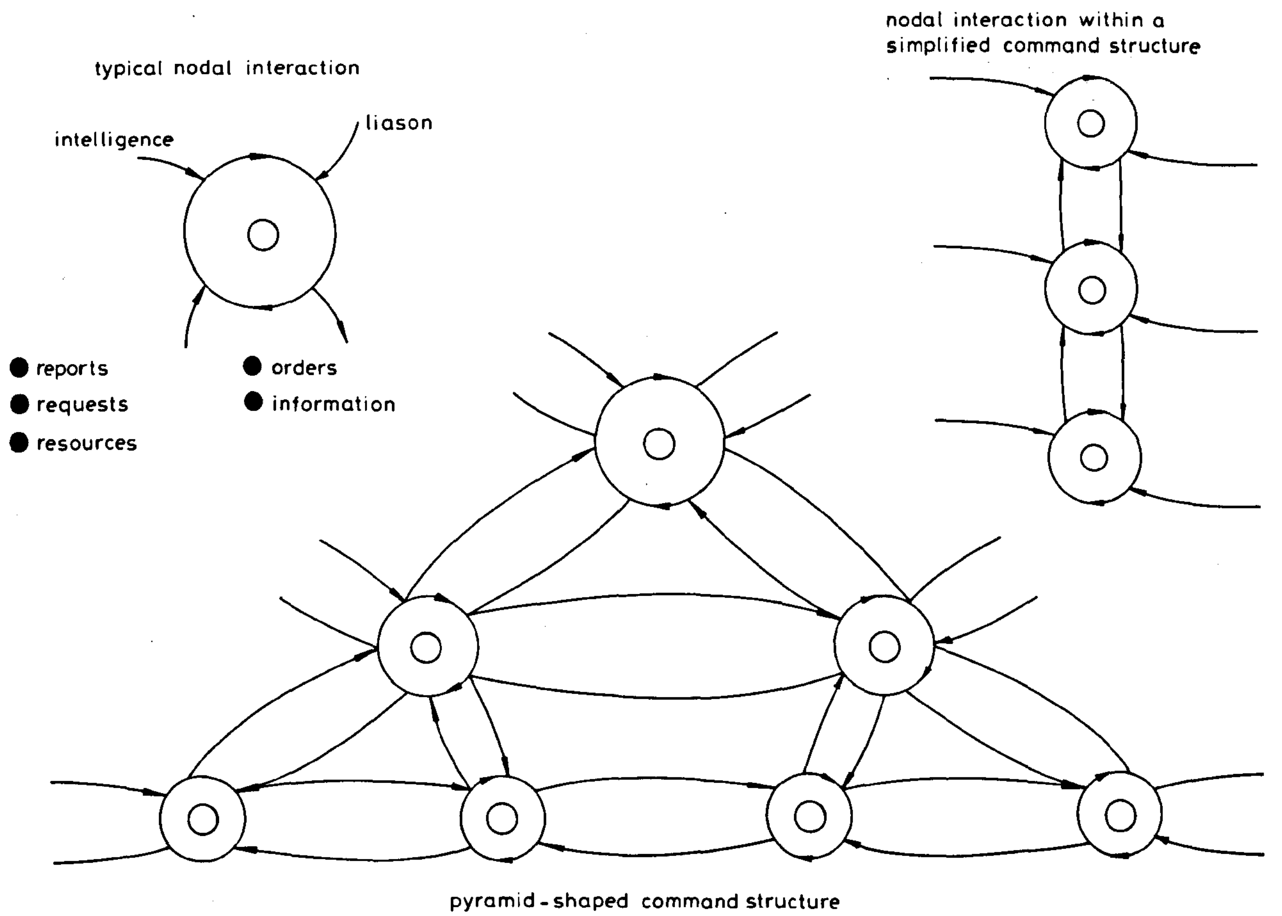


Fig. 3B *Linked decision cycles*
Decision cycles from discrete elements within an organisation can be linked into networks. The representational technique is better replaced by an N^2 chart

	a	b	c	d	e	f	g
1	joint operations centre	• orders • information		• orders • information		• orders • information	
2	• reports • requests • resources	land operations centre	• orders • information	• liaison • co-ord.		• liaison • co-ord.	
3		• reports • requests • resources	mobile land HQs				
4	• reports • requests • resources	• liaison • co-ord.		naval operations centre	• orders • information	• liaison • co-ord.	
5				• reports • requests • resources	naval task groups		
6	• reports • requests • resources	• liaison • co-ord.		• liaison • co-ord.		air operations centre	• orders • information
7						• reports • requests • resources	air bases/ radars

outputs

inputs

typical sub-cell structure

Fig. 3C *N^2 network*

Representing linked decision cycles in an N^2 chart allows a high level of abstraction to be maintained, while providing a powerful tool for architectural analysis

specialists start to fill in detail and suggest alternatives. Obscuration of the fundamental concept will defeat the creative concept.

Figs. 3B and 3C show how decision cycles may be linked to represent a decision hierarchy, and how that hierarchy can be better represented using an N^2 chart. The example chosen shows a joint command and control information system can be constructed from linked decision cycles, each representing an individual command post or HQ.

4 Concept evolution

There are certain fundamental techniques used to evolve concepts which experience shows to provide the most effective results. In practice, it is very difficult to adhere to these tenets in the face of pragmatic short-term solutions.

System design is based on certain fundamental tenets which give the approach its strength. Five of the most important tenets are shown in Fig. 4.

(*a*) *Highest level of abstraction:* It is vital to avoid being confused by irrelevant detail at the early stage in concept formulation. A car is, at the highest level, a means of transporting people by road. In the same vein, what is a tank?

(*b*) *Functional before physical:* Functional decomposition requires the most careful selection of functions. To partition a human (sic!) into senses, nerves and brain is to employ physical, not functional partitioning. To iden-

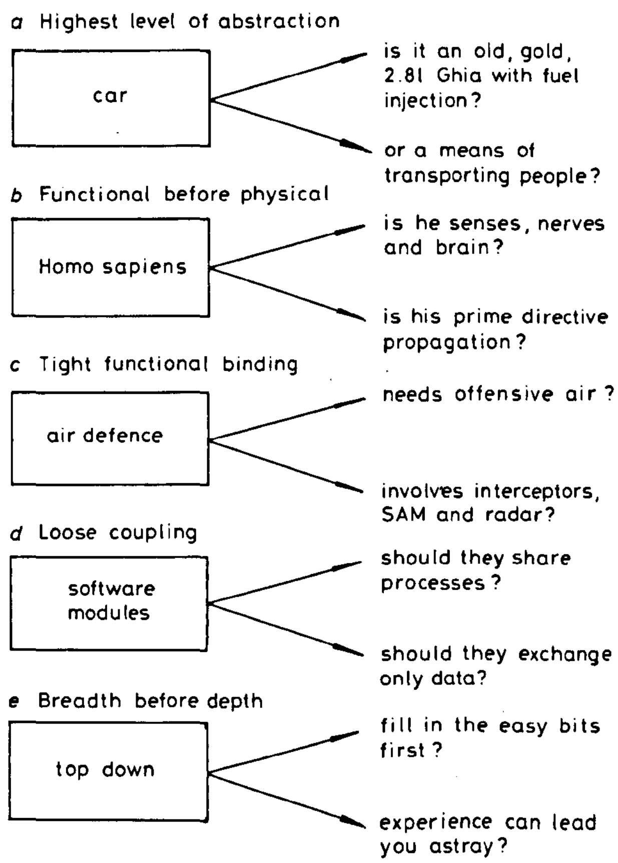


Fig. 4 *Concept evolution*

A methodology is essential to the evolution of concepts if we are to be creative, not repetitive tify human functions as propagation, survival, communication and co-operation is to provide a basis for understanding why we have such types of sensor, nerves and brain. To partition a distributed information system (DIS) into sources, communications, processing and sinks is to overlook its purpose; it is not incorrect, simply unhelpful.

(*c*) *Tight functional binding:* It is valuable to group together people, resources, features and skills which have the same objectives and many mutual interfaces. This concentrates effort, clarifies, simplifies and economises.

(*d*) *Loose coupling:* Conversely disparate groups of people, resources, features and skills with different objectives should be coupled loosely (separated), generally by the exchange of carefully defined data. They should not share functions.

(*e*) *Breadth before depth:* This tenet is linked with the first — high level of abstraction. It is vital to span the 'complete' problem at the higher level, before moving to the next level of detail. So often, it is all too easy to dive into that part of the problem we believe we understand, emerging after much (nugatory?) effort to find that we have been pursuing the wrong path. Worse, we may have invested so much effort that there is no going back, mentally or practically.

5 Creating the creative framework

This and the three following topics illustrate the systems creativity approach by postulating the requirement for, and producing a resulting design of, a systems company. Familiarity with organisation is seen as the major obstacle to a good solution.

If the system approach is sound and general purpose, it should be applicable over a wide range of endeavours, from the Severn barrage to a PCB design, from exploring the solar system to photo-microlithography. As a simple example, we can turn the systems approach in upon itself (Gödel's incompleteness theorem undoubtedly applies) and use it to design a systems company's divisional structure. This approach to creating a creative framework inside which creative ideas may be nurtured is outlined in Fig. 5. The trap to be avoided is overfamiliarity with conventional organisation: we know the solution to similar problems too well.

Familiarity can present real difficulties. For example, how do we functionally divide a house? If we choose, say, bathroom, bedrooms, lounge, dining room and kitchen as the obvious partitions, then subdivision in each room will contain much the same elements, e.g. lighting, plumbing, power, central heating etc. Thus, the 'obvious' divisions for a house which are physical, not functional, are unhelpful, because they have resulted in tight coupling and loose binding, the opposite of our objective. Our familiarity with the problem has led us astray. In essence, we did not divide the house by *its* functions; we divided the house by its *occupants'* functions — and that is a common trap. The same hurdle has to be overcome with divisions in a company. The following Section will continue with this 'case study' by generating four archetypal divisional organisations.

- *Create a creative organisation*
- *Design a systems company* (no manufacturing)
- *Objectives*
  - growth
  - synergy between divisions
  - concentrated skills
  - market led
- *Pitfalls*
  - familiarity with conventional organisation
- *Approach*
  - treat as any other system

Fig. 5 *Creating the creative framework*
Because the framework/organisation has to be created, that process is a good basis for a case study

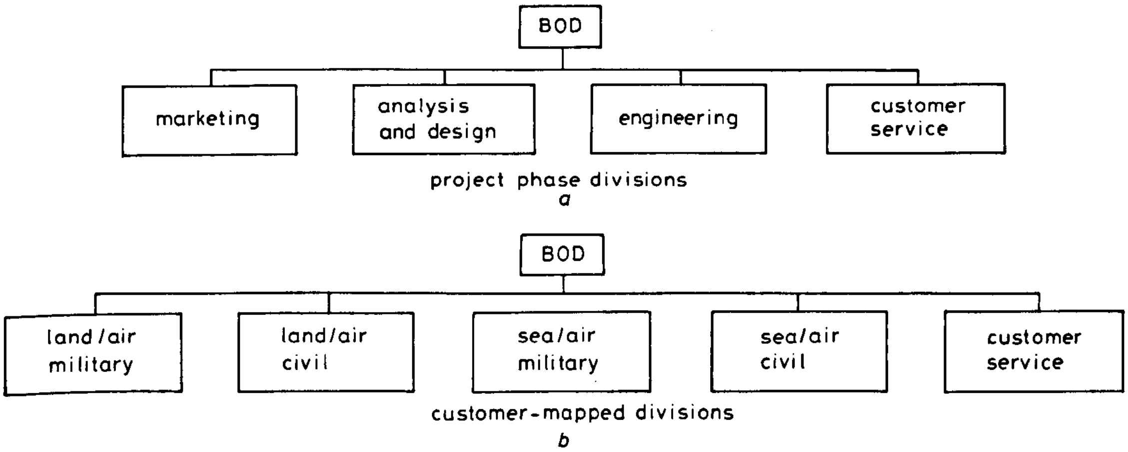
6 Systems company structure options

The four archetypal divisional organisations are outlined at the highest level of abstraction. None of these options is seen as preferred at this stage.

The four archetypal divisional organisations shown in Fig. 6 and are as follows:

(*a*) *Project phase divisions:* This structure is used by most production companies, principally because of the high tooling, stock and WIP investment required for production, which dictates maximum utilisation and inventory turns. A systems company has no large-scale production, however, and so the usual pressure does not apply.

(*b*) *Customer-mapped divisions:* This structure maps directly on to the customer's domains of activity. While excellent in marketing terms, the structure distributes engineers and resources in every division with identical capabilities, and is hence not necessarily effective in implementation, because it requires much lateral flow of engineers between divisions.



tify human functions as propagation, survival, communication and co-operation is to provide a basis for understanding why we have such types of sensor, nerves and brain. To partition a distributed information system (DIS) into sources, communications, processing and sinks is to overlook its purpose; it is not incorrect, simply unhelpful.

(*c*) *Tight functional binding:* It is valuable to group together people, resources, features and skills which have the same objectives and many mutual interfaces. This concentrates effort, clarifies, simplifies and economises.

(*d*) *Loose coupling:* Conversely disparate groups of people, resources, features and skills with different objectives should be coupled loosely (separated), generally by the exchange of carefully defined data. They should not share functions.

(*e*) *Breadth before depth:* This tenet is linked with the first — high level of abstraction. It is vital to span the 'complete' problem at the higher level, before moving to the next level of detail. So often, it is all too easy to dive into that part of the problem we believe we understand, emerging after much (nugatory?) effort to find that we have been pursuing the wrong path. Worse, we may have invested so much effort that there is no going back, mentally or practically.

5 Creating the creative framework

This and the three following topics illustrate the systems creativity approach by postulating the requirement for, and producing a resulting design of, a systems company. Familiarity with organisation is seen as the major obstacle to a good solution.

If the system approach is sound and general purpose, it should be applicable over a wide range of endeavours, from the Severn barrage to a PCB design, from exploring the solar system to photo-microlithography. As a simple example, we can turn the systems approach in upon itself (Gödel's incompleteness theorem undoubtedly applies) and use it to design a systems company's divisional structure. This approach to creating a creative framework inside which creative ideas may be nurtured is outlined in Fig. 5. The trap to be avoided is overfamiliarity with conventional organisation: we know the solution to similar problems too well.

Familiarity can present real difficulties. For example, how do we functionally divide a house? If we choose, say, bathroom, bedrooms, lounge, dining room and kitchen as the obvious partitions, then subdivision in each room will contain much the same elements, e.g. lighting, plumbing, power, central heating etc. Thus, the 'obvious' divisions for a house which are physical, not functional, are unhelpful, because they have resulted in tight coupling and loose binding, the opposite of our objective. Our familiarity with the problem has led us astray. In essence, we did not divide the house by *its* functions; we divided the house by its *occupants*' functions — and that is a common trap. The same hurdle has to be overcome with divisions in a company. The following Section will continue with this 'case study' by generating four archetypal divisional organisations.

- *Create a creative organisation*
- *Design a systems company* (no manufacturing)
- *Objectives*
  - growth
  - synergy between divisions
  - concentrated skills
  - market led
- *Pitfalls*
  - familiarity with conventional organisation
- *Approach*
  - treat as any other system

Fig. 5 *Creating the creative framework*
Because the framework/organisation has to be created, that process is a good basis for a case study

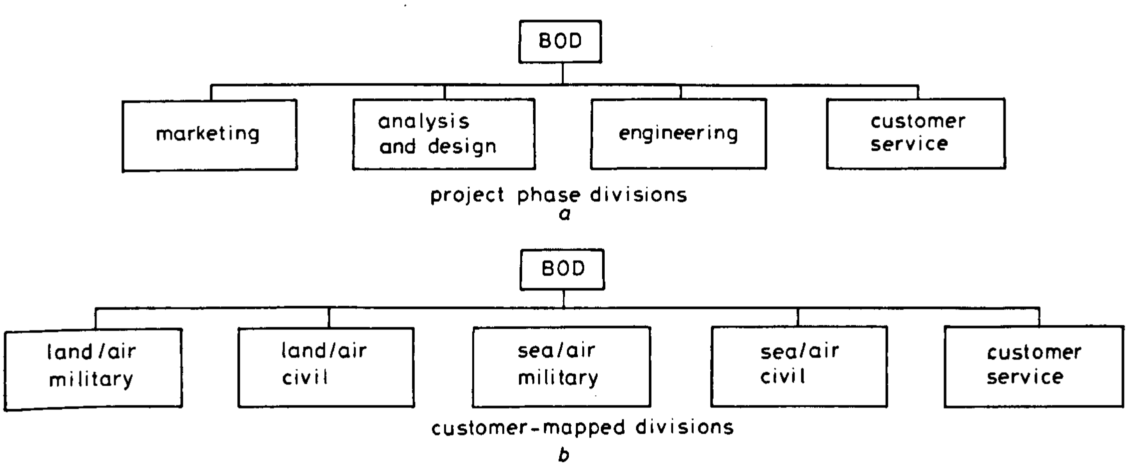
6 Systems company structure options

The four archetypal divisional organisations are outlined at the highest level of abstraction. None of these options is seen as preferred at this stage.

The four archetypal divisional organisations shown in Fig. 6 and are as follows:

(*a*) *Project phase divisions:* This structure is used by most production companies, principally because of the high tooling, stock and WIP investment required for production, which dictates maximum utilisation and inventory turns. A systems company has no large-scale production, however, and so the usual pressure does not apply.

(*b*) *Customer-mapped divisions:* This structure maps directly on to the customer's domains of activity. While excellent in marketing terms, the structure distributes engineers and resources in every division with identical capabilities, and is hence not necessarily effective in implementation, because it requires much lateral flow of engineers between divisions.



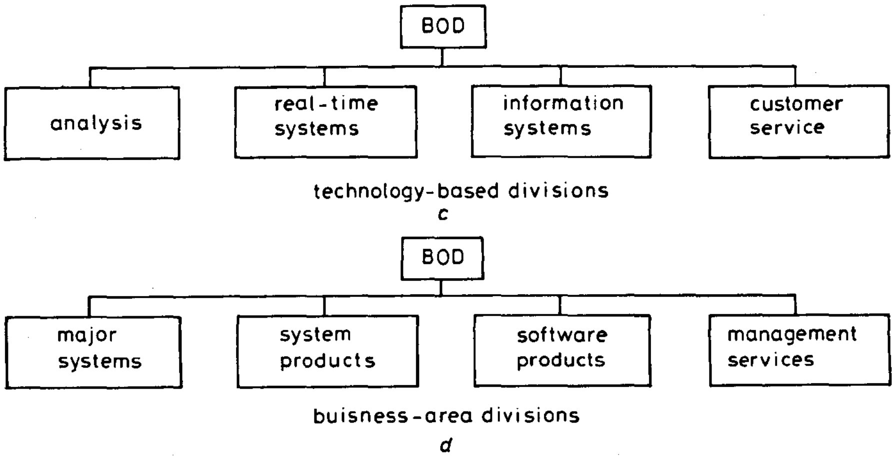


Fig. 6 *Structure options*

At the highest level of abstraction there are four archetypal divisional structures: hierarchical against matrix is not visible at this level. BOD = Board of Directors

(*c*) *Technology-based divisions:* This structure concentrates skills and resources, but discourages synergy between divisions, as they tend to operate in separate customer and business-areas.

(*d*) *Business-area divisions:* This structure maps on to complementary sectors in systems engineering markets, but is, of itself, nonexplicit in terms of marketing and implementation.

None of the organisations is immediately obvious as being the best solution to meet the objectives of growth, synergy, skill concentration and market lead. Further analysis of benefits is clearly needed, first by generating a solution-free set of criteria to judge and compare options.

7 Criteria for a good design

The choice of criteria is fundamental to reaching the best solution; the systems approach is no panacea and can be mishandled. Business-area divisions gave the best solution, but other options have their special merits too.

Choice of criteria is vitally important to drive out the 'right' design solution; as with any approach, it is possible to misapply the techniques. Fig. 7 shows, in a highly visible manner, the criteria and the rank order allocated against each archetypal design. Readers may disagree with both criteria and ranking: it is the objective of the presentation to encourage such results, which will lead to a more solid solution.

Principles of business have been predicated, based on the principles of war (which are not unrelated). Effectiveness criteria have been generated within the principles. Options have been ranked (i.e. 1 is first, 4 is fourth, in order of preference) and inability to differentiate has resulted in equal ranks. Row sums equal 10.

Column sums favour option 4, business-area divisions, with little to choose between customer-mapped and technology-based; project phase is a poor fourth. However, business-area divisions did not rank first against all criteria; first ranks appear in favour of:

(i) technology-based for work flow
(ii) customer-mapped for user/design contact

Principles of Business	Divisional organisation	Options (ranks)			
	Effectiveness criteria	Project phase	Customer-mapped	Technology-based	Business-area
Flexibility	• adapts to new markets	4	2	2	2
	• encourages growth	4	2.5	2.5	1
	• simple, easily understood	3.5	1.5	3.5	1.5
Economy of effort	• non-overlapping business areas	2.5	2.5	2.5	2.5
	• accountable/visible profit and loss	4	3	1.5	1.5
Unity of effort	• synergy between divisions	3.5	1.5	3.5	1.5
	• concentrates resources	1	4	2.5	2.5
Market led	• promotes user/designer contact	4	1	2.5	2.5
	• profitable market potential	3	3	3	1
Resilience	• sufficient regular business	2.5	2.5	2.5	2.5
	• responsive to market	4	2	3	1
Maintenance of the aim	• promotes systems approach	4	2	2	2
	• promotes work flow	4	2.5	1	2.5
Rank sum		44	30	32	24
		fourth	second	third	first

Fig. 7 *Criteria for a good design*

The criteria must be chosen to support the design aim. In this example, the principles of war have been adapted as a basis for addressing business organisational options

(iii) project phase for concentration of (capital intensive) resources.

Optimisation at the next level of detail will incorporate these elements into a preferred solution, where practicable.

8 The preferred solution

The business-area solution is optimised by incorporating the best features from the other structures inside each division as appropriate.

The preferred solution, which best satisfies the requirements of growth, synergy, skill concentration and market lead, is shown in Fig. 8. Business-area divisional structure has been employed, with each division set up differently internally according to need:

Software: addresses software engineering, company-wide and external sales of software products, applications programs, small turnkey information systems and consultancy. Technology-based to promote work flow and concentrate scarce resources, including expensive software development tools. Sales account organised to promote directed growth.

System management services: customer-mapped and with skills concentrated into tightly functionally bound groups of specialists in customer domains and analytical techniques.

Overall, each division is internally bound tightly, while interdivisional synergy is promoted in two senses: there is no business-area overlap, and both system products and software divisions provide resources and products in support of the other divisions (synergy loops). Thus, loose coupling exists too in this structured design.

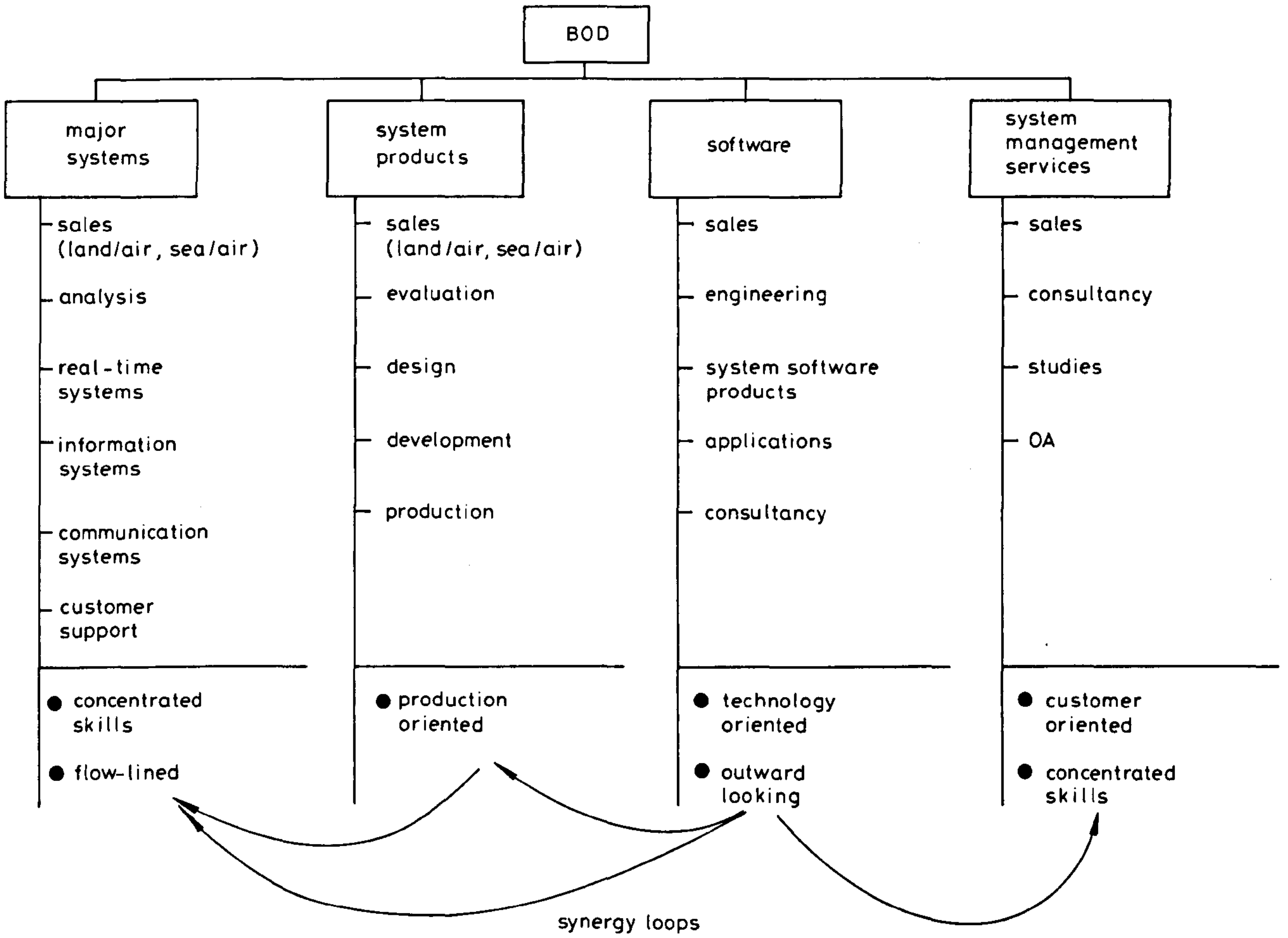


Fig. 8 *The preferred solution*
The structure is business-oriented at divisional level, purpose-oriented within divisions; it is tightly functionally bound within divisions and loosely coupled (via the synergy loops) between divisions. BOD = Board of Directors. OA = Operations Analysis

Major systems: addresses systems engineering projects. Sales organisation is customer-mapped. Resources are technology grouped, and hence skills are concentrated. Project teams are envisaged, dedicated to tasks, with most resources in the division for the complete project from analysis to implementation. The division would probably be matrix-organised internally.

System products: addresses system products — workstations, interfaces, communication switches etc. Production-oriented to meet the high-technology, low-volume throughput. Sales mapped on to customer's domains.

9 State-variable analysis

Company cash flow can be considered as a central problem and addressed using state-variable numerical methods. The results fit observed fact well for start-up companies in the defence business, and present opportunities for different approaches to creative management.

One technique for assessing the dynamics of a company organisation is shown in Fig. 9, in greatly simplified outline. The model, for a systems company, shows a manpower model encapsulated in a surrounding cash-flow model. The company is presumed to be starting from scratch, and capital purchases are considered separately.

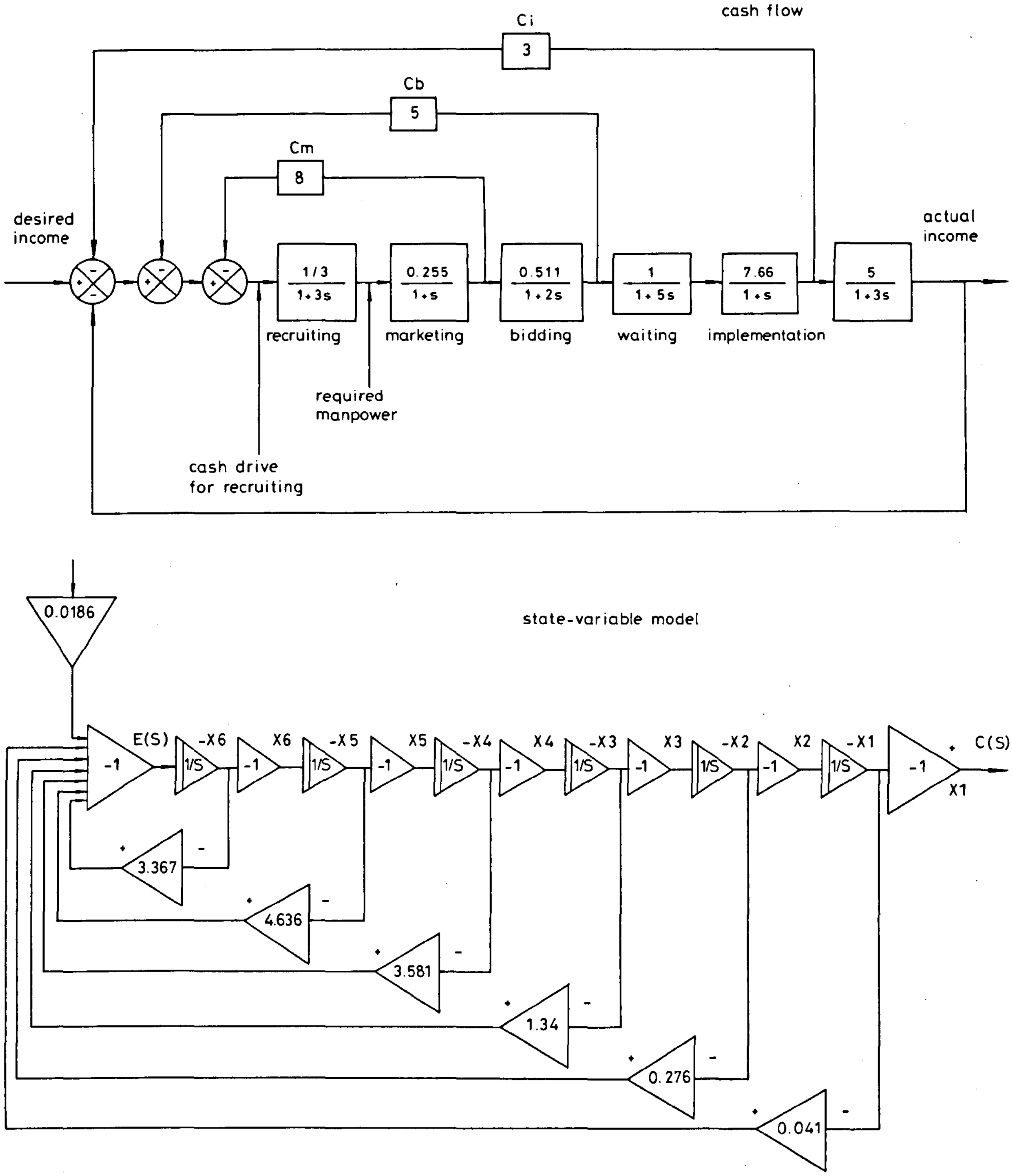


Fig. 9 *State-variable analysis*
The control-loop concept of company economic modelling can be addressed as a classic control-theory problem

Time delays are allowed for recruiting, marketing for opportunities, bidding, awaiting the outcome of the bid, implementation and receipt of first sales revenue. The costs of manpower are calculated and compared with this revenue. Analysis of the cash-flow model is greatly simplified by converting the resulting 6th-order differential equation into six 1st-order equations in the state-variable model.

The state-variable model is simply solved by computer, using recursive numerical methods; the result is nonetheless dramatic for being predictable. There is virtually no income for the first 15 months, after which an upsurge is followed by violent and increasing oscillation, brought about by the long loop delays and the lack of inertia. It is possible to sustain exponential growth, at least in the model, by the simple expedient of preventing $E(S)$, the error, from going negative. The meaning of this in management terms is obscure, but perhaps worthy of investigation.

State-variable analysis is potentially useful, then, in obliging the modeller to home-in on the basic cost and income elements in the company, and in providing a simpler numerical mechanism by which to observe their interplay. Feeding competition and external market forces into the model can also be achieved. The principal value, as the more conventional use of the technique would suggest, is the ability to analyse stability and the impact of discontinuities on system performance.

10 Systems dynamics and competition

Systems dynamics presents valuable insight into the coupling between widely separate influences. In so doing, it helps to define system boundaries at the highest level of abstraction.

Company dynamics can be viewed, qualitatively at least, using the so-called influence diagram of systems dynamics. Fig. 10 shows a company, at centre right, competing for a dwindling defence budget, at bottom left, and for engineers at top right. Systems dynamics influence diagrams are valuable in several senses:

(i) they encourage the highest level of abstraction
(ii) interactions (coupling) can be clearly seen between apparently remote elements in a system
(iv) system boundaries can be found.

This particular diagram suggests that reductions in education spending and increases in engineers leaving the industry can have just as serious an effect on defence business as reductions in the defence vote. Such analysis can (and should) influence companies to be as interested in retaining good staff as in bidding for new work.

Systems dynamics proceeds to a fuller modelling technique which is less simple to justify, owing to the need to quantify and describe in detail influences which may be hard to define mathematically. For example, it would be hard to define the precise relationship between competitors' contracts and the rate of consequent recruiting. Care must be exercised in the design of the model to avoid such parameters where practicable.

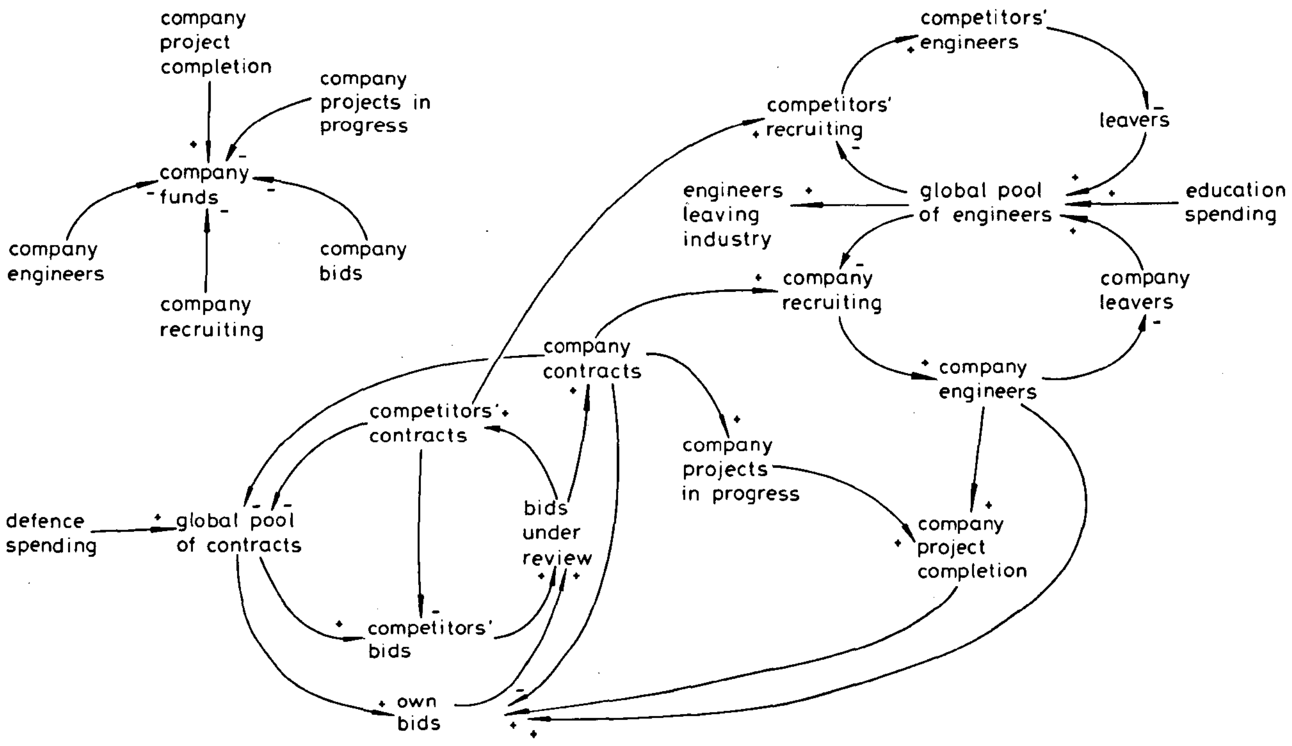


Fig. 10 *System dynamics and competition*
A system dynamics view shows competition to be for common pools of funding and of skills

11 Flow line for aircraft turnround

Flow line activity sequences may change materially at different throughputs. Modelling is a valuable tool in support of flow-line analysis.

Current interest in computer integrated manufacture (CIM) is revitalising flow-line analysis. Techniques and the experience in their use are few, however, and a simple analysis follows to show that the creation of flow lines is not obvious. The subject chosen is intensive flying of training aircraft which fly, in the example, at a fleet rate of 15 sorties per hour, and require to be replenished (turned round) between sorties.

The PERT network in Fig. 11A shows a single turnround. Rules are few and simple:

(i) The marshaller cannot check inside the cockpit until the pilot exits
(ii) Refuelling and re-oxygenation must not overlap, for risk of explosion.

The critical path (CP) for this single sortie is 12 min 41 s. But what happens when the sortie rate rises? Aircraft queue, principally, for the fuel bowser and the oxygen trolley. Fig. 11B is an activity timeline chart showing relationships, queueing times and activities. clearly refuelling is holding up the proceedings. And, moreover, the activity chart is not comprehensive; reoiling, for example, could occur much later than shown, and so float exists. Practical observation of the flow line showed an average turnround time of 19 min, rather than 12 min 41 s, and it was also noticeable that, as the sortie rate rose, a point came at which reoxygenation occurred before refuelling, not after.

Fig. 11C, the final PERT network, shows that, at 15 sorties per hour, the critical path has risen to 19 min, provided oxygen is replenished before fuel. The network therefore represented the fact and justified the

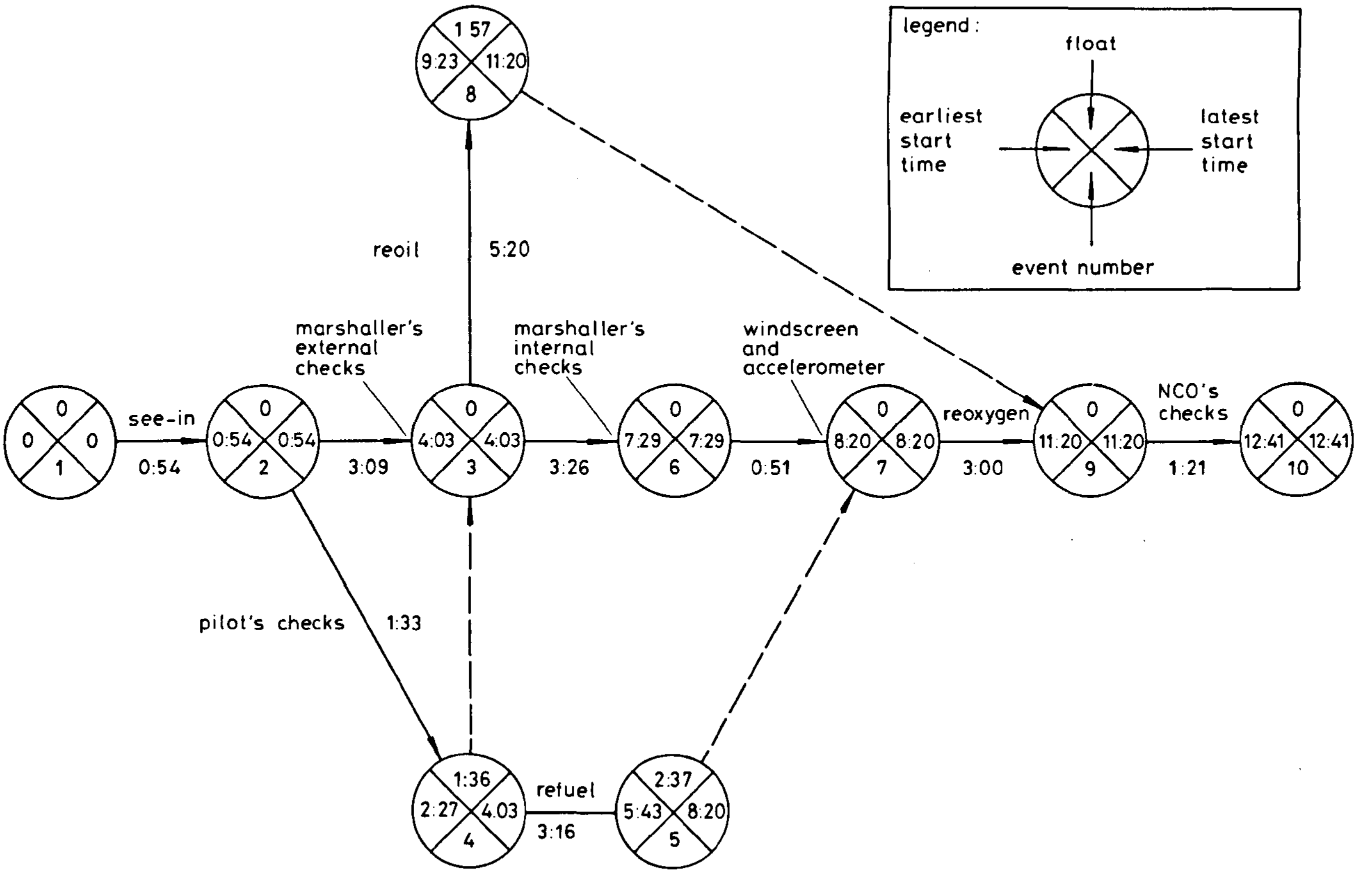


Fig. 11A *Aircraft turnround singleton PERT*

The Figure shows a network for one aircraft on turnround. Note refuelling before reoxygenation. Critical path for single sortie = 12 min 41 s. NCO = non-commissioned officer

Reoiling occurs 10 in 148 times, i.e. 0 : 06

Reoxygenation occurs 97 in 148 times, i.e. 0 : 65

Windscreen and accelerometer check occurs 41 in 148 times, i.e. 0 : 28

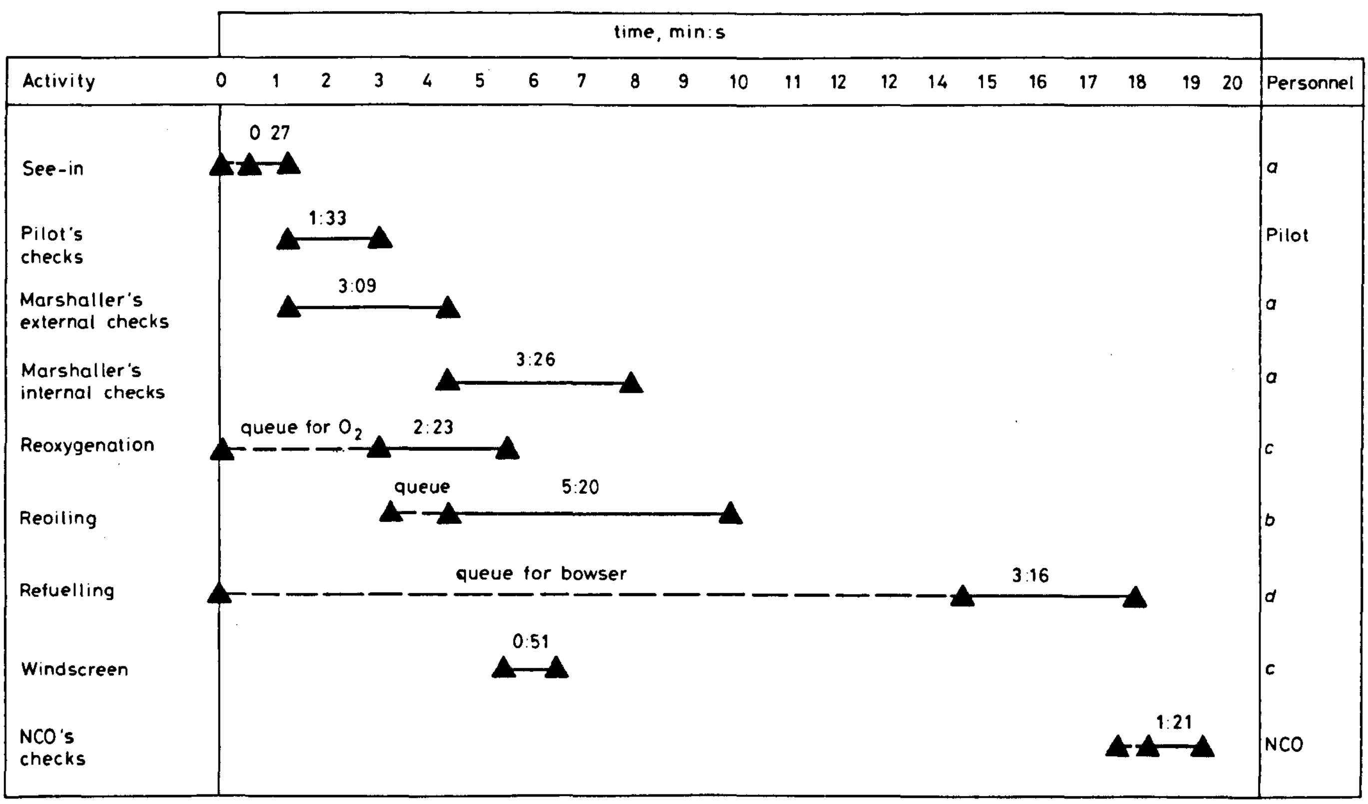


Fig. 11B *Activity timeline chart showing relationships, queuing times and activities*

NB Refuelling and reoxygenation must not coincide

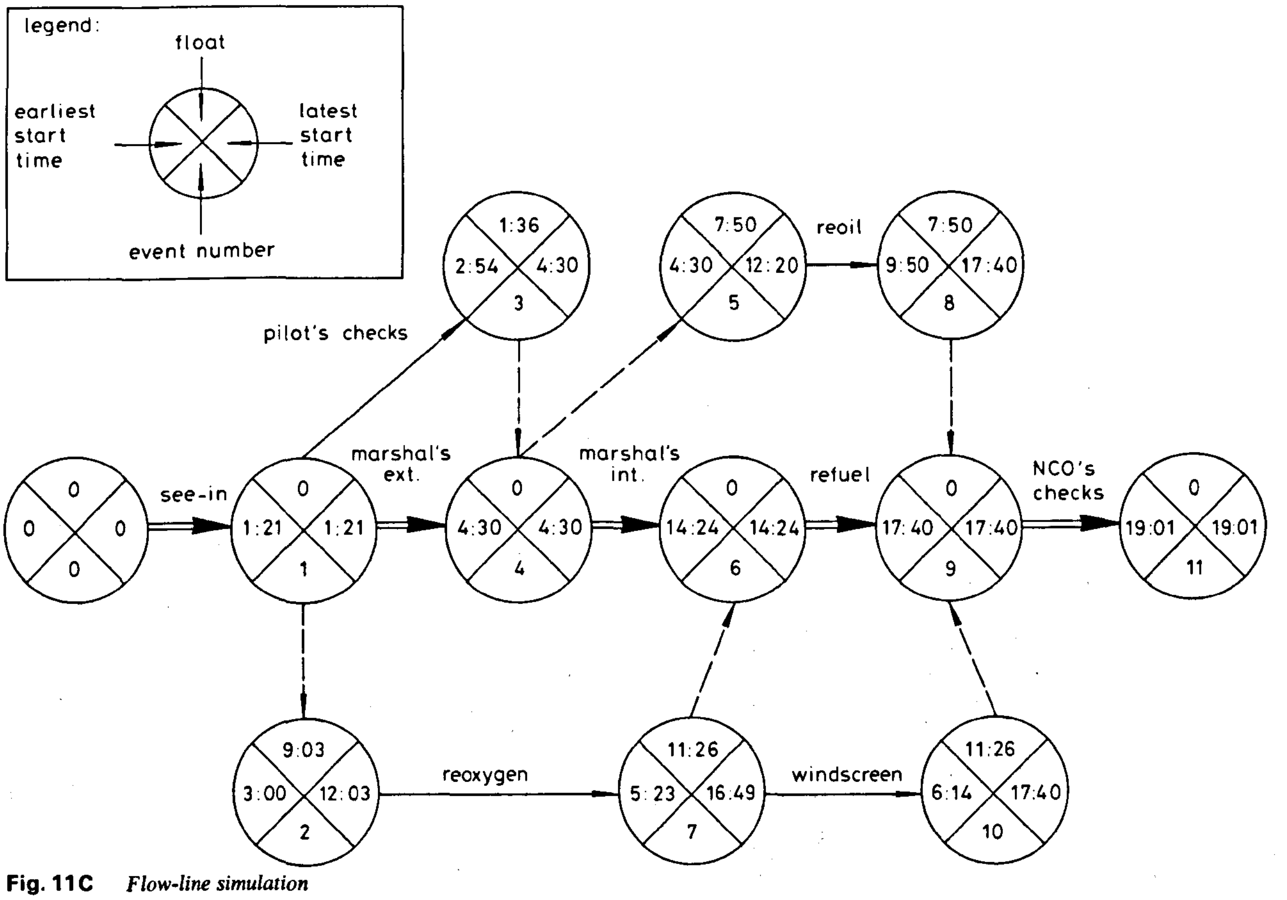


Fig. 11C *Flow-line simulation*

In flow lines, queuing effects introduce delays which vary with the flow rate and can change the preferred sequence of activities. In this real-life example, oxygen would be replenished after fuel when turning round a solitary aircraft. At 15 sorties per hour, the reverse sequence is shorter overall. CP = 19 min 01 s

(unauthorised) change in replenishment sequence which had been observed in practice.

We can conclude, in this case, and generally, that the one-off process does not multiply linearly to the many-off inflow, that the sequences employed to promote flow line throughput may change with rate, and that modelling of flow line process is feasible, practical and highly valuable.

12 Finite state/transition

The technique concentrates on critical factors causing a system to change from one state to another, and gives a valuable insight into design stability and thresholds.

So far we have considered continuous processes, evolution and growth. Life is not always like that.

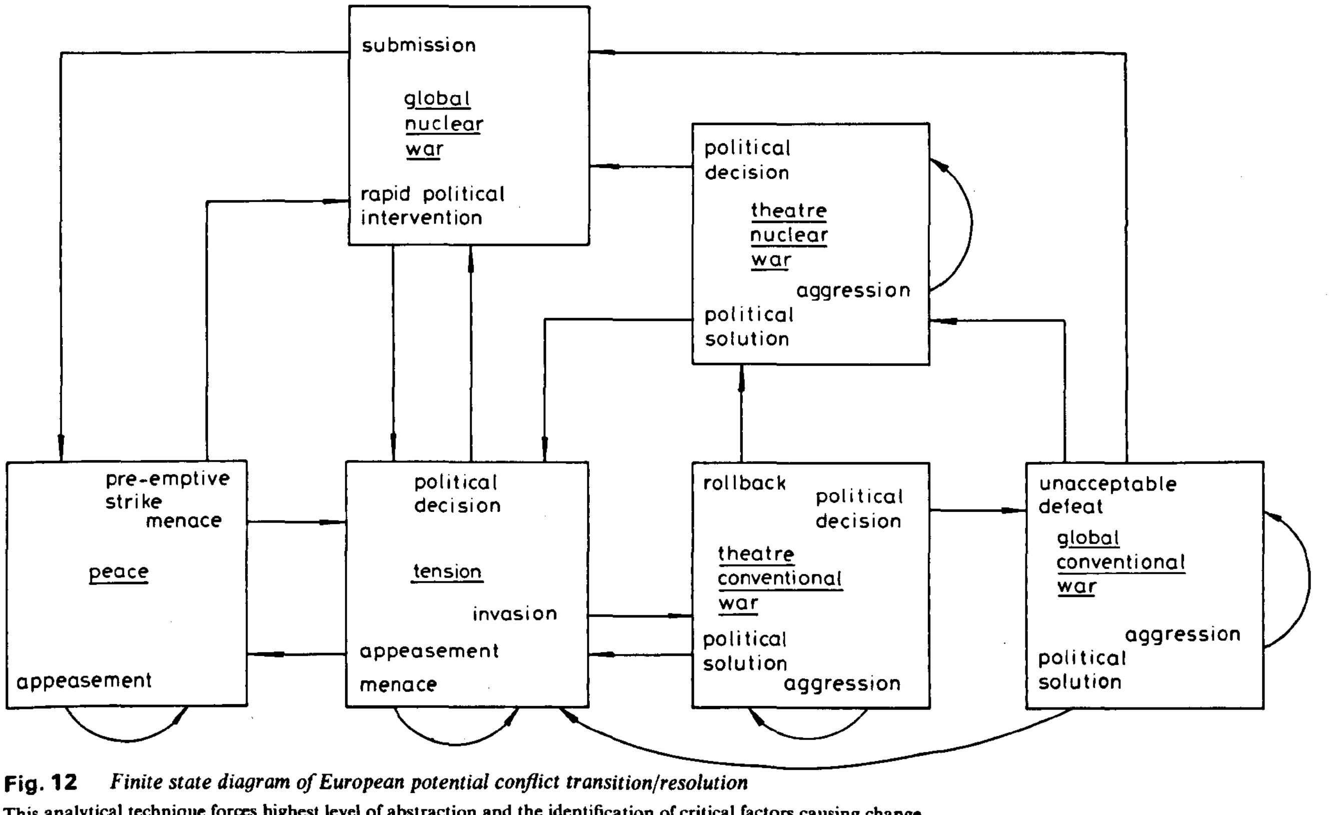


Fig. 12 *Finite state diagram of European potential conflict transition/resolution*

This analytical technique forces highest level of abstraction and the identification of critical factors causing change
NB Political solution generally implies victory or defeat

Catastrophe theory seeks to study the theory regenerative situations, and can prove quite complex. A simpler, but effective, approach is finite state/transition analysis.

Fig. 12 shows one such example for a major system, conflict in Europe. Six states are shown, and Europe is considered to be in only one at any time. It is the cause of transition that is of principal interest. Menace causes tension where peace reigned. Invasion causes theatre conventional war; and so on.

This valuable technique forces the highest level of abstraction and identification of critical factors; it is a valuable adjunct to any creative designer's toolkit, and, as shown here, can give clarity of insight into a generally intractable problem. In practice, most systems have a number of mutually exclusive states. A business, for example, can be static, expanding, or shrinking. A flow line can be smooth, irregular or stepped etc. It is useful, and can help in the creative process, to deliberately identify as many unique system states and their transition factors as possible.

13 Principles of creativity

The principles presented in Fig. 13, for engineering systems creation, are hopefully applicable across a wide sphere of activities.

The principles of creativity proposed in Fig. 13 are general and simple in concept; they require discipline to observe. It is the hope and intention that they have applicability beyond engineering systems creativity, as shown in the Figure.

- *Highest level of abstraction*
- *Breadth before depth*
- *Disciplined anarchy*
- *Functional before physical*
- *Decomposition before integration*
- *Level at a time*
- *Tight functional binding*
- *Loose functional coupling*
- *Functional migrates to physical*

Fig. 13 *Principles of creativity*

The principles apply to many endeavours including engineering, product design, man-machine interface, organisation and management, software engineering and even, perhaps, art.

Highest level of abstraction, breadth before depth, functional before physical, level at a time, binding and *coupling* have all been presented.

Disciplined anarchy implies that creative freedom must not only be permitted but positively required, within a constructive framework.

Decomposition before integration is necessary to understand the problem properly, and provide a sound, considered solution.

Functional migrates to physical is observed in design generally. Functionally, a computer input is quite different from its output. Physically, computer I/O is often one unit, driven by the physical similarities of the two functions. So it is with organisations; functionally discrete groups will aggregate for economy, communication, identity and many other reasons. But well defined functional responsibilities will ensure that performance is maintained throughout the migration.

Having presented the principles of creativity in an engineering environment, it is instructive to consider their wider applicability. At least two noted areas of creativity, oil painting and orchestral composition, have similar underlying principles, although they use different terminology. The artist in oils, for example, first sketches in charcoal and gradually works from a broad background towards a refined foreground, holding the whole eventual composition in his mind's eye. The composer seems to follow a similar line. So, too, does the system designer. Perhaps creativity does have common characteristics in all walks of life.

References

1 HITCHINS, D.K.: 'Managing system creation', *IEE Proc. A*, 1986, 133, (6), pp. 343–354
2 HOFSTADLER, D.R.: 'Gödel, Escher, Bach: An eternal golden braid' (Penguin Books, 1980)
3 SHINNERS, S.M.: 'Modern control systems theory and application' (Addison-Wesley, 1979)
4 ROBERTS, N.: 'Introduction to computer simulation' (Addison-Wesley, 1983)
5 RUIZ-PALA, E.: 'Waiting line models' (Van Nostrand Reinhold Co. Ltd., 1967)