Linguistic Research Institute

Report No. 14

STATE OF AFFAIRS SYSTEMS: Theory and Technique for Automatic Fact Analysis

Peter G. Ossorio

1971

Linguistic Research Institute

Report No. 14

STATE OF AFFAIRS SYSTEMS: Theory and Technique for Automatic Fact Analysis

Peter G. Ossorio

1971

© Copyright Linguistic Research Institute,1978

Whittier, California Boulder, Colorado

Foreword

The present report was written in 1971. The intent was to delineate an information processing system of an advanced type which would be capable of duplicating some of the essential characteristics of human behavior and some important achievements currently accomplished only by human beings.

Probably the system described herein (the "State of Affairs System") would be classified by most interested persons under the category of "Artificial Intelligence". Perhaps the most distinctive aspect of the State of Affairs System is that the system is based directly on a conceptualization of persons and their behavior rather than on the usual theories dealing with computability, data structure, language, or programming. Both its information-processing power and its general lack of intelligibility to members of the computing community stem from this feature.

Since 1971 some advances have been made in the conceptual basis of the system. For example, the conceptual and technical issues involved in the representation of real world phenomenon are dealt with at some length in "What Actually Happens" (Ossorio, ~~1974~~ 1971/1978). Also, the current formula for behavior has an 8th parameter, "Significance", in addition to the seven referred to in the present report. Much of the discussion of the hierarchical structure of processes would today be couched in terms of the significance parameter of behavior. On the whole, however, such later developments do not appear to warrant a revision, since the SAS system was from the outset presented as prototypal and merely illustrative of a new genre of information processing systems.

I do think some clarification is needed in regard to the verbal formula (V = C, L, B), and specifically to the B in that formula. On page 65 the comment is made that " the behavior (intentional action) accomplished by the performance of uttering it (the locution) is its use." and also that "without C the locution would be meaningless and ... without B it would be pointless." These comments may well give the impression that the major use of concepts is in verbal behavior. That would be incorrect. The B in the verbal behavior formula represents the class of behaviors in which the concept C appears in the value of the K parameter. In general, the most important members of that class will be non-verbal behaviors involving the use of that concept, but of course, that class will include verbal behaviors.

State of Affairs Systems was originally written for the Rome Air Development Center.

P.G.O.
1978

TABLE OF CONTENTS

1.0 INTRODUCTION

The objective of the present effort is to continue the design and development of an advanced information processing system designated as a State of Affairs System (SAS or SA System). Briefly, the SAS is designed to analyze and collate information content by transforming text into an updated fact file that can be used to make information systematically and selectively available for retrieval.

The concept of a SAS was presented in a general way as an organization of subsystems under Contract AF30(602)-4032. Two of these subsystems, the Classification Space (AF30(602)-3442) and the Attribute Space (AF30(602)-3432) have been delineated in detail and have undergone successful experimental evaluation separately and jointly (AF30(602)-4032). Additional subsystems, i.e. the Means-Ends, Process-Activity, and Part-Whole have undergone feasibility studies (AF30(602)-4032) in which implementable procedures were demonstrated. No experimental evaluation of these subsystems has been attempted as yet, primarily because it appears that the time is ripe to move to a new level of complexity in data processing configurations and to evaluate the subsystems first as elements in such configurations. The State of Affairs System delineated below is one of the simplest possible prototypes of such a complex configuration. The primary specific purpose of the present effort is to delineate a State of Affairs System in sufficient conceptual depth and functional detail to serve as a basis for experimental implementation in an operational setting. In this regard, specific consideration is given to FTD as the operational setting and intelligence analysis as the functional setting.

2.0 CONCEPTUAL ISSUES AND THE STATE OF THE ART

A survey of relational data systems, with particular reference to intelligence applications was recently completed by the Auerbach Corporation (RADC-TR-70-180: Relational Data System Study). Since it seems clear that any system which provides effective automatic processing of intelligence data will qualify substantially as a relational data system in the Auerbach sense, it is of some interest to note the summary conclusions of the survey.

2.1 Auerbach Summary Conclusions

A. "The systems surveyed evidenced impressive capabilities. However, we believe that none of the systems surveyed have the capability to handle all of the problems associated with the intelligence problem.

Nor do we believe that a comprehensive system that combines the best features of each approach will be fully satisfactory. This is due to gaps in our comprehension of natural language, of our knowledge of effective search procedures, and of the most effective methods of theorem-proving."

B. "Only limited practical experience exists with the use of even experimental RDS's.

Only a few of the systems reviewed, RDF, REL, QA-3.5, and RSS were operational to some extent. With the exception of RDF, no practical experience has been gained with respect to ease of use, problems encountered, effectiveness of the approach, and ease of applicability. Even with RDF, the experiments were limited. Hence, the utility of RDS technology is speculative rather than proven."

C. "A practical system oriented towards solving some of the intelligence data processing problems would borrow techniques from the current technology. An experimental system would be required to select and test those techniques which are most appropriate.

We believe that a system that is based primarily on deductive searches and is not concerned with natural language input or output could be developed. Before an operational system could be built, an experimental system would be used to determine the ease of use and effectiveness of such a system."

2.2 Technical Problems vs. Foundational Problems

The program of development of a State of Affairs intelligence data processing system is consistent with the Auerbach conclusions. However, some reservations are in order. The Auerbach summary may be misleading if its practical and optimistic tone suggests that the current problems of automatic intelligence data processing are merely technical ones which could reasonably be expected to yield to improved variants of existing procedures. Although eventually this might be judged to be the case, such a conclusion is by no means the only one or even the obviously most reasonable one to be drawn from a present servey of the state of the art.

The view taken here is that current approaches reflect some limitations in fundamental conceptualization which makes their limited success not at all surprising, though perhaps it should be emphasized that approaches which are in principle inadequate are not guaranteed to fail in practical applications, so that there is no implication here that these approaches are not worth pursuing.

What is characteristic of the analysis of intelligence data is that this enterprise represents a case of a person who as such is trying to reconstruct some aspect of the real world as such, mostly on the basis of what he takes it that some other persons have said. Moreover, the analyst acts on his reconstructions, and his relevant action consists of saying something which other persons will act on as a datum among other data.

On this basis, one would expect the computer implementation of intelligence data processing to be based on an explicit conceptualization of (a) the real world, (b) language, and (c) persons, <u>as such</u>, and in relation to one another. However, aside from the SAS approach, no such conceptualization has been brought to bear on the intelligence data problem. This is not surprising, since it is only recently that any such conceptualization has existed. Instead, the primary conceptual basis for existing relational data systems has consisted of some composite selection from (a) general systems theory, (b) computability theory, and (c) formal theories of logic, semantics, and natural language syntax. These various sources provide a variety of technical resources which are brought to bear on ad hoc formulations of some aspects of the intelligence analysis problem. Under these circumstances the limited success of these ventures could hardly be surprising, and serious reservations about their ultimate effectiveness would not be out of order.

2.3 An Example of the Technical vs. Conceptual Contrast

In general, the value of a conceptual formulation is to give coherence to the elements of the problem at hand and to distinguish between crucial or essential aspects and incidental detail or peripheral or derivative aspects. An example of this contrast in regard to the intelligence data problem may be constructed by comparing the brief statement (above) about the character of intelligence analysis with the Auerbach characterization of the problem.

"There are several features which distinguish intelligence data processing problems from conventional data processing problems. These are as follows: (1) The input data is fragmentary, in that the description, or even the identification, of an individual, object, or event is often incomplete; (2) temporal relations are important in two senses: in one sense the time factors are important and must be employed in assessing the information, such as the time of the event, the time of detection of the event, and the time the event was reported, and in another sense it is essential that the information be processed as soon as possible; and (3) the validity of the information is unknown: both the source of the information and the information itself must be assessed together with other information received."

At least four characteristics are mentioned here, without any indication that there is any relationship among them. But all four follow from the statement that intelligence analysis is a case of a person who as such is engaged in reconstructing some aspect of the real world as such mostly on the basis of what he takes it that some other persons have said. Further, all four may be seen as general characteristics of human behavior which are merely somewhat more salient in some aspects of intelligence analysis than they are in general:

(a) All observation, information, and reconstruction of any part of the real world is fragmentary and incomplete. Our concept of the real world is the concept of something the content of which is not exhausted by any set of observations or any finite set of facts.

(b) Temporal relations are fundamental aspects of the real world. No reconstruction of the real world which left out the historical relationships among objects and events would be even approximately correct.

(c) An individual person's behavior is part of the history of the real world and thus stands in temporal relations to other aspects of that world. In particular, a person must always act <u>now</u> on the basis of what he knows <u>now</u>, and decisions cannot be postponed indefinitely for the sake of more or better information. Thus, it is a general problem with respect to information not merely to get it at all, but to have it available at a time when it is needed.

(d) Sources of information, and their being sources of information to a given person, are simply parts and aspects of the real world. When a person, P, takes it that another person, Q, has given an accurate account of some matters of fact, F, his doing that is his having a representation of a world, W, in which not merely F, but Q and P, himself, have a place and have <u>certain</u> relationships to one another. If P takes it that the account of F given by Q is inaccurate or doubtful, the difference from the previous case lies in <u>which</u> relationships P reconstructs as holding among W, F, P, and Q. (For example Q was/wasn't in a position to observe F: Q has/hasn't the competence to distinguish F from non-F; Q's statement was/wasn't made with P's concerns and standards in mind, etc.)

Thus, the assessment of the source of information is simply a special case of the reconstruction of the real world on the basis of fragmentary and incomplete information.

In this example, relative to the piecemeal technical formulations, the conceptual formulation not merely contains additional information, but also introduces an order and coherence that is not present in the former. Correspondingly, effective computer implementation would seem to have more far-reaching consequences if it could be accomplished along conceptual lines rather than merely technical ones.

2.4 The Nature of the Conceptual Solution

Part of the point of the preceding example was to exhibit some relationships among (a) persons, (b) language, and (c) the real world, *as such*. The conceptual basis for the State of Affairs Information System delineated in later sections of this report consists of a formulation of each of these three and their interrelationships. Although each of the three is directly articulated as a calculus rather than, e.g., a description or taxonomy, ultimately the three are defined by their analytic relations to one another. Since traditional efforts at systematization have dealt centrally with no more than one of these three at a time it is not surprising that we have lacked fundamental conceptualizations of any of them.

The general notion of there being some intrinsic connections among language and persons or persons and the real world is not new. For example, in 1941 in "History as a System," Ortega y Gasset wrote, "Human life is a strange reality concerning which the first thing to be said is that it is the basic reality, in the sense that to it we must refer all others, since all others, effective or presumptive, must in one way or another appear within it." Such discursive formulations, however, have shown little scientific or technical viability. The present conceptualization represents an exceptional combination of scope and articulation which does provide a basis for technical development. Detailed and extensive presentations are given elsewhere (Ossorio, 1969; Ossorio, 1978). Those portions which provide the SA System rationale are presented in this report.

From the standpoint of computer implementation, perhaps the most significant general aspect of the conceptual formulation is that although it is explicit and systematic it is not a formalization in the usual sense and certainly is not simply or fully computable in the usual sense. This holds for each of the three major units (Language, persons-behavior, and reality) individually as well as the entire system. This was predictable from the nature of the subject matter. Nevertheless, roughly speaking it is (a) possibly as close to a formalizable domain, and (b) possibly as close to a formalization as it is possible to get without literally being one. Because of these features it is here designated as a "quasi-formalization."

Although each of the three major components is represented by a calculational system of a sort, each has some intractable characteristics from the point of view of computability. For example, the reality system is represented by a set of "rewrite rules" which do not progress toward a "terminal string" (or, contrariwise, in which every string is a possible terminal string). Or again, the behavior-persons system is one which begins tamely enough with finite elements and operations, but the recursive use of operations generates not merely an infinity of elements, but a corresponding infinity of possible new operations. Fortunately, as with grammar, the more exotic possibilities almost never occur in practice. Finally, the system of language is embedded in the system of persons-behavior and the infinite set of verbal behaviors stands in a specific one-to-one relation to the set of all behaviors, including verbal behaviors. Since any formalization is a subsystem embedded in the system of verbal behavior, the impossibility of formalization of the three major systems may be considered to be demonstrated, though perhaps only quasi-formally.

The status of the conceptualization as quasi-formalized rather than formalized presents methodological, and not merely technical, problems for computer implementation. These problems appear to dictate a particular methodological strategy. Ordinarily we think of the hardware of a computer as the medium for, or the embodiment of, a formal system which represents the functional characteristics and capabilities of the computer. In the present case both the hardware and a variety of formalizations and data structures must be thought of as the medium for, or the embodiment of, the quasi-formal system which represents the domain of interest.

In effect, we shall have to conceive of a computer directly on the model of a behaving person rather than either (a) thinking of the computer on the model of a formal system or (b) thinking of a behaving person on the model of a computer. All of the aspects of the computing system are then "composed" directly in terms of the behavioral model. These aspects include hardware, formal systems (Programming and inference machinery), data structures, and certain aspects of the relation of subsystems to one another (the SA System is designed as an organization of at least eight subsystems).

Of course, only an individual who was capable of operating the quasi-formal system could sensibly undertake such a task of composition. Such an individual would be a behaving person. Unlike the approach via formalization, success here is not a theoretical impossibility. It should be emphasized at once, however, that the aim of the SA System project is to make useful progress along this road, not to achieve definitive success. As noted above, the SA System delineated below is among the most primitive of this genre. For example: (a) no hardware design is brought to bear; (b) it does not incorporate Mitchell's human judgment simulation (RADC-TR-67-640), which would be an obvious resource, e.g., for assessing sources of information; and (c) it does not exploit the potential of multidimensional geometric representation (RADC-TR-67-640) for achieving the functional equivalent of intuitive apperception (Polanyi's "Tacit Knowledge") for the concrete organization of pattern recognition or executive operations.

The success of the strategy outlined above would not imply that the computing system could not then be rigorously described as a finite automaton. Neither would it imply that an element of indeterminacy had somehow crept into the operating characteristics of the machinery. The only implication would be that then there would be a significant behavioral description which was applicable and which was not deducible from the other descriptions.

Even this much is a familiar circumstance. When we say, for example, that a computer system has calculated a customer's bank balance or made a medical diagnosis we are giving a behavioral description which cannot be deduced from its formal operations and programming. Thus, the primary difference in the present case would simply be the level of complexity of the behavioral description applicable and the degree to which information to the system would have to be either hand fed (does it merely get its information or does it find it?) or spoon fed (can it tell <u>what</u> information it gets or does <u>that</u> information have to be provided).

2.5 Deduction and Representation in Intelligence Analysis

As noted above, the primary phenomenon of intelligence analysis is that of a person who is part of the real world, along with other persons, reconstructing certain aspects of the world, including his own relations to those aspects. The general pattern of his behavior is that in so doing he makes certain factual decisions which in turn may serve as part of the basis for still others, and so on.

(_a_) The phraseology of the foregoing is in deliberate contrast to the customary formulations in which the analyst would be represented as acquiring certain information which then serves as premises for a deductive (or probabilistic" or "plausible") inference which results in the acquisition of new information. The contrast between decision transitions and deductive operations illustrates the differences between the behavioral and the formalistic approaches. The point here is not that it is false that a given decision transition _could_ be represented _by a person_ as being a case of deductive inference from premises. Rather, it is a reminder that deductive inference is a very special sort of transition and that the greater part of most people's thinking and decision-making is prima facie _not_ of this sort. (There is an informative historical parallel between the two approaches: Formalism is the descendant of Aristotle's theory of theoretical reason; the present behavioral approach is the heir of Aristotle's theory of practical reason). Thus, and particularly in light of the development of quasi-formalization, one might well have serious reservations about the optimal degree of predominance of inference procedures in an intelligence data processing system. Any suggestion to the effect that the computer duplication of the analyst's synthesis of information is simply a matter of improving our technology of automatic theorem-proving might find a place in history alongside the suggestion that heavier than air flight was just a matter of improving our technology for flapping wings up and down.

(b) Reference to "the deductive power of system x" are ambiguous in a way which might be seriously misleading. The phrase could equally well refer to (1) the excellence of the deductive procedures employed or (2) the range of consequences that are drawn form given input as "implicit information." There is a doubly ambiguous intermediate case, namely (3) the degree to which deductive procedures are effectively exploited by being brought to bear on all of the relevant information contained in the rest of the system. The first of these three possibilities would seem to correspond directly to theorem-proving technology. The second would seem to correspond primarily to the representational power of the functional elements (including linguistic units)of the system, but would be subject to the constraints of case(3), namely that there be some deductive machinery which was effectively exploited.

An examination of the relational data systems surveyed by Auerbach shows that the primary function served by the "relational" aspects of these systems is to give representational power to the linguistic elements functionally incorporated into the system. In the main, this is accomplished by entering statements of relationships, R (p.q.) which serve as meaning postulates, in the Carnapian sense, for the classes, p and q, of items which are related. For example, given the relational formulas (1) that an airplane, p, may be in a given city, q, at a given time, (2) that an airplane, p, may depart from a given city, q, at a given time, and (3) that an airplane cannot have both relationships to the same city at the same time, then, given a list of members of the classes "airplane" and "city", it is possible to go from "AA-697 departs from Chicago at 10:53" to the deductive conclusion "AA-697 is not in Chicago at 10:53." It is relational statements which provide the basis for factual and not merely tautological conclusions.

For this very reason, however, the deductions which can be made by the system from "AA-697 departs from Chicago at 10:53" are restricted to those which follow from that fact together with the explicitly stated relationships in which "AA-697" and "Chicago" appear, together with any higher order linkages among relationships (e.g. incompatibility, as in the example above). The contrast to be drawn now is between the conclusions that can be drawn here by the information system having those relational statements and the conclusions that can be drawn by a person who knows about airplanes and cities. There is no comparison. The difference does not come about because the person can perform deductive inference better, for the information system is not unlikely to be superior in this respect. Rather, the difference reflects the difference in representational power which the functional elements "airplane" and "Chicago" have in the formal system (the computer) and in the quasi-formal system (the person).

Clearly, then, the total deductive power of a system can be increased indefinitely, though perhaps not without limit, by increasing the representational power of its functional elements. One way to do this is to increase the number of relational statements implemented in the system, including second-order statements relating relationships. However, even in limited portions of the real world the number of possible relational statements is frequently unlimited. Moreover, many such statements cannot be generated literally as meaning postulates. For example, instantiation or paradigm case statements are of this sort. The statement that a 20-foot object having (a) rubber-tired wheels, (b) a four-cycle engine, (c) a transmission, (d) a leaf and spring suspension, (e) shock absorbers, (f) recirculating-ball steering, and (g) a chassis enclosing (h) two upholstered seats is a paradigm case automobile by virtue of its having these characteristics is not a meaning postulate. The features (a) to (h) are not part of the meaning of "automobile." They are, rather, a familiar and significant exemplification; a different paradigm case might consist of a very rare, but strategic logical possibility of an automobile. Since we also do not have an adequate set of meaning postulates for the term "paradigm case" it would be largely impossible to exploit paradigm case methodology by proceeding on explicitly deductive lines. In contrast, that methodology can be incorporated directly in some number of ways in an information processing system.

Thus, although the SA System delineated below does include a set of deductive procedures, the emphasis is on maximizing the total deductive power of the system. Specifically, contributions to that total are distributed among the following in addition to inference machinery: (1) schematic representation of real world phenomena which operates as a second-order code for generating a variety of relational statements of the functional equivalents thereof; (2) procedural rules for moving from decision to decision in a non-inferential way; and (3) the functional interconnections among subsystems each of which deals with limited kinds of representational power and transitional capability.

2.6 Simulation vs. Information Processing for Intelligence Data

If the function of an intelligence data processing system is simply to give the analyst the information he wants, why does the issue of what it takes to duplicate his analytic procedures arise at all? Briefly, to give the analyst the information he wants *is* to duplicate his selective activity, since success here would consist of giving him those data which he (or some other selected criterion, such as a supervisor or a panel of intelligence experts) would have selected if he had examined manually the corpus from which the system made its selection.

The function of the system would be to relieve the analyst in just those tasks in which it could effectively duplicate his judgment, leaving him free to concentrate on the rest.

It is a common assumption that the replacement would be primarily in regard to simple, repetitive, clerical-like tasks. This is a reasonable conclusion, but it is easy to overlook the degree to which even "simple" tasks are judgmental in nature.

For example, the Auerbach report distinguishes as one major systems task "An on-line response to queries about specific facts in the relational data store, restricted by logical conditions which must be satisfied by reported items." An example of such a query is given as "Print the number of reports issued by the Academy of Science in the field of ionic propulsion during 1968 which were produced on projects supported by the MInistry of Defense."

The Classification Space studies reported in RADC-TDR-64-287 and RADC-TR-67-640 were based on the notion that it is a considerable oversimplication to suppose that a given document or portion thereof simply is or is not about a given topic or in a given subject matter field such as "ionic propulsion" in the example above. Those studies provide convincing evidence that documents vary on a dimension of degree of subject matter relevance to a given topic of field of knowledge and that the dividing line for a yes-no decision will in general be drawn in different places for different purposes and by different people.

The degree of relevance of a given term, sentence, paragraph, or document to a given subject matter can be expressed in a relational statement, but (a) such a statement cannot be generated as a meaning postulate, and (b) even if it could, it would be entirely unfeasible for any sizeable collection of documents to generate a relational statement for each document with respect to each and every topic which might be subject to a query. The classification space technology for subject-matter indexing and retrieval is an early example of a structural arrangement and a set of procedures which systematically generates the functional equivalents of an unlimited number of relational statements, and where the implementation is procedural rather than inferential.

The Classification Space, Attribute Space, and similar technologies provide a further example of the difference between a formalistic, deductive approach to intelligence data analysis and a quasi-formalistic, decision-making approach. In the former, the Classification Space provides an unlimited number of relational statements and the burden of incorporating them into inferential operations as premises. In the latter, the Classification Space represents the competence, or ability, to make a certain kind of judgment (subject matter relavance) over a limited range of such judgments. This behavioral description carries with it systematic implications as to the optimal placement of the Classification Space relative to other subsystems of the SA System. The design problem becomes one of analyzing a behavioral achievement (intelligence data analysis) as a function of the contribution of the several basic abilities (including deductive inference) contributed by the various components of the SA System. To have spoken earlier of "composing" an information processing system might well have evoked visions of a purely intuitive, inspirational approach. In fact, however, the task is one of straightforward pragmatic analysis of the sort which is familiar to anyone who has faced a complex problem of suiting means to ends. Inspiration is, of course, always welcome, and sometimes does occur.

3.0 CONCEPTUAL BACKGROUND OF THE SAS

What is characteristic of the analysis of intelligence data is that what is going on is that a person, who is part of the real world along with other persons, is reconstructing certain aspects of the world, including his own relation to those aspects. The function of an automatic intelligence data processing system is to duplicate the analyst's judgmental accomplishments insofar as that is technically feasible so as to relieve the analyst of some significant portion of the burden of his work.

Given these two characterizations, one would normally approach the problem by going to an existing theory of the real world and implement that theory in the form of a simulation or other dynamic representation. The technical problem with either the theory or the implementation would be to achieve (a) sufficient detail and precision to provide a satisfactory representation of the analyst's reconstructive possibilities and (b) sufficient cogency to achieve accurate calculation of the analyst's actual reconstructions. Such a theory of the world would also be a theory of behavior.

It is because the automatic processing of intelligence data is a task for which we would naturally turn to a world theory that it is important to emphasize that no such theory has existed. More conservatively put, no world theory has been evolved which has sufficient detail, precision, and cogency to give it any apparent promise for scientific and technical development.

3.1 World Theory: State of the Art

The exercise of conceptualizing the world and saying "what there is" is an ancient and popular one, and one which has had some outstanding exponents. In spite of the degree of controversy which various world views have generated, there has been a perhaps surprising area of near-universal agreement as to what the serious candidates are in regard to "what there is ." One might say that the controversies stem primarily from assertions of priority (e.g., ideas over substance, material bodies over ideas, etc.) rather than disagreement over the most general categories of what there is.

There are six of these general categories about which there appears to be appreciable consensus: (1) Universals (or ideas, or, in the present formulation, concepts); (2) Relations; (3) Objects; (4)Processes; (5) Events; and (6) Facts (or, in the present case, states of affairs). Moreover, there appears to be some consensus of a different sort, namely, that there is a difference between what there is and how it is (or what it is like), and that objects, processes, and events are the candidate for a bare catalogue of what there is, whereas concepts and relationships are indispensibles in a catalogue of how the world is. Facts, or states of affairs have an intermediate status. For Wittgenstein, for example, "The world divides into facts, not things", whereas for most other metaphysicians facts, being propositional in nature, represent how the world is and not merely its contents. Finally, there appears to be a very strong current consensus to the effect that although objects, processes, events, and states of affairs each have been picked by different theorists as being what there is, any one of these can be used to do the job and none can be shown to be more correct than the others or to have any priority over the others. Neither, however, have the different formulations been connected with one another. For example, Wisdom (1931) wrote "an account of the world in terms of things, an account of the world in terms of facts, and an account of the world in terms of events is just an an account of one world in three languages"(emphasis added). Since world theory appears to be a domain in which "you pays your money and you takes your choice," and since, moreover, none of the choices appears to have any scientific or technical utility, it is not surprising that the exercise of saying what there is and how it is has declined considerably in popularity. The primary legacy of the history of metaphysical thinking appears to be our current formal logical theories, which deal with concepts and relationships and purely formal individuals (i.e., individuals which are not per se objects or processes, etc.).

There are a number of pejorative aphorisms dealing with the failure to examine what one is doing. "He who will not study history is compelled to repeat it." and "He who denies that he has a philosophy has one nevertheless, but it is a defective one." The relevant aphorism here is "he who eschews metaphysical thinking is engaging in metaphysical thinking, but doing it poorly." the aphorism is apropos of modern science which publicly eschews metaphysical thinking.

Scientists, and most educated persons in our society, do have a world view, which may be summarized as follows: (W1) What there is in the world is objects as historical particulars. (W2) These objects are of the sort that physicists mention in their theories, i.e., subatomic particles. (W3) The world consists of such objects in configurational and dynamic relationships. (W4) The configurations are those which can be represented geometrically. (W5) The relationships are of the sorts mentioned by physicists in their theories. (W6) Human beings are middle-sized configurations of basic objects. (W7) Objects observable by humans are large or middle-sized configurations of basic obejcts. (W8) Relationships of other sorts are reducible to (are nothing other than) relationships of the basic sort. (W9) Basic objects, configurations, and relationships are what linguistic terms are _about_, or _refer to_, insofar as those linguistic terms have any real meaning and are not merely emotive, mythological, or otherwise merely subjective. (W10) The presence of human beings in the world is a historical accident. (Corollary W10A) The principles on which the world operates and the constituents (obejcts) on which these principles operate in no way depend on the nature of human beings or even on there being any such configurations. "It was there before we arrived on the scene, and it'll be there after we're gone!" (Corollary W10B) Human beings _as such_ are in the world as spectators; their actual participation in the world is as physical configurations and their apparent participation in the world as human beings is always illusory, but the illusion has a reality basis when it is reducible to physical participation. (W11) The presence of language in a world which contains human beings is a historical accident. (Corollary W11A) The principles on which human beings operate and the constituents on which these principles operate in no way depends on the nature of those configurations called "words" or "sentences" or "utterances". (Corollary W11B) Human knowledge of the world is acquired first and independently of language (i.e., by observation and thinking) and only then translated into or coded into verbal expression. (Corollary W11C) The relation of language to the world is entirely external, hence a connection must be made if linguistic expressions are to be applicable to the world. Semantic theory: "Here's the word and over here's the thing it stands for." (Corollary W11D) The relation of language to concepts and relationships is entirely external, hence a connection must be made if linguistic expressions are to have this kind of application: "Here's the word, and '---' is the attribute it designates." (W12) Although (W1) through (W11D) are the way the world is, I (and that goes for all of us) can't operate with it literally because (a) None of the things I observe are in fact reducible to basic objects and relationships-- all I have is a verbal formula which can be interpreted by a person as saying that these reductions can be made; (b) I can't separate out my language from my knowledge of the world; indeed, the very distinction between linguistic and non-linguistic is itself a linguistic distinction; (c) { A substantial number of additional distinct paradoxes could be generated from W1 through W11D. (W13) In spite of (W12), (W1) through (W11D)} must be accepted because that is the substantive import of modern natural science. Our inability to use it will be ultimately explicable in terms of (a) our human biases and deficiencies and (b) our scientific formulations are corrigible by reference to empirical research and so they may be expected to improve to the point where, partly by being better per se and partly by helping us understand our biases and deficiencies, they will be literally usable. What we have to do, therefore, is accept as a temporary expedient the separation of our developing knowledge of the world into a number of scientific areas each with its distinctive techniques, subject matter, and terminology, but connected by the fact that (a) the persons carrying out these studies are following a common set of methods and (b) in principle and eventually in fact, there is only one subject matter. That is the scientist's credo.

The foregoing is a kind of composite "physicalist view" which few would agree to without some reservations or rejection of some one or more points. It would be easy enough to demonstrate empirically, however, that the scientific community, particularly the natural science community within whose scope linguistic data processing has largely fallen, is characterized by the substantial acceptance of a majority of these principles and that that is enough to generate a widely shared attitude and approach to particular tasks which is never subjected to scrutiny or criticism because it is simply "they way the world is" and as such is the context, and not the object, of particular questioning and investigation. Correspondingly, any procedure, position or conclusion which is not in accord with these principles tends to be routinely dismissed as "philosophical", metaphorical, or simply obviously wrong. This point is of particular relevance here because, with the exception of W12 which notes that the physicalist view is unusable, the conception which underlies the SAS constitutes a rejection of the physicalist view on the whole and of every single principle W1-W13 separately, including the thesis that the demonstrable human successes of the natural sciences requires a physicalist view of the real world. The rejection does not take the form of opposing the physicalist's mystical and incoherent credo with a competing credo but of rejecting the credo approach altogether. Briefly, the SAS formulation is the formulation of a concept, specifically, an articulated conceptual system, and as such it is nonpropositional in nature. From this a variety of fundamental consequences follows: (a) It could not possibly be either believed or doubted, because (b) it could not possibly be either true or false; (c) it could not possibly be derived inferentially from any set of premises (since only propositions or statements can be the conclusions of inferences), hence also (d) it could not possibly rest on a foundation of any assumptions, and (e) it could not possibly be supported or not supported by empirical evidence, but (f) neither is it a priori, since only propositions or statements can be a priori or empirical.

Before going on to the SAS formulation it would be apropos to develop the criterion of adequacy to which it is responsive and to do this in the light of current approaches.

First, let us note that the physicalist view as presented above does entail the separation in principle of (a) the real world, (b) persons, and (c) language. In part this illustrates the fragmentation expressed in (W13); these are three _separate_ subject matters. In part, too, the separation reflects the empiricist credo normally associated with physicalism: it would be _improper_ to assert logical or conceptual connections, because these interrelationships must, properly, be established empirically. But consider the following dialogue:

Gil: The relation between language and the world is R.

Wil: Wait a minute. You can't do that. These questions have to be settled empirically.

Gil: How would you do that?

Wil: I would perform experiment E.

Gil: So if the question, Q, is "What is the relation between language and the world" then the results of experiment E will give us the answer to Q. Is that it?

Wil: Yes.

Gil: Now consider the question Q^1, namely, "Will the results of E give us the answer to Q?" That's another question like the first one. Now, it seems that either you've violated your own principle by giving me an a priori answer to _this_ question or else you're going to have to specify a new experiment, E^1 to decide this new question. _But_, if you take that tack you'll have to answer the question, Q^{11}, of whether the results of E^1 will give the answer to Q^1.

Wil: Come on now. You know what I mean. Don't give me all this philosophical nit-picking.

Gil: I wonder if there's a way of establishing that description empirically?

Wil: Oh, my!

The moral of the dialogue is that empiricism, whatever its actual virtues may be, is not a possible universal principle, since it could only be used that way if it were accepted as a non-empirical principle. [Notice, by the way, that the dialogue is not merely an academic exercise. All we have to do is substitute "information" for "experiment" and let Q be such a question as "Do they have re-entry capability?" and we have a problem in intelligence analysis: Will information E provide the answer to Q? Whether the answer is yes or no, how do I know that? Do I need still further information, and will I then be sure?]

The foregoing may be considered as an exercise in reflexivity, and the conclusion may be stated as "empiricism is not an empirical principle." A quick review of the phenomenon of a person reconstructing aspects of the real world, including it, himself, and his (verbal) informant will show that what gives coherence and unity to the whole is a variety of reflexive aspects: The world is known to part of itself, namely a person; the person knows about the world, including himself; language, which gives the person access to the world, is also part of the world, and so language gives the person access to itself.

The incoherence of the physicalist view may thus be diagnosed as stemming largely from the failure to deal effectively with any of these reflexive aspects. This failure is perhaps nowhere more clearly and explicitly demonstrated than in the paradoxes of self-reference that were generated in the course of development of formalizations of syntax and semantics, primarily for artificial languages. [And it is these which, together with state of the art systems theory, provide the conceptual-methodological foundations for almost all current work in linguistic data processing.] The response to the problem has been the object-language-metalanguage stratification and the infinite progress which it generates. This technically elegant device is justified on the grounds that in principle there is no language which cannot be semantically explicated in this way. What is normally not emphasized is that, correspondingly, no such explication of any language could take place without presupposing an unexplicated language having at least as much representational power as the language explicated. If a speaker were to use a semantic explication of a language L in understanding his own, unexplicated language M, it would have to be by using the unexplicated language, M, reflexively. But that is the very feature of language M which the 'explication' in effect denies the existence of. One might say, aphoristically, that if language were the phenomenon explicated by formal semantic theories of the familiar kind, natural language as we know it would not exist. It will not be surprising, therefore, that syntactic and semantic theories of the familiar sort have no contribution to make to the basic conceptualization of the SAS, though, of course, they can appear peripherally as historical special cases in the description of those actual or hypothetical persons in whose behavior these theories have a place.

The standard of adequacy to which the SAS conceptualization is responsive is, essentially, that the reflexive aspects noted above be preserved. More specifically: A world-characterization is formally adequate only if from it it follows that it is possible for the world to be known to be as characterized. Note that (W1)-(W5), the basic physicalist view, fails by this standard. If the world were that way, it could not possibly be known to be that way. The addition of (W6)-(W13) does not help, though it complicates matters. Briefly, even if other facts reduce to (W1)-(W5) facts, the possibility of that fact cannot be represented within (W1)-(W5) facts. Thus there are at least some facts which are additional to and independent of (W1)-(W5) facts, i.e., that there is an individual who knows that other facts reduce to (W1)-(W5) facts. Finally, if the independent existence of that individual is illusory, as stated by (W6)-(W13), then so is the reducibility of other facts to (W1)-(W5) facts.

The statement, above, of the standard of adequacy is too brief to be serviceable, and must be expanded: The world must be such that it necessarily could contain and does in fact contain an individual who knows that the world is as characterized and therefore knows that the world necessarily could, and does in fact, include himself.

Expanding still further: The individual in question must be characterized in such a way that it is necessarily possible for him to know the world to be as characterized. Moreover, the characterization must be not merely compatible with the more general characterization of the world, but specifically, a special case of it. More colloquially: (a) the individual must be the <u>kind</u> of individual who <u>could</u> know the world to be as characterized, and (b) the world as characterized must be such that the presence of such an individual demonstrably exemplifies the possibilities inherent in <u>that</u> kind of world.

Finally, in addition to the criterion of avoiding the paradoxes of non-reflexivity, there is a second criterion: The world-characterization does not pre-empt the answers to any questions that could be settled empirically. It should be noted that in the basic formulation this is an automatic consequence of a non-propositional formulation.

The representation of the real world and the problems connected therewith need not be as tortuous as the preceding discussion. It is one of the technical advantages of the conceptualization presented below that it provides systematic resources for dealing with just such questions and problems, including the question of the relation of language to persons and the real world. In the following sections the conceptual basis for the SAS is presented directly.

4.0 THE STATE OF AFFAIRS SYSTEM

Six basic category notions were identified above as being consensual candidates for explicating "the real world." These are (1) Object, (2) Process, (3) Event, (4) State of Affairs, (5) Concept, and (6) Relation. In the SAS conceptualization all six are brought together into a single system. The six are exhibited as (a) standing in analytic relationships to one another, (b) being defined by those relationships, and (c) collectively forming a system which is recursive and generative, or calculational.

The products of this system are "reality formulas" or structural reality concepts. When descriptive constants are substituted (in a formal, not procedural, sense) into these formulas, the results are conceptions of "the real world" or any part or aspect of it. Such conceptions will frequently be referred to as descriptions. The entire system is designated as the "State of Affairs System." Since this term is also used for the information processing system delineated in later sections, it will sometimes be necessary to refer to the State of Affairs Conceptual System (SACS) as contrasted with the State of Affairs Information S-stem (SAIS).

The SACS is later elucidated as having the status of "concept" within the more comprehensive framework which encompasses the world, persons, behavior, and language. This framework is designated as the "IA System," and it is the IA System which is the conceptual basis for the State of Affairs Information System.

Specifically, the systematic character of the State of Affairs Conceptual System is codified by a set of "transition rules" dealing with the analytic relationships among objects, processes, events, states of affairs, and relationships. The relation of all of these to concepts is implicit and is dealt with separately.

To anticipate some subsequent development in the interest of closure and orientation: A particular type of object (i.e., a person) is defined by its analytic relation to a particular type of process (i.e., human behavior). In turn, that type of process is defined by reference to seven fundamental parameters (comparable to the physicist's length, mass and time). Finally, one of those parameters is, essentially, simply the SACS itself, and the values of that parameter are the formulas which are the products of the SACS, with descriptive content. In this way the basic reflexive closures between persons and reality are obtained and the criterion of adequacy is met. In a later development, language is characterized formally as a special case of behavior in such a way that its distinctive character as language and not just some variety of behavior is exhibited. The distinctive character consists of an irreducible but complex 1-1 connection between locutions and concepts, including concepts of behavior and of language. In this way the basic reflexive closures connecting language with persons and reality and language are achieved.

In the presentation below, the articulation of the SACS will be given in detail, because it will carry the major burden with respect to the SAIS. Persons, language, and behavior, though equally important in the future and equally well articulated elsewhere (Ossorio, 1969) will be presented in more schematic fashion.

The formulations described so far are sufficiently detailed and analytically articulated to determine a genre or class of information processing systems. The analytic connections are the source of much of the distinctive representational power, hence also of much of the total deductive power of such systems. The conceptual definition of such a system is that it is a truncated model of the concept "person."

A system of this sort becomes an intelligence data processing system when additional constraints are introduced such that what is modeled is a particular kind of person, an intelligence analyst, dealing with a particular range of phenomena. Because of the reflexive aspect of reality discussed above, the simulation of the analyst <u>as</u> an analyst is formally equivalent to the representation of those aspects of the real world with which the analyst deals, including himself and his sources. It is here that we first become involved with the flow of documents, or more generally, the actual information pool, with which the analyst deals, as well as with the problem of the particular aspects of reality with which he aims to deal. As a simulation of the analyst, the system is formally responsive to the notion that the point of an automatic intelligence information processing system is to relieve the analyst of some of his tasks by duplicating his judgments in regard to these tasks. At that point, and from that development, there might be grounds for saying,"Only the technical problems remain. How well can we actually do it with the means at hand?"

4.1 The State of Affairs Conceptual System

The basic transition rules of the system are as follows:

1. A state of affairs is a totality of related objects and/or processes and/or events and/or states of affairs.
2. An event is a change from one state of affairs to another.
3. An object is a state of affairs having other, related objects as constituents. (An object divides into related smaller objects.)
4. A process is a sequential change from one state of affairs to another.
5. A process is a state of affairs having other, related, processes as constituents. (A process divides into related smaller processes.)
6. The occurrence of an event is a state of affairs having at least two states of affairs (i.e., "before" and "after") as constituents.
7. The initiating or terminating of a process is an event.
8. The occurrence of an object or a process is a state of affairs which is a constituent of some other state of affairs.
9. That a given state of affairs has a given relation to a second state

of affairs is a state of affairs. (The relation may be, e.g., succession, similarity, incompatibility, inclusion, etc.)

4.2 Some Limiting Cases

The basic transition rules can be applied successively, and the recursive structure of rules 1, 3, and 5 has the consequence that "object", "process", and "state of affairs" conceptions can each be composed and decomposed without limit by making use of the other concepts as intermediates. At a minimum, an infinite set of infinite (denumerable) sets of possible conceptions and descriptions is generated. The limiting cases presented below direct our attention to the difference between our concepts of "reality" and "the real world". The latter is the concept of a large finite set of historically particular objects, processes, and events. Both the physicalist view presented above and the layman's view of a world consisting of tables, buildings, trees, stars, etc. share this feature. In contrast, the notion of "reality" is a more general notion of "what is so" without any definite specification regarding historical particulars. Conceptions of "the real world" as a set of historical particulars are related to the recursive use of reality concepts in that they may be thought of as the consequence of introducing limiting cases, or boundary conditions on the compositional and decompositional generation of reality descriptions. Specifically, the following four limiting cases are most familiar and probably of greatest immediate importance.

1. The state of affairs which includes all other states of affairs (i.e., "the real world").
2. A type of object which is not a state of affairs (i.e., has no constituents, hence is an ultimate constituent, or "basic building block" object).
3. A type of process in which nothing changes (part of the history of an unchanging object or state of affairs; compare the concept of "absolute zero temperature").
4. A type of process which is not a state of affairs (i.e., has no constituent processes, hence is, in effect, a unit class of events, or practically speaking, is the same thing as an event).

Clearly, these limiting cases are ways of putting an end to the composition (case 1) and decomposition (cases 2,3,4) of objects, processes, events, and states of affairs. By setting a limit at each end, the result is a collection of historical particulars. Further, it appears that setting the lower limit can be accomplished by setting a lower limit either to object or process decomposition as such. In the physicalist view, and in physics the limit is set on objects. Thus, we identify certain kinds of objects as ultimate particles. Certain kinds of corresponding possible processes then dissappear in favor of states of affairs or limiting case processes. For example, the motion of certain particles is a non-fact and we deal instead with a certain state of affairs concerning the likelihoods that a given particle is at given locations at a given time. On the other hand, other processes, e.g., the motions of other particles, are divisible without end - they are continuous processes.

Clearly, too, the setting of lower bounds need not be accomplished at a single stroke for all possible states of affairs. It would be possible to set certain types of object as basic with respect to a given range of facts, or states of affairs, and other types for other ranges. Thus, the division of reality into a variety of different types of objects and facts can be accomplished without the physicalist's insistence that, of course, there must be a single sort of ultimate object.

The state of affairs system represents a more complex handling of these logical considerations. One might say, it is anchored in the middle, at the level of observation and human behavior, rather than by the setting of a lower bound which generates historical particulars. In this system there is an ultimate process, namely human behavior as intentional action. Ultimate objects are the correlative of the ultimate process, i.e., persons, which are by definition the kind of object which participates in intentional action processes. The products of compositional and decompositional maneuvers are always related to ultimate processes and objects as products. Every description and reconstruction of "the real world" is someone's description or reconstruction. So also is the setting of upper and lower limits. In this form of conceptualization, the fact of limit-setting can be represented explicitly either as a general case or in particular cases of particular persons or groups of persons. Moreover, its consequences, i.e., the conception of _a_ "real world" composed of certain historical particulars, can be represented _as_ consequences. In contrast, the physicalist view and other theoretical views cannot represent the limit-setting because such views are the consequences of a particular limit-setting, which they presuppose and therefore can neither represent or criticize.

The present formulation makes a number of significant contributions to world-theorizing, including the clarification of three points which have been mentioned as phenomena.

(a) It explains the consensus with respect to the six reality categories: each one is a fundamental term in the _system_ of reality.

(b) It explains the equivalence of "object","process", "event", and "state of affairs" for simply identifying "what there is". Not only are descriptions which mention any one of these replaceable by descriptions which may mention any of the others, but also, by virtue of rules 1,3,5,and 7, our first limiting case, the concept of "the world" can equally well be thought of as (1) an all-encompassing state of affairs, (2) an all-encompassing object, (3) an all-encompassing collection of events, and (4) an all-encompassing process.

(c) It explains why the reflexive closures involving persons and the real world are essential. The state of affairs conceptual system is not itself either an object or a process or an event or a state of affairs. Thus, it is not a candidate for the status of being simply "what there is". As a concept its only status in the real world is given by the part it plays in the behavior of individuals who are capable of using it. On the other hand, the use of it by these individuals must be a state of affairs, and they real individuals, not merely the concept of a state of affairs and they imaginary individuals. Hence the SACS is as essential to there being behaving individuals as there being behaving individuals is essential to there being a SACS.

4.3 Comments on the SA Conceptual System

In the interest of closure and orientation, certain comments regarding the state of affairs conceptual system appear to be apropos in connection with its later use.

(1) Methodologically, one good reason for anchoring the conceptual system at the "level" of observation is that that is true to life. People act in the light of their circumstances, and what they know of their circumstances _is_, "ultimately", acquired by observation. There is a fundamentally non-accidental relationship here. The six reality categories are not constitutive of reality and merely accidentally or incidentally connected to what we observe. Objects, processes, events, and states of affairs are what we observe _per se_ and all six categories represent the terms in which observations _per se_ are

made. To say that we observe such things is to say that there are exemplars of _each_ sort which we come to know about on occasion without on that occasion having to find out something else first. ("observation" contrasts with "inference.")

For example, I observe an object when I see a rifle, hear an airplane, taste an apple, touch a tank, or see an airplane, or an apple, or a tank. I observe a process when I feel the wire warm up, see the salt tablet dissolve, or hear a jet airplane move from left to right overhead. I observe an event when I hear the airplane explode, see the light blink on the computer console, or feel the gravel hit my forehead. I observe a state of affairs when I see that the tank has a double turret, when I hear that the motor is out of tune, or notice that he is angry or that he is lying or not paying attention. Note that in these several cases I observe instances of concepts or relationships although I do not observe concepts or relationships; I observe _that_ something has a given characteristic, or _that_ it has a given relationship to something else, and that is to observe a state of affairs.

The connection referred to here between knowledge and observation is an analytic one, not a piece of epistemological theorizing. It is not implied, for example, that observation is the starting point for knowledge because it is indubitable (there is no "protocol language" or "observation language" here) or because observation reports consist of "simply reading off the features of what is actually there."

Rather, the logical constraint is that knowledge must have a starting point somewhere, since we could not and to not go through an infinite regress of cognitive operations. But it need not start in the same place for different persons (competence enters in here) or for the same person at different times or with respect to different sorts of facts. "Observation" marks that boundary wherever it may lie for a given person at a given time, and the fact of there being observation is no more than the fact of there being such a boundary condition.

Which is to say that, as with all the other concepts with which we shall have a major concern, "observation" is defined by its place in the system (IA System) of possible behavioral facts and not by reference to any external (to this system) circumstance or intrinsic feature by virtue of which it has that place.

(2) The latter statement holds for the six reality concepts also. Because they are defined by their analytic relations to one another, they are formal concepts like "three" or "addition" rather than names or descriptive concepts like "Chicago" or "red". A consequence is that when we speak of using these concepts we are not talking of their being applied to anything. Rather, when limit-setting has been accomplished, the historical particulars which result then serve as the "referents" of our descriptive and referring expressions.

Corresponding to their status as formal, rather than descriptive concepts, the SA System concepts are content-free. For example, an object is not, _per se_, any particular type of object such as a psychological object, an economic object, an artistic object, a biological object, a material object, a physical object, or a mathematical object. To identify and refer to any particular kind of object, some logical constraint must be added to the general concept of "object". This point bears explicit mention because one symptom of the pervasiveness of the physicalist view is the uncritical assumption that an object is, per se, '_really_' a physical object, so that whereas with other terminology we identify meanings, it is only when we use the terminology of physics that we directly identify referents, because physical objects, being "what there _is_", are also what any terminology must be _about_. Associated with this is the equally uncritical assumption that material objects such as tables, trees, and mountains are, _per se_ , physical objects whereas they are obviously social, behavioral objects.

Rule 1 (basic transition rules) states that a state of affairs is a totality of _related_ objects and/or processes and/or events and/or states of affairs. The nature of these relationships may be, e.g., geometric, economic, emotional, kinetic, or any others which are appropriate to the type of object involved. Here again, it is an analytic relation that is involved. Types of object are distinguished by reference to the types of relationship in which they might be found. Conversely, whenever any two or more objects are related in any of the ways in which they can be related, <u>that</u> they are so related is a state of affairs.

(3) The presentation of the transition rules is not really modeled on a formalization. This is evidenced in such procedures as making use of such terms as "constituent", "initiating", "sequential", "change", and some others without any explicit rules for their use. To do so would be anti-heuristic for most purposes since any such explication would inevitably take us back to object, process, event, and state of affairs concepts. It should be kept in mind that in the context of the theory of practical reason, "heuristic" loses its pejorative connotation of "but not really rigorous".

This feature is implicit in the notion of quasi-formalization and the peculiarity of the behavioral system mentioned above, namely that we generate an infinity of operations as well as an infinity of elements (which are behavior formulas, analogous to the reality formulas generated by the SACS . If to use or "operate" a formal system is to perform permissible operations or permissible elements, then to use or "operate" the system of behavior is nothing other than to behave. If the point of writing down a formalization is to codify the constraints which define the correct use, in the sense of "operation" not "application" of the calculus, then the corresponding virtue of a codification of this quasi-formal system is nothing other than its being "heuristic". Terms such as "constituent", "initiate", "change", etc. appear to be exceptionally non-problematical except in the context of formalization.

5.0 PERSONS, BEHAVIOR, AND LANGUAGE

5.1 Persons and Behavior

Given the concepts of "object" and "process" as defined by the SA System transition rules, we may now think of defining some special kinds of object and process by adding some logical constraints over and above the transition rules. Thus, a type H object, or H-object (human object, person), is defined by reference to the concept of an IA process (intentional action, behavior). The IA process is defined by seven fundamental parameters, two of which provide the links to the concept of a person and one of which provides the primary link to the SA conceptual System. The general formulation of the concept of behavior is given by formula (1).

(1) B = <I, W, K, KH, P, A, ID>

The interpretation of these symbols is as follows:

B	Behavior:	Intentional action as the general concept of behavior.
I	Identity:	The identity of the individual whose behavior it is.
W	Want:	The motivational parameter.
K	Know:	The cognitive parameter.
KH	Know How:	The competence parameter.
P	Performance:	The procedural (process) parameter.

A Achievement: The result, or outcome, parameter.

ID Individual Difference: The personal characteristic parameter.

Formula (1) represents the fact that one behavior is distinguished from another behavior as such by virtue of having different values of one or more of these parameters and that a behavior is incompletely specified unless its values on all of these parameters are specified, but a type of behavior may be specified by specifying some values on one or more parameters. In actual practice behaviors are seldom identified by giving full parametric specifications, and no effort to do so carries any guarantee of success (recall the incompleteness of reconstructions of the real world). Most often they are indexed by descriptive locutions which apply to B and imply more or less precise restrictions on the parametric values of B. (Consider the implications of "He wrote down Peano's third axiom set" or "He telephoned the embassy for the values of the individual parameters.) Such incomplete specification of behavior is often exactly what is needed for the purpose at hand.

Values of the identity parameter are given by individuating expressions such as names or space-time locations. Values of the Know, Want, and Achievement parameters are given by state of affairs descriptions. (The values literally are states of affairs for Achievement; they are states of affairs concepts for Know and Want). Values of the Know How parameter are given by state of affairs descriptions which ascribe particular kinds of learning histories to the behaver. Values of the Performance parameter are given by historical process descriptions (thus, it is the P parameter directly and the A parameter indirectly, which give the behavior its status as a historical particular).

Values of the ID parameter are given by a new set of parameters designated as "personal characteristics" or "individual differences". These are functions which map the behavior, as given by the other parameters, into a class of patterns of occurrence of like behaviors in the life history of the individual. These functions are (a) Trait, (b) Attitude, (c) Interest, (d) Style, (e) Ability, (f) Knowledge, (g) Values, (h) State, and (i) Status. Since this range of concepts does include what are commonly referred to as "personality characteristics" it is in no way paradoxical to present these functions as parameters of persons as such, i.e., as the ways in which one person can resemble or differ from another person as such.

With respect to the ID functions we shall be primarily concerned with (a) the fact that there is such a set, (b) the particular ID category of "Ability", and (c) to a lesser extent, the categories of "Knowledge" and "State".

Thus, in general, the parameters of behavior provide the analytic links between behavior and both persons and the world, and they do so in a way which goes far beyond simply the fact that objects and processes are instances of SA System Concepts. The Identity and Individual Difference parameters provide the connection to persons and the Know, Want, Know How, and Achievement provide the connections between IA process and SA System Concepts. Among these four, the connection via the Know parameter is primary, since for the behaving individual (a) whatever he wants is ipso facto something he discriminates, hence whatever specification is given of W must be included in the specification of K (N.B. the relation of formulas (2) and (3) below), and (b) accomplishments and learning histories (or in general, any state of affairs) are accessible to the individual (or any other individual) only via what he knows. [Incidentally, the use of "Know" for this parameter is a descriptive choice: it does not imply correctness here as it does in ordinary usage, but neither does it carry the blanket reservation about correctness that would be implied by "belief". As used here, the term is completely non-commital with respect to correctness.]

5.2 Aspects of persons, Behavior, and The World

In this section several topics are dealt with separately in order to round out the presentation, thus far, of the conceptual basis for the State of Affairs Information System.

(A) Individual Difference parameters provide a follow-through on the notion of limit-setting which took us earlier from the basic notion of "reality" to the more commonplace notion of "the real world". The transition rules of the SA System do not determine either any limit-setting or the employment of any particular descriptive terms. Both must be intorduced "from the outside" and in that sense are "arbitrary" relative to the transition rules. The logical role of the person is that of the individual who provides these additional, "arbitrary", constraints. The statements that "every reconstruction of reality is someone's reconstruction" and "every description is someone's description" are analytic statements as well as useful reminders. It is for this reason that the computer representation of the domain of facts with which intelligence analysis is concerned is conceived of formally as a simulation of an actual or hypothetical intelligence analyst.

Individual difference concepts come into play further in a more narrowly technical sense in other respects. One example is that in dealing with conflicting reports they can be recorded directly as conflicting reports or "what P said" independently of any decision as to which, if any,is correct.

Another example lies in the representation of beliefs or theories. In monitoring the state of the art in, say, ionic propulsion theory for a particular historical group of persons, the development of the theory over time can be represented as a set of beliefs held by a hypothetical individual who is paradigmatic for that group. Where inconsistencies or controversies are found these would be codified as viewpoints, each of which is represented by a hypothetical individual.

(B) The ID notion of "Ability" is that of a class of achievements which the individual in question could be expected to succeed in accomplishing if he tried. If no contingencies are introduced explicitly, "under normal conditions" is understood. There is an associated derivative concept of "merely able" or "capable of". If P is capable of achieving X, then there is at least one set of conditions in which he could be expected to achieve X if he tried.

(C) The ID concept of "Knowledge" is that of the repertoire of factual judgments available to the individual. An equivalent definition is that it is the set of factual discriminations which the individual has the ability to act one. The example mentioned above of monitoring the state of the art in ionic propulsion theory would come under the heading of assessing the ID characteristic of "Knowledge". In contrast, the task of monitoring the state of the art in ionic propulsion technology would be codified as part of the social practices in which theory and technology both had a place. ("Social practice" is developed formally below, and the representaion of processes, of which social practices are instance, is dealt with in detail in the technical section.)

(d) The ID concept of "State" is defined as follows: "When an individual is in a particular state there is a systematic difference in his powers and/or dispositions." Here "dispositions" includes Traits, Attitudes, Interests, and Styles; "powers" includes Abilities, Values, and Knowledge. This definition conforms closely to the ordinary notion of a state as a temporary condition which affects an individual's personal characteristics. The concept of "state" enters in primarily because much of intelligence analysis has to do with monitoring changes which occur, and in this connection it is frequently of interest to decide whether a given change has a long-term significance or only a temporary one.

(D) Both paradigm case methodology and parametric analysis are methodological devices which are available when definitions are not. Parametric analysis is a way of coping with the fact that ultimates cannot be defined. When we reach an ultimate category we cannot answer the question "what is it really ?" (What is space, really?) What we can do is articulate the category by formulating the ways in which exemplars of that category may resemble or differ from one another. Frequently a parametric analysis gives us technical leverage in dealing with the phenomen in question, so that we frequently engage in such analysis even when definitions can be given.

Paradigm case procedures are for dealing with a range of cases for which no single definition is available because the cases are heterogeneous (e.g., the concept of "family" or "weapon" or "game").
The procedure is twofold. (a) First a definition is given or a type of

case is identified, and this is the "paradigm case". (b) Subsequently remaining cases are accounted for by characterizing them as being the same as the paradigm case with a specific transformation or alteration having been carried out; the number of transformations which have to be introduced for the coverage of all the cases will depend on the heterogeneity of the cases and the choice of paradigm. Many of our ordinary descriptive efforts incorporate this kind of move ("It's a T-33 with an afterburner"; "It's a stripped down version of a ...").

(E) Objects, processes, events, and relationships (hence by implication, properties also) are parameters of states of affairs. This is, one state of affairs may resemble or differ from another in terms of which objects, processes, and events in which relationships are the constituents. Thus, concepts of objects, processes, events, and attributes (this includes properties and relationships) are necessary in order that there be concepts of states of affairs to serve as values of the "Know" parameter of behavior. [To be sure, transition rule 1 shows that a state of affairs may have a second state of affairs as a constituent, and it follows that the latter might in turn have a third state of affairs as a constituent, etc., but constituent states of affairs would ultimately have to be specified in terms of objects, processes, and events and their relationships otherwise the explicit specification of the original state of affairs could not be given.]

For this reason, an incomplete specification of the value of the Know parameter of a given behavior can be given by specifying that a certain concept is being used or that a certain discrimination or distinction is being acted on. The latter two are equivalent. For example, to use the concept of "dog" in one's behavior is to act on the discrimination or distinction between dogs and non-dogs. To say that the concept of "dog" was used in an individual's state of affairs concept which he acted on; to act on a particular state of affairs concept, e.g., that the dog is lying on the grass, is to act on the distinction between that state of affairs and any other state of affairs.

We may note in passing that although the Know parameter of behavior has a strong resemblance to the concept of "stimulus" in traditional behavior theory, it is categorically different in that the state of affairs concept on which one acts need not, and generally is not, restricted to a representation of what is then and there present, although unless the person is confused and disoriented, it will <u>include</u> a representation of himself and his circumstances. For example, if a person asks a friend whether Eve really did eat the apple in the garden of Eden, or if he tells someone that the boy's brother runs the store he is making use of concepts of things that are not then and there present but which stand in some relation to what is then and there present. The state of affairs concept on which a person acts is not merely a symptom of what is then and there present either. This kind of notion is simply the physicalist view as exemplified in traditional psychological theory, and it shows the same incoherence: Only an individual whose behavior was *not* merely a symptom of events and circumstances would be in a position to discover, or even to suppose, that *other* individuals' behavior was merely a symptom of what was then and there present in their circumstances.

5.3 The Concept of Behavior as a Calculational System

The following have been specified previously: (a) That the values of the W, K, and A parameters of behavior are given by state of affairs descriptions; (b) That behavior is a process; and (c) That the occurrence of a process is a state of affairs. Since no restrictions have been placed on these three statements, it follows that the concept of behavior is fully recursive in these parameters. Thus:

(2) B = <I, W, <B> , KH, P, A, ID>

(3) B = <I, <B> , <B> , KH, P, A, ID>

(4) B = <I, W, K, KH, P, <B> , ID>

However, a more general set of results can be accomplished by stating directly that the concept of behavior is a calculational system. Here "calculational system" is to be understood in terms of the Element-Operation-Product model. That is, (a) an initial finite set of elements and operations is specified explicitly; (b) It is tautological that operations are performed on elements and that the result of doing so is a product; and (c) anything that is generated as a product is also an element.

In the light of the discussion above of using concepts and acting on distinctions, we may say that a person may (and normally does) use the concept of behavior (as given by formula (1)) _in_ his behavior. Further, he uses it as a calculational system in which the concept given by formula (1) generates an unlimited number and variety of behavioral concepts as products and it is that range of products which is available for use. When we represent the calculational system, as in Table 1, and derive some first-order products, including one which is designated an"achievement description", we will be able to "tighten" the foregoing by saying that it applies _as_ an achievement description. Likewise, we will be able to say that the four operations in the system are simply four exemplars of formula (1) given by achievement descriptions, so that nothing other than the concept of behavior and some of its particular instances is required for the calculational system as such.

Notational conventions with respect to Table 1 are as follows: (_a_) The replacement of W, K, KH, P, or A by θ indicates that no commitment with respect to that parameter is made (i.e., that parameter is 'deleted') by giving that form of description. For example, if "He is studying ionic propulsion" is given as an Activity description, it would be more clearly put as "He is studying ionic propulsion, but I don't know (or I'm not saying) why". However, if it is given as a performative description, that would be more clearly put as "He is trying to study ionic propulsion, but I don't know (or I'm not saying) whether he is succeeding". (_b_) The replacement of one of the parameter symbols by a C or an E and the omission of commas between two C's or E's indicates that the distinction between the parameters as such has been dropped. For example, I, CC,EE, A, ID indicates that no distinction is made between the W and K parameters or between the KH and P parameters, so that the seven-parameter formulation of B has been transformed into a five-parameter formulation in which two of the parameters stand in a C-E relation.

Table 1. IA as Calculation

Element	Operation	Product
< I,W,K,KH,P,A,ID >	Substitution	< I,W, < B,KH,P,A,ID > Cognizant Action description
	Substitution	< I, < B > , < B > ,KH,P,A,ID > Deliberate Action description
	Substitution	< I,W,K,KH,P,< B > ,ID > Social Practice description
	Deletion	< I,θ,K,KH,P,A,ID> Activity description
	Deletion	<I,θ,θ,θ,P,A,ID > Performance description
	Deletion	<I, θ,θ,θ,θ,A,ID > Achievement description
	Deletion	<I,W,K,KH,P,θ,ID > Performative description
	Identity	<I,W,K,KH,P,A,ID > Intentional Action description
	Reduction	<I,CC,EE,A,ID > Cause-effect description
	Reduction	<I,CCC,E,A,ID > Cause-effect description

(c) Substitution is without reference to any particular values of the parameters of behavior. Thus, for example, when a Cognizant Action description is given in expanded form as:

< I,W, < I,W,K,KH,P,A,ID > ,KH,P,A,ID >

there is no implication that both W's have the same value or both I's, etc.

(d) Substitution implies inclusion only. Thus, for example, a Cognizant Action description indicates that < B > is included in the value of the K parameter and not that K is completely specified by < B > .

The major sorts of utility of the various forms of description are as follows:

(a) Cognizant Action descriptions may be used to represent behavior (1) which consists of describing behavior or (2) of a person who knows what he is doing.

(b) Deliberate Action descriptions are used to represent behavior which the person not merely distinguishes, but has grounds for engaging in one behavior rather than the other.

(c) Social Practice descriptions are used to represent sequences of behaviors by a single individual or patterns of behavior involving more than one participant.

(d) Deletion operations make it possible to refer to, or represent, or act in terms of, any aspect or combination of aspects of behavior short of the full representation given by IA descriptions.

(e) Reduction operations permit a formal simplification of the representation of behavior into a form which is in principle inadequate but frequently useful to a person.

In short, the formulation of intentional action as a calculational system carries to a new level of technical detail the general notion that having the concept of behavior gives the person in-principle access to the entire range of possible behavioral facts (and that is the range of all possible facts whatever), including facts about himself and his own behavior and his relation to the world about him. The use of deletion operations, and specifically achievement descriptions, permits us to codify the facts of behavior in this way without being driven to the position that the calculus is embodied (enpsyched?) in some corresponding psychic machinery which goes into operation when a person behaves (witness current developments in grammatical theorizing).

The major interpretive tasks in intelligence analysis fit directly the notion of behavior description as developed above. That is, the aim of intelligence analysis is to achieve descriptions of what social practices they are engaging in, and where and when, what their capabilities are in various fields, what are their values, policies, strategies, etc., and how are all of these changing and trending? The very distinction of We and They which gives rise to the practice of intelligence analysis arises as an Individual Difference distinction on a group or paradigmatic basis. The aim, however, is to achieve correct descriptions. Thus although the formulation thus far provides a basis for collating intelligence data because it is rich in extensive relational implications, it will be of some relevance to present the next level of technical detail, where systematic principles for giving correct, or objective, behavior and person descriptions are formulated. These will be procedural, not deductive principles.

We proceed from formula (3), Deliberate Action, which involves not merely distinguishing behaviors, but also having grounds for choosing among them. Four major categories of grounds are found. These categories are characterized as <u>standards</u>, or <u>perspectives</u>, from which behaviors may be appraised. Particular appraisals which are made reflect the level of competence, or <u>Ability</u>, in the use of these perspectives. The four are the Hedonic, Prudential, Ethical, and Esthetic perspectives. Each of these provides a kind of ground for behavior choice. To see a prospective behavior as being pleasurable, as being in one's self-interest, as being the right thing to do or one's duty to do, or as being the fitting or appropriate thing to do <u>gives</u> one a reason to engage in that behavior and to choose it over alternatives which are less pleasurable, prudential, ethical, or correct or which are <u>un</u>pleasant, etc. When the behaviors in question consist of describing someone's behavior the esthetic perspective provides the relevant sense of "appropriate", namely, the giving of true, or correct, or objective descriptions. An individual in whose behavior this standard has priority is thereby constrained to restrict his descriptions in certain ways. These constraints reflect both (a) the logical characteristics of the subject matter, namely behavior, (b) the logical characteristics of the behavior, namely the giving of behavior descriptions, and (c) the relevant standard for choosing among behaviors, namely the esthetic standard. From these three bases jointly there follows a set of rules which are tautological in character. These rules embody the analytic relations among various sorts of behavioral facts, hence they provide the basis for bringing together various observations or reports of behavioral facts and appraising them from the standpoint of their mutual compatibility and thereby making it possible to use one observation as a check on the *empirical* validity of the other (and vice versa). The nine rules are given in Table 2. Further explication would be found in "Notes on Behavior Description" (Ossorio, 1969). The relevance of these principles to the interpretation of intelligence data is clear.

Table 2. Maxims for Behavior Description

1. A person takes it that things are as they seem unless he has reason to think otherwise.

2. If a person recognizes an opportunity to get something he wants, he has reason to try to get it.

3. If a person has a reason to do something, he will do it unless he has a stronger reason not to.

4. If a person has two reasons for doing X he has a stronger reason for doing X than if he had only one of these reasons.

5. If a situation calls for a person to do something he can't do, he will do something he can do.

6. A person acquires facts by observation (and thought).

7. A person acquires concepts and skills by practice and experience in some of the social practices which involve the use of the concept or the exercise of the skill.

8. If a person has a given person characteristic he acquired it in one of the ways that it can be acquired, i.e., by having the prior capacity and an appropriate intervening history.

9. Given the relevant competence, behavior goes right if it doesn't go wrong in one of the ways that it can go wrong.

5.4 Language as a Form of Behavior

The last major sector of the conceptual basis for the SA Information System is the placement of language in relation to behavior and to reality. Since language is clearly a form of behavior, the issue there is which form? In relation to reality, the brief answer is that the SACS has the status of "concept" and language provides the person's access to concepts because verbal expressions serve uniquely and indispensably to identify concepts; further, the statement that they do so is analytic, not empirical.

The general concept of verbal behavior, V, is given by formula (5). The relation of verbal behavior:

$$(5) \quad V = \langle C, L, B \rangle$$

to behavior is shown by the juxtaposition of formulas (1) and (5) :

$$\begin{array}{llllllllll} (1) & B = \langle & I, & W, & K, & KH, & P, & A, & ID & \rangle \\ & & & & \vdots & & \vdots & & & \\ (5) & V = \langle & & & C, & & L, & & B & \rangle \end{array}$$

In formula (5), C is a concept, L is a locution which stands in a one to one relation to C, (L is specified by giving a Performance description of the behavior of uttering the locution), and B is the class of behaviors which consist of acting on the concept C. Uttering L is thus a special case of B.

What is portrayed by the juxtaposition of (1) and (5) is that to say that a given behavior is a verbal behavior is to give a partial specification of two of the parameters of that behavior, namely, K and P. The performance is the verbal performance of uttering the locution, but of course, there is more to the performance than that. The SA Concept being acted on includes the concept C which is identified by the locution, but the SA concept would inevitably include more than C. Thus, verbal behavior is shown as literally a kind of behavior and its special character as language is carried by the structure of formula (5).

The way in which L is indispensible with respect to C in formula (5) follows a pattern which is familiar in logical theory and in modern technology. The following are two examples of this pattern:

(A) Russell and Whitehead's classic definition of cardinal number was accomplished by first defining a procedure for generating well-formed formulae. When the procedure is followed recursively the result is a series of formulae all having the same general structure, but differing in the number of parenthesis-enclosed expressions. Each formula represents a class having a specific cardinality, and the cardinality is the same as the number of parenthesis-enclosed expression. Thus, it is easy to"read off" the cardinality of a particular class identified by this type of formula, since one has only to count the parentheses and that will be the cardinal number of that class. Against this background, their definition of cardinality was easily given: A cardinal number, say"5", is the class of all classes having <u>the same</u> cardinality <u>as this</u> class, which by definition has that cardinality. In the case of "5", "<u>this</u> class" would refer to the class given by the five-parenthesis formula.

(B) In a parallel fashion, the standard meter stick in Paris is used to define the unit of length of one meter. The definitional formula is parallel to the definition of cardinal number: The length of one meter is the length of any interval having the same length as the length of this object, which by definition has that length.

We may now give a definition of "concept" in the same form: The concept C is the class of all behaviors which include in the values of their K parameter the same concept as this behavior, L, which by definition includes that concept in the value of its K parameter.

Recognizably, this definition of concepts is simply a consequence of the analysis of behavior in which the values of the K parameter are concepts. What is accomplished by the definition above, by reference to the locution, L, is the introduction of units of reference (since there is no quantification here, we cannot talk about units of measurement, and so must use the more general expression). It is universally agreed that the introduction of such units is crucial for gaining technical leverage in dealing with the domain in question. Such units permit us to deal with the phenomenon in detail, and in practice. Having units of length, for example, permits us to distinguish different lengths, identify cases of particular lengths, and relate different lengths systematically. These advantages are not duplicated by reductive or transcendental definitions which purport to tell us what length really is.

Likewise, the introduction of verbally anchored units of reference in the domain of concepts makes that domain available to us _in detail_ and _in practice_. Having such units permits us to distinguish different concepts, identify exemplars of the same concept, and systematically relate different concepts to one another. These advantages are not duplicated by reductive or transcendental definitions which purport to tell us what language or concepts _really are_.

Since traditional analyses of language are hardly more than analyses of the Locution in formula (5), it is worth commenting on why the other two terms are essential. Briefly, we may say that without C the locution would be _meaningless_ and that without B it would be _pointless_. The concept identified by the Locution is its _meaning_; the behavior (intentional action) accomplished by the performance of uttering it (the locution) is its _use_. Without C, the performance of uttering L would be merely vocal behavior, noise, not verbal behavior, language. The fact that there are a variety of non-verbal behaviors which, for a given C, qualify as acting on that concept gives point to the explicit identification of that concept, for without that means of identifying concepts we would be unable to distinguish any value of K from any other value of K, and hence would be unable to distinguish one behavior from another. [From the standpoint of reflexivity, and in light of formula (3), note that if we were in that kind of bind we would be unable to represent that fact, since we would be unable to distinguish the behavior of truly describing that bind from the behavior of falsely describing it or from any other behavior.]

It is also worth pointing out that the present formulation of language as verbal behavior (rather than, per se, being, e.g., grammatical competence) takes us beyond the limits of syntactic and semantic theorizing to the full range of linguistic behavior. Formula (5) which expresses C as the meaning of L works just as well for "Please pass the salt" and "At last!" and "Would that Eve had never seen the apple" as it does for descriptive locutions such as "The cat is on the mat." The SA System as such codifies the traditional linguistic notions of (a) extensions or referents (objects, processes, events), (b) intentions, or meanings (concepts, relations), and propositions (state of affairs concepts).

Since in formula (5) both L and B corresponds to conceptual distinctions,the concept of verbal behavior is progressive in both these terms:

(5) $V = \langle C, L, B \rangle$

(6) $V = \langle L, L', B' \rangle$

$V = \langle L', L'', B'' \rangle$

etc.

(7) $V = \langle B, L', B' \rangle$

$V = \langle B', L'', B'' \rangle$

etc.

That is to say that the concept of a given locution L in formula (5), being a concept, calls for a second kind of locution,L', which stands in one to one relation to that concept, and correspondingly, a class of behaviors, B', which consist of acting on the concept of L. But then L'

corresponds to a new concept, which calls for still a third kind of locution L" and a corresponding new class of behaviors, B", and so on.

Likewise, the concept of a given behavior B in formula (5), being a concept, requires a second kind of locution, L' and a new, corresponding, class of behaviors, B', etc. In this way the one to one relation between verbal behaviors and behaviors is exhibited (by formula (7)). Technical issues stemming from the progressive character of formula (5) are discussed at some length in "Meaning and Symbolism" (Ossorio,1969).

Formula (5) is not only progressive in L and B, it is also recursive in C and B:

(5) $V = \langle C, L, B \rangle$

(8) $V = \langle \langle V \rangle, L, B \rangle$

(9) $V = \langle C, L, \langle V \rangle \rangle$

(10) $V = \langle \langle V \rangle, L, \langle V \rangle \rangle$

That is to say, the concept in question may be the concept of a verbal behavior; the behavior other than L which qualifies as acting on the concept C may be verbal behavior; and both may be the case simultaneously. These recursive features follow from (a) the recursive formula (2) and (b) the statement that the C parameter of formula (5) is an incomplete specification of the K parameter of formula (1), together with the fact that the B of formula (5) is the B of formula (1).

On the view presented here, grammatical theorizing on natural language is a form of behavior which exemplifies formula (10). That is, it is verbal behavior which consists of acting on the concept of verbal behavior. Specifically, the behavior consists of attempting to give a paradigm case formulation of the <u>performance</u> of saying something. That is, the theory is an attempt to portray how one must go about it if one is to be <u>guaranteed</u> to be speaking <u>correctly</u>. The acceptance of any given grammatical theory is largely a matter of individual preference and the level of the state of the art. The present formulation is entirely compatible with current generative grammars in this respect and can incorporate them as technical detail. In no sense, however, is any of the present formulation dependent on whether the particular generative grammatical formulations are correct or even on whether the entire genre survives.

Finally, some comment is called for in connection with the one to one relation between locutions and concepts. Ordinarily the assertion of such a relation, though admittedly attractive methodologically, would be dismissed out of hand as directly contrary to fact (the facts of synonymy and homonymy) and generally suspect even on prior grounds, namely that locutions are in some sense arbitrary (the arbitrariness of conventional assignment of <u>this</u> word to <u>this</u> concept which it <u>stands</u> <u>for</u> (sic)).

In a reflexive conceptual framework in which any principle may be the subject of its own application it is misleading to think of general statements or principles as simple pictures or representations of the way the world is. It is considerably less misleading to think of them as procedural rules which may have a different result depending on whether they are applied recursively and in conjunction with which other rules. (One might talk here about the "deep structure" of the IA System which is given by the present formulation, and contrast that to a "surface structure" consisting of the concrete behavior of particular individuals or groups.)

In the present case, a recursive move solves the primary difficulty: Two locutions L and L_2 are used as locutions for a given concept C, per formula (5), and (b) the concept of L_1 and L_2 having the same meaning is formulated as the C of formula (6), where "acting on C" reduces to "treating C as being so." In short, L_1 and L_2 have the same meaning when there is a way of saying that they do and when our social practices of acting on that notion consist of successfully treating it as being the case. (In the IA formulation, the basic notion is no longer "the real world" in its role of being what our language is about, but instead the notion of reality constraints as the boundary conditions which correspond to the restrictions on which behaviors persons are able to engage in.)

Conversely, the same locution has different meanings if it can be used to identify two different concepts C_1 and C_2. In that case it is possible to identify the two concepts and to distinguish (a) the behaviors which differ only in this respect, call them B_1 and B_2, and (b) the two classes of behaviors, per formula (7) which constitute acting on the concept of B_1 and acting on the concept of B_2. If we designate these two classes of behavior as B_a and B_b, respectively, then we can say that although the utterances are indistinguishable, the concepts are not because B_1 and B_2 are distinguishable, and if these are not distinguishable "on sight" they are distinguishable by virtue of the difference between B_a and B_b. To explain, for example, that "He couldn't bear it" could mean either "He couldn't stand it" or "He couldn't carry it" is to distinguish B_1 and B_2. By showing that each of B_1 and B_2 is equivalent to a case of B, the ambiguous behavior, and yet are distinct from each other, it follows that the possible cases of B fall into at least two distinct groups.

The element of conventionality in language has already been elucidated by assimilating it to the introduction of units of reference. Generally, the selection of a given unit of reference could have been replaced by another, but the selection of <u>some</u> unit cannot be dispensed with. Moreover, statements couched in terms of one unit would in principle be translatable into statements couched in terms of another. The treatment of systematically related meanings would follow exactly the same form as the treatment of 'identical' meanings of synonymous locution. So also would the equivalence of spoken and written forms.

6.0 THE IA SYSTEM IN INTELLIGENCE ANALYSIS

The Intentional Action System, which encompasses formulations of persons, language, the world, and the interrelationships among them, is, of course, of much broader scope than an ad hoc formulation of the intelligence analysis problem. Thus, it is appropriate to look back and summarize the major relevance of the IA formulation to intelligence problems before proceeding to the delineation of an information processing system which represents a first effort to exploit the basic IA conceptualization in this respect.

(A) The IA formulation permits an explicit, systematic representation of the phenomenon of intelligence analysis, i.e., a person reconstructing some aspect of the real world, including himself and his sources and the persons who will be affected by his decisions and actions. It thus permits a conceptually coherent approach to the intelligence problem as contrasted with technically oriented, ad hoc approaches.

(B) It permits an explicit representation of the domain of facts which are of interest. Because we have a representation not merely of particular facts, but of their place in a domain of facts, it permits the accumulation of isolated explicit facts within a single framework which maximizes the potential for recognizing what the isolated facts "add up to."

(C) Technically speaking, the pre-empirical conceptual structure is infinitely rich in relational information which can in principle be exploited in dealing with "implicit information" when descriptive constants are introduced such that they define a domain of interest and a body of fragmantary information concerning that domain.

(D) By systematizing the notion of the <u>use</u> of information (acting on a concept), the formulation helps to bridge the gap between 'hard facts' and 'interpretation' or between "mere facts" and "useful facts".

(E) Since individual differences are systematized within the IA formulation, objectivity is not confused with uniformity. Relativity appears in at least three significant areas: (1) Some number of intelligence problems are hopeless as they are now stated, because there is no such thing as the single, definitive, 'objective' answer demanded; instead there may be complex partial answers which are differently derivable, depending on the use to which the answer is to be put. (2) Since individual differences are simply parameters of the information system, reconstructions or solutions can in principle be generated relative to a specified set of beliefs, practices, purposes, individuals, or groups within the intelligence community. (3) Since the foregoing applies to the representation of "They" as well as "We", it provides some degree of built-in safeguard in dealing with the 'facts' of the behavior of persons whose customs and purposes are substantially different from ours.

(F) The greatest difficulty connected with computer implementation of the IA System, even for limited purposes and domains such as intelligence analysis, lies in the complexity of the design which is required and the degree of information which must be incorporated in the structure of the system from the outset. Because the value of a given data structure increases when more data structures are added, a large scale system may be expected to be more effective than a small scale experimental system.

In any case, relative to the data structures incorporated in the relational data systems surveyed by Auerbach, the amount of data required in the construction of a SAIS appears to be of different order of magnitude.

(G) The application of the IA System to the problem of automatic intelligence data processing involves the computer implementation of a quasi-formal system rather than a formal system. This procedure in no way prevents us from making use of formalization whenever and wherever it is apropos. Conversely, it should be emphasized that the behavioral approach is not merely a way of talking about something that is 'really' a formal approach. Behavioral concepts such as "ability", "achievement", etc. can be used directly in dealing with computer systems. (We have already largely anthropomorphized computers in our informal references to them.) The overall design of the SA Information System as an organization of subsystems corresponds roughly to the notion of a meaningful range of real life achievements being accomplished by bringing to bear a limited variety of fundamental abilities within the means-end framework of a task analysis.

(H) One consequence of proceeding on the basis of the IA quasi-formalization is that the accomplishments of the SAIS in intelligence data processing cannot be deduced from a description of the system, and this will hold even when a description of the data and data formats is given. Of course, one may judge a given design to be entirely reasonable as contrasted with others which would be clearly defective or hopeless or irrelevant, but that is to proceed on a quasi-formal basis. With respect to the SAIS this simply means that it could be warranted only experimentally and not deductively. With respect to the appropriate description of the SAIS it means that intelligibility is likely to count for more than fine technical detail. At any rate, this is the view which has guided the presentation which follows.

7.0 THE STATE OF AFFAIRS INFORMATION SYSTEM

The following is a discursive specification of the information processing system which is given in greater technical detail subsequently.

(A) First, the most general design format corresponds to the notion of a social practice (of formula (4) of the IA System) at the level of "We" and "They", or the general phenomenon of intelligence analysis. This degree of generality is required in order to generate the descriptive concepts which identify what constitutes relevant or useful information. A process (of which a social practice is an example) is a sequential change in a state of affairs(Transition Rule 4). In characterizing a social practice as a process, we go through several analytic procedures: (1) We distinguish the sequence as consisting of some number of stages— S_1, S_2,---S_n. (2) We distinguish some number of objects, including human objects, as elements or participants in the process. (3) At some level of analysis the sequence consists of actions by individuals. (4) The structure of the process is given by specifying (a) which individuals are eligible to engage in which of the actions, (b) what options there are for alternatives within each stage, and (c) what contingencies restrict the selection of options at a given point as a function of any state of affairs at other (ordinarily, earlier) points.

(B) It is the contingencies in the social practice which determine what constitutes relevant or useful information to a participant, since the relevance or usefulness consists in the fact that the participant will do something different depending on the content of the information. Thus, if one of the participants is We and another is They, and if We is represented by the intelligence analyst, or the consumer of the analyst's reports, then We will do things differently depending on what They can do, intend to do, are engaged in doing, know about or don't know about, etc. These primary contingencies define one of the major subsystems of the SA Information System. This is the Fact-Event Recording System (FERS). The function of this system is to carry an updated summary of the reconstruction of these facts which are of primary interest, so that they are readily available for retrieval. Adjustments in both directions are made for technical convenience, i.e, representation of primary facts elsewhere and representation of other facts in the FERS.

(C) In practice, an experimental SAS would be restricted to some subdomain of facts relevant to intelligence analysis and the specification of the domain would be achieved through consultation with members of the intelligence community who would, collectively, be represented as "the analyst" being simulated by the SAS.

(_D_) The SAS is designed to operate on simple factual statements. In the absence of an effective technology for processing English text, the present system is designed to accept input which is formatted in terms of the predicate calculus or simple lists interpreted as values of functors. Heuristic programming for accepting English text or some lesser degree of formatting may be incorporated later and evaluated for its marginal utility.

(_E_) Although the FERS is a reconstruction of "the real world", it will exhibit various kinds and degrees of arbitrariness and heterogeneity which are the consequences of the analyst's selective interests. A standard historical reconstruction is also provided by the Operative Time System. Histories of persons, objects, institutions, and practices are recorded here _as_ historical processes, all coordinated to a single calendar-clock.

(_F_) Certain other subsystems provide representations of processes (Process-Activity System), part-whole structures (Part-Whole System), and Means-Ends structures (Means-Ends System). These three, which are abbreviated as PAS, PWS, and MES, contain two kinds of ingredients: (1) basic schematic units and paradigm cases and (2) subroutines for performing compositional and decompositional operations on the basic units. In general, PAS, PWS, and MES schemas represent structures which _include_ as possible elements the very sorts of entities which appear in the FERS as being of special interest to the analyst.

To take a variation on a classic example, suppose the analyst is interested in a particular factory and his interest is primarily in the level of production of solid state fuel and in whether factory is a site for some basic research. The PAS would (ideally) contain an extensive, though still schematic representation of the various processes for producing various solid state fuels. Suppose also that it is discovered that three of the plant employees are glass-blowers and one is a machinist. Checking through the various processes might disclose none in which custom glassware played any essential part, in which case the conclusion would be drawn that either (1) a new process has been invented or (2) something else, quite possibly research is going on. Conversely, if there were fuel-making processes in which quick mending of glassware were essential, the conclusion might be drawn that either (3) they are running three assembly lines or three shifts or (2) something else, possibly research, is going on. Further, checking the history of that plant as stored in the OTS, might indicate that (4) there hasn't been time to add assembly lines since the time when they were known to have only one, or (5) that there has been time, and if they are now running three lines this represents a change in policy. In any case, concluding that three lines are in operation would carry some implications for the level of production, which was one of the points of primary interest.

Thus, the Process-Activity, Means-Ends, and Part-Whole schemas generate the functional equivalents of a number of relational statements which can be used to draw conclusions of interest to the analyst. They also illustrate the statement that a SAIS requires that a good deal of structured information be put into the system before it can begin to process a flow of information.

(G) The Classification Space provides subject matter indexing in which documents are not assigned to particular subject matter headings, but instead are quantitatively related to every subject matter within the domain of interest. Document indexing is carried out automatically by making use of a finite vocabulary of terms which were indexed psychometrically. Thus, the index consists of an n-dimensional space having as reference axes a set of interpreted subject matter dimensions, and documents terms, and requests are indexed by being assigned a set of coordinates in the C-Space.

The C-Space serves two major functions. The first is simply to provide document indexing and retrieval independently of the rest of the system. The second is to provide an order of priority for the inspection of process and other schemas of the PAS, PWS, MES, and OTS. Since such schemas will be indexed in the C-Space and since incoming documents are indexed automatically, each document is automatically interpreted as a request for implications (see below) and the relation of various schemas to this request in the C-Space is interpreted as their degree of relevance to this request. Given a priority ordering, the search process may be terminated in various ways, e.g., an automatic C-Space distance criterion, a manual override, an external contingency, etc.

(H) The Attribute Space (A-Space) provides conceptual content indexing in a way that is technically entirely parallel to the Classification Space. It, too, serves more than one major function. The first is to provide document indexing and retrieval in conjunction with the C-Space. The second is to serve as a limiting resource for the compositional and decompositional procedures of the MES in seeking to identify alternative means for reaching a given end when such alternatives have not been specified as such. The heuristic principle followed here is that two things are interchangeable if they have the same attributes (operationally, approximately the same location in the A-Space) and they are interchangeable for a given purpose if they have the same relevant attributes. The relevance of attributes to particular topics or practices may in turn be represented directly as such in the C-Space or as portions of contingency rules governing PAS schemas or in the functor subspace of the Attribute Space (see RADC-TR-65-314) together with normal C-Space indexing.

(I) The Deductive System (DS) consists primarily of deductive procedures for implementing the predicate calculus. In addition it will contain meaning postulates with respect to both elements and relationships which are represented in the FERS. Thus it is designed primarily for dealing efficiently with those facts which are of primary interest. The DS will also contain representations of theories as sets of statements which have the same deductive status as meaning postulates. Meaning postulates and theories are subject to two controls: (1) Deletion and rewrite by a user and (2) data contingencies, e.g., specifying whose theory it is.

Although it would be possible, with some maneuvering, to give the PW, PAS, and MES schemas the status of meaning postulates also, it is an open question as to whether it is worth the effort in a prototypical system since (1) it is unlikely that the internal structures of these schemas would support much strictly deductive inference and (2) to use the schemas for plausible inference would require operational criteria which were (a) over and above the deductive process and (b) could be carried out without an overlay of strictly deductive procedures. Instead, it appears to be technically simpler to use "possibly" and "apparently" as validity tags in the FERS representations based primarily on part-whole pattern matching within the PAS, PWS, and MES.

(J) The Inductive System (IS) is considerably more important in the long run than at present, since it is the locus for concept formation and learning principles. However, at present, we specify only that it performs inductive counts on the use of the meaning postulates and the criteria for plausible inference. The inductive count consists of establishing an error ratio for each rule. The error ratio consists of the proportion of conclusions drawn by the use of that rule which are rejected at that time or later as being inconsistent with observation or better-founded data. The system can, of course, be programmed to search for optimal degrees of part-whole matching for use as criteria for plausible inference. An example of part-whole matching would be: If they have the pipeline, the steel, and the chemicals needed, do they have the capability for a refinery, which would call for engineering manpower, financial support, and x,y, and z in addition ? The degree of part-whole matching is simply the extent to which all of the separate elements of a known unit are known to be present. Since the "known units" are represented as units in the PAS, PW, and MES whereas they are represented in the other systems only as names (except the OTS), criteria can be applied in these three subsystems without reference to a general deductive capability.

(K) As shown in Figure 1, the gross structure of the routine processing of input is relatively simple. The first major operation consists of checking input for consistency with existing information in the FERS and OTS systems and adjudicating any inconsistencies, primarily by reference to a grading of sources and the number of independent indicators involved pro and con. Secondly, purely deductive consequences of the input statement are derived using the Deductive System and the "meaning postulates" available to it. Following this, the non-deductive consequences are searched for by matching the input statement to the PW,PAS, and MES schemas as they are instantiated in the OTS (primarily) and FERS. The point at issue here is whether the input statement renders a given schema any more or less applicable (in terms of pattern matching) to portions of the OTS and FERS which involve the elements mentioned in the input statement. If criteria for "possibly" or "apparently" are crossed, a corresponding non-deductive conclusion is drawn. When all the statements in a given document are processed in this way, the entire collection of inputs and implications is cross-checked for internal consistency, and any inconsistencies are adjudicated. Inputs and implications which survive are recorded as states of affairs in the FERS and OTS. All statements, including those rejected as factual, are recorded in the OTS as having been asserted or implied by some source. Finally, success ratios for the rules used in drawing either deductive or non-deductive conclusions are updated and the document is indexed in the C-Space and A-Space, together with an optional additional commentary as to its validity and significance.

(L) As noted initially, the structure (elements and attributes) and content of the FERS is designed to correspond directly to those questions which are of primary interest to the analyst. Thus, retrieval of information from the FERS is a relatively simple matter of responding to a formatted specification of (a) a set of elements, including a set defined by a given attribute, (b) a set of attributes and (c) a structure of either "and" or "or" connecting the attributes. In this regard, it may be noted that the names of PW, PAS, and MES schemas would be represented in the FERS as the names of attributes. In this regard, the "contingency unit" of the FERS basic unit (see below under "FERS") is particularly relevant. Referring back to the earlier example, a query "Do they have the capability for a refinery?" The answer, in terms of the resources described so far, would be (1) Not known completely; (2) Apparently so,

Figure 1. SAS Gross Input Processing

```
↑ Input a statement
↑     ↓
↑ Interpretable?  No→ Buffer C and output for interpretation
↑    Yes
↑     ↓
↑ Consistent with FERS and OTS?  No→Adjudicate in IS
↑    Yes                                  ↓
↑     ↓                                   ↓
↑ Using DS,FERS, and OTS  ←————————————
↑
↑    Draw implications directly
↑    Draw implications via MES,PWS, PAS
↑     ↓ using A-Space, C-Space for priority ordering
↑     ↓
↑ Using DS,FERS, and OTS
↑
↑    Are Implications Compatible with FERS,OTS
↑    Yes                  No ————————→   Adjudicate in IS
↑     ↓                                       ↓
↑ Store in Buffer B ←————————————————————
↑     ↓
↑ Last Statement?
↑    return←——No      Yes
                       ↓
                       ↓
  Using IS             ↓
     With DS, cross-check implications for mutual consistency.
     Record acceptable input and implications in FERS,OTS.
     Record unacceptable input and implication in OTS as reports only.
     Update success ratios.
       ↓
     Index documents in A-Space and C-Space with record added of unacceptable
     statements and implications.
```

if they have financing, technical manpower, and xy,y, and z, since they have pipeline, chemicals, and steel. That is, the "matched" elements of the refinery schema would appear in the "since" clause and the non-matched elements would appear in the "if" clause.

Questions concerning the histories of objects (in the broad SAS sense defined by transition rule 3) would be directed to the OTS. An example of such a question would be the request described as "Print the number of reports issued by the Academy of Science in the field of ionic propulsion during 1968 which were produced on projects supported by the Ministry of Defense." In the working SAS format the query would be substantially more structured: (a) It would include a conventional instruction "Count the n members of List R and print n, R." (b) List R would be defined as "History Unit of P from Date 1 to Date 2" with P= Academy of Science, History Unit = Scientific Publication Schema (PAS), Date 1 = January 1, 1968, Date 2 = December 31, 1968. (N.B.: The PAS schema would be a social practice schema such that, among other things, research procedures A are carried out by B, supported by C, written about by D on topic E, published by organization F in publication titled G on date H.) (c) Count unit would be defined by F = P = Academy of Science, C = Ministry of Defense, E = ionic propulsion, and H = 1968. Where the required information is absent (e.g., a report which fits all the specifications except that the source of funding is unknown) the "Apparently, if..., since..." format is available.

More generally, all the subsystems except the Deductive System and Inductive System may be treated as distinct indexing systems reflecting different indexing principles. Each may be used directly as such to answer a specific kind of query, and all may be used jointly to answer complex queries which can be broken down into some and -or combination of simple queries of the Part-Whole, PAS, FERS, etc. sort. In an interactive system the compositional and decompositional subroutines for PAS, PWS, and MES schemas would offer the active counterpart of "browsing."

(M) It was stated in (A), above, that the most general design format corresponded to the notion of a social practice at the level of "We" and "They", i.e., at the level of the general phenomenon of intelligence analysis. It was stated further that it was this form of representation which differentiated between mere facts and relevant or useful information, since the latter correspond to the conditions under which "We" would make different choices among the alternative ways of participating in that practice. These statements may be amplified as follows.

(1) That social practice may be explicitly represented in the SAIS as one of the PAS social practice schemas. Let us call this schema PAS-1.

(2) Among the elements (roles) of PAS-1 at some level of analysis(decomposition) will be some number of intelligence analysts. Among the possible alternatives at a variety of points in PAS-1 there will be the particular alternatives of one analyst being asked a question by another analyst, going through certain investigative and decision-making performances

and achievements, and answering the question.

(3) This performance of this role will be subject to further constraints which are contingent on the ID characteristics of the analyst being interrogated.

(4) One set of ID characterizations of the analyst being interrogated defines that analyst as the SAIS itself and the particular performances selected by that contingency are simply the operations of the SAIS.

(5) Thus, the SAIS may contain a representation of itself as a historical individual and of its current and past behaviors and their social significance. This provides a certain kind of technical implementation of the notion of reflexive closure of person and world, and the "name and description" structure of the PAS schemas provides a certain kind of technical solution to the problem of infinite regress in self-representation.

(6) Self-representation is one of the primary conditions for significant advances beyond the primitive SAS delineated in this report. Among the anticipated improvements are (a) refined control of its own internal and external operations and (b) the incorporation of learning principles to create an adaptive system at a far greater logical depth than the "success ratio" counts mentioned above.

(7) An example involving self-representation in the control of internal

and external operations is described below (the interception problem).

8.0 THE SAS SUBSYSTEMS

In this section a more detailed description is given of the data structure and the operation of the several subsystems of the SAS. It is convenient to divide the subsystems as a whole into seven representational, or "content," systems and two procedural, or "operational", systems. The latter two are, of course, the Deductive and Inductive systems.

Emphasis is given to the seven representational systems, since the Deductive System is envisioned as essentially state of the art and the Inductive System is envisioned as primarily relevant to systems which are more elaborate and more advanced than the present SAS, so that at present it performs only a few primitive functions. The representational systems taken as a group provide a certain kind of hierarchy of degree of completeness and organization, and correspondingly, of representational power.

The most highly organized and complete representation is found in the OTS chronologies. These involve object-process configurations which are instantiated in terms of historically particular individuals and times and places. The next most highly organized are the PWS object-process configurations which are the same as the OTS chronologies but without a historical anchoring; the PWS configurations are the patterns of which the OTS chronologies provide historical examples. At the next level are the other PWS configurations and the PAS schemas which are temporally ordered and logically structured. At the next level are the MES schemas which resemble the PAS schemas but have less order and structure. Roughly, a PAS schema represents how some goal is accomplished, and this requires specification of what elements are involved and how they enter in. In contrast, the MES schemas deal primarily with the question of what are the relevant elements for achieving a given goal, and it is only when they are elaborated in detail as to where and how the elements are relevant that the MES schemas approach the structure of PAS schema.

The A-Space and C-Space are limiting cases, more or less. In these there is representation of some extremely simple, general characteristics such as subject matter relevance, properties, attributes, etc. These represent a minimal degree of organization or representational power; they correspond to single, isolated facts. The utility of the A-Space and C-Space stems from the fact that they deal with kinds of information which are significant and that they function as implicit rules for generating an indefinitely large number of new facts of the same kind.

Finally, the FERS is a miscellaneous collection of representations which may be at any of the levels of organization distinguished above. The content of the FERS is heterogeneous, and although there is a formal organization of "elements and attributes" there is no logical organization which holds the various elements and attributes together in some coherent way (that function is performed by the OTS). This looseness is comparable to that of the Means-Ends schemas in which the relation of various component means to one another need not be specified. Indeed, the FERS is simply a collection of means which the analyst has identified as means to his ends -- he needs that information in order to accomplish his job.

Thus, the subsystems as a group may be regarded as reflecting the methodology of the "Deletion" operations in the calculational system of behavior. That is, they provide a technical framework for giving incomplete descriptions as incomplete descriptions, and there is a range of forms of description from more to less complete. The importance of descriptive forms which are per se incomplete is that (a) they permit us to represent explicitly whatever we do know (or are willing to make a commitment on) as contrasted with what we don't know and (b) they serve as antidotes to the temptation to reify our incomplete descriptions; they serve as reminders of what it is we don't know and its relation to what we do know, and it is by virtue of this kind of relationship that we are able to piece together fragments of information into a coherent whole.

8.1 The Fact-Event Recording System

The FERS is probably best thought of as a "Big Board" display which consists of two-dimensional matrix of rows (elements) and columns (attributes, functors), with the intersections corresponding to (a) a yes or no decision (i.e., the element either has the attribute or it hasn't), or (b) a number which is the value of a functor or (c) a list which is the value of a functor. For example, "The list of reports published by X from T_1 to T_2" defines a functor having an individual and two dates as its arguments and a list of publications as its values. The "Big Board" aspect of the FERS is represented by the basic fact unit in the FERS data format:

Basic Fact Unit = R (P, Q, S)

Where R is a 1 - or 2 - or 3 - place attribute or functor

and P, Q, and S are 1st-, 2nd-, and 3rd-place members, respectively. Complex relationships involving more than three elements are represented either by PAS, PWS, or OTS schemas or by recursive use of the BFU:

R' (T, U. R)

R (P, Q, S)

The FERS data format contains three distinct units in addition to the Basic Fact Unit. The additional units are primarily in the interest of simple retrieval.

The second unit is the Contingency Unit, which states the conditions, if any, under which the information contained in the BFU may be considered factual. The data form of the Contingency Unit is a list of BFU's with logical connectives and the special symbol "since" as an option (for the "if..,since..." output). Data in the CU will normally come from system calculations rather than direct input.

The third unit is the Information Pathway Unit, which traces the history of the information contained in the first unit (BFU). The data form is any one of several PAS schemas, of which a paradigm would be "X, Y, Z are responsible for fact BFU at T; this became known to P who communicated with List A who communicated with Org which published it in Pub, where it became known to List B ending with immediate source, then myself, and it has been communicated to User List C"(In the schema, all events are dated). Ordinarily, only the immediate source, the organization (if any), and the publication (if any) would be routinely received as original input. The remainder would either be calculated or be received as additional input.

The final unit is the Currency Unit which contains a list of dates on which the information in the BFU is considered to have been checked on or verified. The original report date is one. Being implied by an accepted report constitutes a check, as does additional direct observation.

Thus, in general, the additional units provide each primary fact with some degree of appraisal and historical context. Keeping the auxiliary units updated would be an internal, or housekeeping, function of the system rather than a retrieval procedure.

In summary, the data format for row-column intersections in the FERS is as follows:

Basic Fact Unit: R (P, Q, S)

Contingency Unit: List of BFU's with connectives

Information Pathway Unit: An instantiated PAS schema

Currency Unit: List of dates

8.2 The Process-Activity System

The PAS consists of several distinct components. Primarily, however, it consists of a library of process schemas indexed by name and having the form (i.e., name + description) of the basic process unit described below. Other components are ancillary in that they provide resources for operating on or with these schemas. Although behavioral processes (social practices, activities) will be substantially represented among the PAS schemas, the latter are not restricted in principle to any particular kind of process.

(A) The Basic Process Unit (BPU) has the following structure:

Name A

Description A, which specifies

(1) A list of stages I-N, i.e., Name A1, Name A2...Name AN.

(a) Options, or alternate possibilities (alternate forms) of each stage (e.g., Name A11, Name A12, ...Name A1K for stage Name A1).

(b) An event description of each option, i.e., a description of the kind of change represented by the occurrence of that option.

(2) Elements (roles) in Name A.

(3) Eligibilities: which elements are eligible for which parts in the schema.

(4) Contingencies (occurrence constraints), including those which connect

(a) Any subset of I-K stages (i.e., some stages may be optional.)

(b) Name A to any external states of affairs (including processes) up to 6 in number (an arbitrary limit), either earlier or later than Name A, and needing only to be identified, not described.

(c) States (Name A1, Name A2, etc.) to any external states of affairs as in (b).

(d) Stages to alternatives within stages (e.g., Name A1 to Name A26).

(e) (Most important) options within stages to options within stages (e.g., Name A13 to Name A52).

Two comments are apropos here. First, the combination of event description (1b, above) and contingency structure (4, above) permit the representation of continuous processes even though this is not the primary concept of a process. Second, the structure involves the name of a process which is coordinated with an explicit description which includes some names of other processes. Because the Basic Process Unit <u>need</u> not be very complex, it is not ipso facto technically unmanageable as a unit. At the same time, the presence of process names within the description gives the BPU an implicitly recursive structure, since (a) Name A may itself be one of the names within the description of a larger process and (b) the names within the description of Name A may in turn be expanded by giving them descriptions which contain further names. Thus, the compositional and decompositional operations discussed in connection with the basic transition rules are permitted by the Basic Process Unit. The subroutines identified below provide a kind of technical implementation of these operations.

(B) Analysis Subroutine

This subroutine "expands" the description of Name A by searching the library of BPU's for names which correspond to the names within the description of Name A. The subroutine would itself have a recursive structure so that selective or non-selective expansion could continue up to some criterion (see below) or arbitrary limit.

The Analysis Subroutine would be responsive to two kinds of queries, namely, (1) "If Name A has occurred, has Name Q also occurred?" and (2) "Is there a way (a) to accomplish change X or (b) to get from stage Name A3 to stage Name A5?" Queries of type (1) would arise, for example, in a refined procedure of drawing conclusions from the occurrence of Name A. Queries of type (2) would arise in the course of Means-Ends composition (see below) which is analogous to PAS composition.

(C) Identity Subroutine

This subroutine operates on a housekeeping basis primarily. It searches the BPU library, performs various expansions via the analysis subroutine, and for each BPU produces a list of its roles in various larger processes. Its list statements would have the form "Name A is the same process as Name Q1, Name R32, Name W5, Name M634, Name K1385, ..."

(D) Composition Subroutine

Given a role statement such as "Name A = Name M634" the Composition Subroutine expands Name M throughout to the level of detail of Name A.

Reconstructions of this sort are essential to the pattern-matching procedure described above, where the question would be, "Given that Name A has been reported to have occurred, is Name M thereby occurring also, and if so, to what extent?" The contingencies of both Name M and Name A would be involved in drawing conclusions from the occurrence of Name A.

(E) Process Paradigms:

For each Basic Process Unit, at least one "concrete" example of it will be given. Such paradigms increase the specificity of pattern matching and help predict the future course of events. For example, suppose that Name M is a 10-stage process of which Name M25, Name M46, Name M59, and Name M63 (i.e., instances of stages 2, 4, 5, 6) reported as occurring. The set of four substage options is not merely matched against Name M so as to conclude that stages 1 and 3 apparently have occurred. They are also compared (either by pattern matching or by A-Space similarity) with the paradigm cases of Name M. If the set of four stage options closely matches one of the paradigm cases of Name M, then this provides a basis for predicting that Name M will continue along the lines of that paradigm case.

8.3 The Part-Whole System

In the SA Conceptual System, objects and processes are the Categories primarily involved in the composition and decomposition operations which are so centrally involved in generating representations of the real world. In both cases the technical problem is to achieve representations which are both finite and explicit, hence technically manageable, and at the same time have an implicit recursive structure so that they can be related to one another along compositional and decompositional lines. Thus, it would be possible to deal with part-whole representation in a way which was very closely parallel to the procedures described for process representation. Instead, the procedure described below is generally parallel, but is modified in the direction of a more complex basic unit and simpler subroutines. This both seems more natural and promises a greater breadth of experience when both are incorporated into an experimental system.

(A) The Basic Object Unit (BOU)

The basic object unit is given by formula (11):

(11) ObUnit = < N > , < D2 > , < D3 > , < D4 >

That is to say that the general representation of an object is accomplished by specifying the members of the following four classes:

(1) < N > : A set of names which are the names of this same object. The limiting case is, of course, a single name. The degenerate case is the absence of any name and the identification of the object by means of one or more descriptions of type D2, D3, or D4.

(2) < D2 > : Class membership descriptions of this object: attributes and functor values as in the FERS.

(3) < D3 > : A set of constituents and a subset of immediate constituents in configuration. (A configuration is an object decomposed per transition rule 3, i.e., into a set of <u>related</u> smaller objects. The configuration is specified by specifying the objects and relationships (or functors) which are aspects of the configurational state of affairs and are therefore implied by it.)

(4) < D4 > : A list of configurations of which the object is part. Three important types of configuration are (a) single larger object, (b) a static arrangement of objects, and (c) a process involving objects.

In the part-whole format < N > performs some of the functions corresponding to the PAS identity subroutine and < D4 > performs a function analogous to the PAS Composition Subroutine. Correspondingly, < D3 > is the equivalent of the Description in the BPU. The combination of < N > and < D3 > is comparable to the name and description which constitute the Basic Process Unit.

(B) The Analysis Subroutine:

This subroutine uses the Basic Object Unit recursively to expand the parts of an object as configurations which themselves have objects as elements, etc.

(C) Composition Subroutine:

This subroutine uses < D4 > recursively to "extend" the BOU by identifying successively more extensive configurations of which BOU is an element.

(D) Contingency, or structural, rules:

This component is a list of statements to the effect that certain descriptions (of type N, D2, D3, or D4) or configurations hold only if certain other such descriptions hold.

(E) Paradigm Object Configurations:

Paradigmatic objects serve a function comparable to the PAS paradigmatic processes. One additional feature is that paradigmatic cases for a given ObUnit may be used to map the range of variation of instances of that ObUnit in the A-Space or C-Space.

8.4 The Operative Time System

The OTS is designed to provide a unified collection of histories of objects, organizations, institutions, and perhaps eventually, fields of knowledge. The basic form of representation is that of an object-process configuration. The history of an object consists of a configuration of the historical processes in which it participates or which it undergoes. Unification is provided by a reference configuration, namely a calendar-clock and its repetitive, conveniently standard, time-keeping processes. To bring any other event or configuration into composition with this configuration is to give the former a locus in historical chronology (the C-Scale).

A secondary form of representation (the T-Scale) is a temporal interval representation which is not chronologically anchored. This corresponds to the situation where a given PAS schema is known to have been instantiated but the where and when are unknown. PAS schemas are, in effect, formulas for temporal interval representation. Note that temporal interval representations may be chronologically anchored at any point, not merely at their beginnings. For example, reports that "The game ended at 4:55 P.M." or "The campaign was in full swing during the Spring of 1970" would constitute such anchors.

Thus, the OTS requires no basically new forms of data organization, since it makes use of PAS schemas and PWS configurations. Primarily, the OTS provides the unified reference framework (the calendar-clock) and coordination among instantiations of the process schemas and configurations.

8.5 The Means-Ends System

The Means-Ends System contains several forms of data structures, including alternative forms at different levels of complexity and precision.

(A) The simplest data form is expressed as "N equals the degree to which X is a means to Y." This corresponds to the evaluation of a functor, "the degree to which X is a means to Y," when particular content is specified for the arguments X and Y. Functor information is obtained from relevant informants by means of psychometric procedures described in RADC-TR-67-640.

(B) A more complex functor form is required for data which would be discursively given as "The degree to which the value of the functor F, given argument X, is relevant to the suitability of X as a means to Y." The contrast between this form and the preceding may be clarified by an example. Data in form (A) would involve such statements as "Fuel is to a high degree (8.0 on a 9 point scale) a means to transportation." In contrast, corresponding data in form (B) might be "The octane rating (or combustion temperature, etc.) of a particular fuel is highly (7.6 on a 9 point scale) relevant to its suitability as a means to transportation." Thus, form (B) adds complexity in the form of a contingency which renders the means-ends structure of the fuel-transportation relationship more precise.

(C) There is a still more elaborate form which may be particularly significant for dealing with technical data. This data has the form of a formula which expresses (a) the suitability of X as a means to Y as a function of (b) the value of the functor F which is referred to in (B), above. Here again, an example will be most to the point. With respect to the fuel-transportation relationship and its contingency on the octane rating of the fuel, an example of the new data form would be "Any octane rating over 95 will make the fuel suitable as a means to transportation." Alternatively, we might have a mathematical formula, e.g., "Suitability of fuel = 1.13 (octane rating - 90)".

As indicated in the discussion above of the hierarchy of completeness of representation in the several subsystems, the introduction of greater complexity of means-ends considerations produces results which approach the PAS schemas. In the present case, the relevant PAS schema would be that of getting from place P to place Q by using a gas-powered vehicle after putting in fuel, but with a contingency (a) that the fuel be 95-octane or better or (b) essentially the same formula as above, expressing suitability as a function of octane rating.

The motivation to include the more complex means-ends data is not simply the general one of greater precision of representation. In addition to that, form (C) provides a specific advantage when combined with A-Space data (below) in which the major characteristics of elements (FERS elements) are quantitatively represented. The specific advantage is that the A-Space data provides ingredients from which form (C) statements can generate means-ends judgments which (a) have not been explicitly entered in the system and (b) may be unknown to the human user. Of course, there is the corresponding risk: the wrong formula will produce erroneous judgments. This is why the learning capacity which can be provided by the Inductive System becomes more and more important with increasing complexity and sophistication of the SAS type systems.

(D) Data in form (A) or (B) may also be correlated and factor-analyzed to provide geometric representation (a Means-Ends Space and/or an Ends-Means Space) analogous to the Attribute Space and Classification Space. The advantage here is that every element which is entered at all is related, and in a quantitative way, to all the Ends represented in the system. Since the same element may be entered as both a Means and an End it is possible to use M-E spaces to construct means-ends chains(see RADC-TR-67-640, the M-E feasibility study).

(E) Finally, as with the PAS and PWS, there will be means-ends chains as representational schemas, accompanied by paradigm instances of each schema. These means-ends sequences (note that the preceding data forms may be considered limiting cases of means-ends sequences) serve as representational types in the FERS and OTS and they are used in drawing conclusions via pattern-matching procedures. In these respects they parallel the PAS and PWS schemas.

8.6 The Classification Space

This system contains data in the form of an evaluated functor: "N is the degree to which term X is relevant to subject matter field Y." With data of this sort, subject matter fields can be correlated across terms and the inter-correlations among fields factor analyzed to produce an n-dimensional subject matter space. When such an arrangement is used to index and retrieve lexical material it is designated as a Classification Space. These procedures and their experimental evaluation are described in detail in RADC-TDR-64-287 and RADC-Tr-67-640.

One function of the C-Space is simply to provide document indexing and retrieval for users. In this respect it functions solely in conjunction with the A-Space and independently of the remainder of the SAS.

Within the SAS, the C-Space performs indexing of sentences, paragraphs, documents, etc. In principle, therefore, the C-Space solves the problem of "spoon-feeding" for the SAS. That is to say, the C-Space provides a representation of what subject matter a given message "is about" and therefore in principle, information received by the SAS does not have to be accompanied by an elaborate set of meta-messages, such as topic tags, in order that it be interpretable. In practice, as pointed out, there is always some ambiguity as to what a given message is <u>about</u>, and so the degree to which spoon-feeding is required by a given SAS is best evaluated experimentally.

Since the PAS, PWS, and MES schemas are also indexed in the C-Space, the C-Space indexing of an incoming message provides an interpretation of which of these schemas the message is "about". Thus, from the point of view of drawing conclusions from input via pattern-matching with respect to PAS, PWS, and MES schemas, the C-Space indexing provides an order of priority (based on apparent relevance) for selecting schemas as candidates for pattern-matching. Since the C-Space type of indexing provides an ordering, in terms of degree of relevance, with respect to all the "topics" representable in the C-Space, it provides for pattern-matching searches at any degree of depth up to an examination of every schema which is contained in <u>or</u> can be constructed by the MES, PAS, and PWS. The depth of search is easily controlled by specifying contingencies and absolute limits.

8.7 The Attribute Space

This system contains data of the form of an evaluated functor: "N is the degree to which X is characterized by Y." This data is factor analyzed to give a "conceptual content space" which is analogous to the "subject matter space" of the C-Space. Detailed procedures and experimental evaluation are described in RADC-TR-67-640.

One function of the A-Space is to contribute, with the C-Space, to the automatic indexing and retrieval of intelligence documents for the analyst, independently of the remainder of the SAS.

Likewise, within the SAS, the A-Space indexing operates jointly with C-Space indexing to provide an interpretation of what a message is "about".

The main function which the A-Space does not share with the C-Space stems from the fact that the A-Space provides a summary characterization of the kind of attributes that a thing has. Thus, it provides ingredients for procedures which depend on contingencies involving such attributes. One example was provided by the form (c) data of the means-ends system as described above. A variation is based on the principle that two elements are intersubstitutable in any schema if they have the same attributes and are substitutable in a given schema if they have the same relevant attributes.

8.8 The Deductive System

It is not anticipated that initial versions of a State of Affairs information system will require an elaborate theorem-proving capacity. In large part this reflects the fact that most conclusion-drawing in the SAIS will not be accomplished on a deductive basis. For a minimum capability, the following would suffice:

(A) Implementation of a modus ponens inference schema (if X implies Y, and X; then Y).

(B) Implementation of constraints on either arguments or relations in propositional schemas. That is, constraints of the form"this inferential schema holds only if, in the propositional schema F(X,Y) the value of X = C and/or if G(X,-) holds, and/or the value of Y = K and/or if H(-,Y) and/or if the value of F = f and/or Ø (f)." The constraints on arguments would provide the equivalent of meaning postulates for referring expressions such as "airplane" or "city"; likewise, constraints on relations would do for attributes such as "father of", "larger than". Among the specific relational constraints to be implemented are those dealing with transitivity, symmetry, and reflexivity of relations.

Thus, in principle, the Deductive System is able to take as premises any of the facts represented in the FERS and OTS, so long as these facts involve elements or relationships which satisfy the constraints referred to above.

8.9 The Inductive System

The Inductive System performs three distinct functions. These are (1) to make decisions as to the acceptability and disposition of conflicting information, (2) to make decisions as to the acceptability of a conclusion drawn on the basis of pattern-matching, and (3) to calculate and record success-ratios for specific meaning postulates and pattern-matching criteria. Since initially both (1) and (2) are based on explicitly written decision tables which may later be elaborated by introducing contingencies and calculations, none of the three functions requires more than routine programming.

8.10 Heuristic Program System

Ultimately some kind of heuristic, or translational, system will be required in order to extend the range of input and output formats with which the SAIS is capable of dealing. Two kinds of input goals are obviously relevant. The first is to achieve full capability for analyzing natural language text and paraphrasing it into the propositional calculus format which the SAIS is geared to operate on. The second is to achieve visual and other sensory input and not merely linguistic input. Both natural language analysis and pattern recognition are areas in which substantial activity is currently going on and in which some progress may be expected.

It would seem, however, that if a system cannot be designed which will process intelligence data effectively without these input resources, then it cannot be done with them, either. Likewise, it is difficult to imagine an adequate experimental test of an input system which could be carried out in the absence of a "control" system which had all of the information processing features other than the input system. Finally, there is some reason to suppose that a functioning SAS based on logically formatted input might be of significant value in the design of input and output systems. For example, the information-synthesizing capabilities of the SAIS as described above make it, in effect, an instrument for discourse analysis. Discourse analysis, in turn, might well provide the kind of context which was indispensible in deciding what was being said at a given time (by a given sentence or clause).

For these several reasons, our specification of a prototypical SAIS does not include a Heuristic Program System, though it would be quite possible to do so and eventually it will be essential to do so in the interest of methodological completeness as well as functional efficiency in an operational setting.

9.0 SA SYSTEM IN SURVEILLANCE

As is well known, the work of intelligence analysis does not comprise a single kind of task at the technical level. Even though the activities of the intelligence community may be conceptualized as aspects of a single, organized body of practices for SAS purposes, we must still deal with different kinds of SAS applications which would correspond to different major options within a PAS schema of intelligence activities. Among these major options would be included (a) surveillance, (b) technical capability appraisal, and (c) scientific-technical development appraisal.

The first of these is exemplified by the problem of reconstructing what type and range of activities are being carried out in a given installation. The second is exemplified by the problem of evaluating the likely and/or possible performance of a piece of equipment or an operational system under a variety of conditions. The third is exemplified by the problem of tracing the development of the scientific theory and technology of ionic propulsion and summarizing its present status.

Since exemplars of each type of problem appear significantly in FTD operations, the application of a SAIS to such problems would constitute a direct contribution to the FTD mission.

In this section the problem of reconstructing the range of activities being carried out in a given installation is examined from the viewpoint of SAS implementation. In the analysis of a specific problem such as this, we assume the existence of an information processing system having the general characteristics delineated in Section 7 and Section 8, including the input sequencing of 7 (K), the retrieval capabilities of 7 (L), and the subsystems described in 8.1 to 8.9.

If we designate the installation as LocA, it appears that the most critical distinctive requirement of the problem is a set of PAS schemas representing <u>each</u> of the envisioned candidates for "what goes on at LocA." Only slightly less critical would be MES schemas for each of these activities.

PWS schemas would be required at some points but not as centrally as the PAS and MES types. Most of the candidates would probably be alternate possibilities within one of two general schemas, i.e., (1) a scientific research, publication, and behavior change schema or (2) an industrial production schema. More specific schemas would represent particular types of research (primarily) and publication and behavior changes or more particular types of industrial production.

The function of the PAS, MES and PWS schemas would be to give the SAIS the ability to represent the <u>possibility</u> of each of these being "what goes on at LocA" and of their being alternative possibilities, not necessarily mutually exclusive (mutual exclusivity would be represented by meaning postulates relating the two schemas in question.) The presence of the schemas as FERS attributes or functors or as OTS chronologies would give the SAIS the ability to represent the <u>fact</u> that given schema did or did not identify "what goes on at LocA". Schemas would, of course, also be indexed in the C-Space and A-Space.

The crudest sort of stage-analysis of such PAS schemas could probably be accomplished by a sources-activities - products-use breakdown. Although it is the Activities stage which corresponds primarily to the simple notion of "what goes on at LocA', the other stages are of potentially critical significance in connecting what goes on at LocA to the rest of the real world, and it is those connections which permit us to draw conclusions about what goes on at LocA in the absence of direct observation.

"Sources" refers to the elements which must be assembled in order for the activity to take place. (These are referred to in Section 8.2 as "Elements (roles) in Name A", where Name A is the name of a Basic Process Unit.) These include both persons and material. In general, these elements would be merely identified and only partially characterized by the listing of elements. Further characterization would be given by contingency rules to the effect that certain roles require individuals or materials having certain characteristics. Following the logic of behavior descriptive maxim No. 8 (Table 2, Section 5), these specifications would connect with other PAS or MES schemas which represent ways in which an individual might acquire such characteristics. With respect to persons, probably the most relevant contingencies concern abilities and putative competence. If a plasma physicist works at LocA or visits occasionally, that is one thing; if a microbiologist does so, that is something else. And we may draw conclusions about an individual's putative competence by discovering some facts about his educational history. Likewise we may draw some conclusions about an institution's putative competence by finding out some facts about its financial history. If the funds for activities at LocA come form the Ministry of Defense, that is one thing; if they came from the local municipality or university, that is another.

Conversely, the products which result from the activities at LocA will bear the marks of their histories also. To be sure, we may be more interested in the products than the activities themselves. (We may want achievement descriptions, not performative descriptions of "what goes on at LocA".) A very significant product, especially in the research schema, is the published research report. There are three grounds of significance: (1) Such reports are likely to provide one of the primary lines of access to what goes on at LocA. (2) Such reports may literally describe what goes on at LocA or come close to doing so. (3) The published report provides connections to individuals (e.g., co-authors), institutions (the affiliations of the authors, including LocA), organizations (the publishers of the report), other published reports (in the same publication, in the same subject matter area), and a body of users of the report (cf citation indexing).

In general, the activities carried on at LocA are related in a variety of ways to a variety of objects, processes, events, and states of affairs. To the extent that different activities which might be carried on at LocA would be <u>differentially</u> related in a variety of ways to other objects, processes, events, and states of affairs, to that extent it would be possible to distinguish which activities occur there in the absence of direct observation. To the extent that such relationships can be expressed as specific relationships connecting historical particulars in the OTS format the SAIS may be expected to accomplish the degree of identification of LocA activities that is permitted by the information available.

Procedurally, the SAIS application to surveillance problems at FTD would encounter three major and distinct sorts of tasks, namely (1) data preparation, (2) schema preparation, and (3) system programming. These are discussed in turn below.

(A) Data Preparation

At one level, this is a simple clerical-keypunching task which is important primarily because the amount of data which might be expected to be involved in a surveillance type problem is on the order of 500 abstracts which in their present form are moderately formatted. In an experimental application one might anticipate the preparation of from two to four times this amount of data, since the selection of relevant from non-relevant data is part of the point of experimentation. The average length of an abstract is estimated at 400-500 words.

The problem of data preparation is substantially complicated by the logical formatting required for SAS processing. (This complication, correspondingly, provides substantial motivation to begin development of heuristic programming for translating text into R (a, b, c) statements. One of the features of abstracts or summaries is that they tend to employ elliptical linguistic constructions in the interest of brevity and this compounds the data preparation problem irrespective of whether a heuristic program or a person trained in SAIS formatting is doing the data preparation. The nature of the problem, the potential use of SAIS discourse analysis in the solution of the problem may best be gauged by imagining the following paraphrased in simple factual sentences:

"An approximate method based on the use of logarithmic frequency characteristics has been used for determining forced vibrations in nonlinear gyrostabilizing systems. The method, which consists in subdividing the linear part of a system into a number of standard units for which logarithmic frequency characteristics are well known, makes it possible to reduce the volume of computation considerably. The forced vibrations are assumed to occur at the frequency of external excitation...."

On the other hand, for the type of surveillance problem indicated above it is likely that very often the analysis of the factual content of the abstract will not be at all crucial since the relevant considerations would be the author, journal, and publisher and the A-Space and C-Space indexing of the abstract, which do not require a factual analysis of the content. Secondly, it may well be that the analysis of the factual content of such text as the example above is best treated within the framework of the scientific-technical development appraisal problem.

(B) Schema Preparation

Unlike the data preparation, the schema preparation presents professional level problems. Three kinds of expertise are relevant here. First is a working mastery of the PAS, MES, and PWS formats into which the conceptualization of relevant activities must be put. This expertise provided, initially at least, by the experimenter. Second is substantive familiarity with the activity, i.e., some professional competence with respect to the research or industrial procedures, concepts, and materials involved. Third is the expertise provided by the intelligence analyst who is sensitive to the level of specificity and types of contingency which are likely both to be actually available as information and diagnostically significant with respect to the problem at hand. It is anticipated that an analyst with substantial experience in a given field (a) is likely to be able to provide both the latter two types of competence and (b) is likely to be the most adequate single source of information for schema preparation. Thus, the procedure called for here is a cyclic consultation between the experimenter who prepares the schemas and the analyst who provides the relevant information and comments on the structure of preliminary schematizations.

The number of schemas involved in a surveillance problem is substantially variable. In part it depends on the size and complexity of the LocA in question—how many different things could be going on there. In part it depends on the form of the question concerning LocA. QA would be "What is going on at LocA?" QB would be "Is either S,T, or V going on at LocA?" QC would be "Of S or T, which is most likely to be going on at LocaA?" For a QA formulation of the problem it would not be unreasonable to think in terms of 20-30 PAS schemas or OTS object-process configurations. A rough estimate of the time involved would be 2-4 hours of consultation per schema. For QB or QC formulation the number of schemas would be substantially less.

One of the characteristic features of SAIS schemas is that they become increasingly useful as more schemas are added. Schemas introduced relative to one problem are available for use in any problem and because of their manifold inter-connections at various levels (OTS, PWS, PAS, MES) provide both contrast (what isn't going on at LocA) or additional positive possibilities. Thus, other things being equal, an extensive operating system will be more effective than a small-scale experimental system. In many respects this is a distinct advantage, given the normal tendency to extrapolate optimistically rather than pessimistically from experimental results.

(C) System Programming:

The programming of the SAIS is a major task which has two components. The first is that of programming a content-free SAIS, or what might be called the SAIS format. This programming would be at the level of the descriptions given in Sections 7 and 8, above. Given a programmed SAIS format, the second task would be to introduct the descriptive content which would define the domain of facts which was relevant to the problem at hand. In principle, the second task would not be a matter of programming at all, but rather one of introducing formatted data which would identify FERS elements and attributes, dimensionalize and index the C-Space and A-Space vocabularies, specify different Basic Object Units and Basic Process Units, etc.

In practice, some cyclic procedure and mutual readjustments are to be expected. Indeed, it may well be considerably more efficient to program a SAIS directly for its initial application and then use this as a paradigm case to generalize the program so that the relevant parameters of the paradigm case are controlled by external contingencies which have the status of input data in the generalized program. Certainly, for the present report, at least, it has not seemed feasible to specify the SAS format even at the level of a flow-chart in any greater detail than Section 7 (L). This is because the extreme flexibility of the format, provided in part by its recursive capabilities, leaves open so many complex options that detailed flow-charting would be at least uninformative and more likely positively anti-heuristic. In any case, the system programming, together with schema preparation, would be the major contribution of the experimenter in the construction of a SAIS for experimental application to surveillance problems as part of the FTD mission.

10.0 SA SYSTEMS IN TECHNICAL CAPABILITY APPRAISAL

The characteristic feature of the surveillance problem is that although the desired information is in principle available through direct observation, the direct observations cannot in fact be made and so the facts in question have to be established, insofar as they can be, indirectly by reference to some number of relevant observations. Technical capability appraisal presents a distinctive problem when the desired information is not in principle available through direct observation, at least not in any simple sense, but the relevant observations can be made. A problem of this sort arises for example, in connection with the performance capability of the communication network involved in the aerial interception of unidentified flying objects where these objects are presumed to be aircraft and not inscrutable phenomena.

The communication network involved in the interception problem is an object-process configuration involving a variety of human and non-human objects, the latter primarily instruments and vehicles and stable, building-like structures. One way of appraising the performance capabilities of this configuration is to raise certain questions and answer them. (It is, incidentally, because the issue is one of capability rather than actual performances that direct observation does not simply provide an answer. Capability is, methodologically, an ID characteristic, and specifically, an Ability characteristic, not a form of behavior.) A configuration of this sort is vulnerable to deficiencies in its various elements or distortions in its various processes. Thus, from a practical viewpoint one might ask, by way of appraisal, "How great a deficiency in element X (or process P) can the system endure and still function successfully?"

However, although the question is one which comes naturally to mind, it is not a profitable one, because it cannot be answered in a non-trivial way. The trivial answer is "Well, it all depends on what happens elsewhere in the configuration." But what happens at any given "elsewhere" also "all depends" on what happens elsewhere, just because of the systematic relationships upon which the basic functioning of the system rests.

Thus, we conclude that not only is there not a simple explicit answer, but also there is not a simple contingency for giving a qualified answer. Moreover, it becomes clear that the apparently simple factual question is closely related to the non-simple question of what is the optimal form of that configuration. That is, how serious a deficiency can be tolerated in a given portion of the configuration "all depends" in large part, on the degree to which the entire configuration is optimal. One might even define the optimal configuration as the one which can tolerate the greatest deficiencies in its elements and still function effectively.

Pursuing the interception problem in this way leads us to two preliminary formulations. The first is that the indicated procedural solution involves (a) a representation of the entire configuration, (b) variation of the relevant aspects of the configuration independently or with co-variation constraints, and (c) "answers" having the form of probability distributions across the various possibilities that are left open when any one or more of the relevant aspects are specified; in turn, this very nearly rules out anything except computer simulation as the technical approach, or at least, as the primary approach. Second, a plausible prescription for optimal system functioning is (a) maximal use of communication within the system by elements when it is available, together with (b) maximal independence from the rest of the system (in the sense of minimal reliance on it) when adequate communication is not available.

Examining our preliminary formulations, which were made on metholological grounds, we may see why they might be expected to provide insuperable technical difficulties under existing methods of approach. (a) A representation of the entire configuration is impossible to obtain. Although many important aspects are susceptible of mathematical treatment the entire configuration, and especially "the human element" does not yield to mathematical treatment. (b) "Minimal reliance on the rest of the system" is a paradox. How could it mean anything other than just whatever limited function can be performed by a given element when other elements are deficient? Likewise, what could "maximal use of communication" be except the condition that nothing is wrong with the system?

It is in connection with such difficulties as these that a SAIS appears capable of making a distinctive contribution: (a) A SAIS is capable of representing an object-process configuration, and since that representation is a structure of alternatives and contingencies, the relevant aspects can be varied on a simulation basis. (b) The "maximal use-minimal reliance" prescription is met by giving the major elements a SAIS as part of their operating principles. In particular, each such element has a representation of the entire configuration and the role of each element, including its own, in the configuration. As a result, communications from other elements do not have simple automatic consequences either when they come or when they fail to come or are distorted. Given SAIS capability, each element is capable of "observing" the other elements rather than simply being operated on by them. Given observational capability and a representation of the configuration as mission-oriented, each element can then utilize to the fullest its ability to make use of the contributions of other elements or to compensate for their deficiencies in terms of their joint mission. In the interception problem the element to which the SAIS makes the most significant contribution is the aircraft which attempts the interception.

In effect, the introduction of the SAIS optimizes the object-process configuration by achieving a degree of reflexive closure which transforms it from a machine to a team. Surveillance capability, as described in Section 9 is added to the "direct observation" capability which corresponds to the normal "communication within the system." As we shall see below, the more detailed working out of this SAIS application appears to require something akin to aspects of scientific-technical development appraisal as well.

10.1 The Interception Problem

In this section the interception problem is developed at a level of detail which appears to be heuristic in delineating the specific applicability of SAIS procedures to this example of a technical capability appraisal problem.

(A) Discursive summary:

The following is a summary of the paradigm case interception configuration: There are give primary roles or elements. These are the UFO, a radar station, a control station, an aircraft station, and a pilot. The sequence of events is (1) the UFO is detected by the radar station, (2) the radar station communicates the information to the control station (3) the control station signals to the pilot, (4) the pilot gets the aircraft airborne, (5) the aircraft, under the direction of the control station, moves to a location where it encounters the UFO, and (6) shoots it down, whereupon (7) the control station directs the aircraft back to the base and (8) the pilot lands and disposes of the plane.

Using this schema, we may introduce elaborations, alternatives and contingencies, such as (a) recursive patterns of steps 1, 2, and 5, i.e., sequences of detection, computation and command, and execution, (b) a variety of deficits and distortions in either detection, computation, command or execution, possibly resulting from efforts by the UFO, (c) evasion and confusion tactics by the UFO, (d) structural damage to the aircraft, and (e) introduce additional aircraft and UFO's or have one replace another.

For purposes of reference, the UFO is designated as X, the aircraft as Y, the pilot as W, ground control as V, and the radar ground unit as Z.

(B) Representational Requirements:

(1) A hierarchically organized set of social practices dealing with the interception mission, in the form of PAS schemas and PWS object-process configurations. The highest level in the heirarchy should be superordinate to the interception mission, e.g., the defense of a certain air space. The superordinate schema provides (or provides for) principles for selecting among options in the subordinate schemas (e.g., for selecting alternate means to the same end, for selecting options on the basis of efficiency, etc.)

(2) With special reference to the aircraft, Y, we require a distinctive portion of the OTS chronology, designated here as the "biographical register", which provides a real time representation of what Y is currently engaged in doing. "What Y is currently doing" refers to his current place in the carrying out of the interception, and this in turn corresponds to the role of Y in the current stage of the PAS schema which represents the interception. This is accomplished by formulating the role of Y both as an element in the representational schema and as an executive program which controls the behavior of Y. In this case the execution of successive

phases of the program is also the behavior of accomplishing a given portion of the interception activity, and it is recorded as such in the biographical register.

(3) The following illustrates a paradigmatic interception schema, designated as PAS-1. In the schema "STL" indicates a spatio-temporal location and "informs" designates a sequence of"sends" and "receives". As noted above, X is the UFO; Y is the aircraft; W is the pilot; V is the ground control station; and Z is the ground radar unit.

Stage	Activity
1	Z detects X at STL (1)
2	Z sends "X at STL (1)" to V at T (2)
3	Z receives "X at STL (1)" at T (3)
4	V sends "X alert" to W at T (4)
5	W receives "X alert" from V at T (5)
6	W gets W-Y airborne at T (6)
7a	at T (7a) Z detects X at STL (7a)
7b	at T (7b) Y detects X at STL (7b)
8a	at T (8a) Z detects Y at STL (8a)
8b	at T (8b) Y detects Y at STL (8b)
9	Z informs V that "X at STL (7a)" at T (9)
10	Z informs V that "Y at STL (8a)" at T (10)
11	V computes and at T (11) directs Y toward "intercept point", namely STL (11)
12	Repeat sequence from 7 (a)
13a	V computes D (d(X, Y), i.e., Y is closing with X
13b	Y computes D (d(X, y))
14	Repeat from 7 (a) until
15	Y informs V that "AOK" at T (15)
16	Z informs V that "AOK" at T (16)
17	V computes and directs Y to base
18	W lands and disposes of Y

(4) A representation of the aircraft as a part-whole configuration with particular emphasis on those parts whose relationships are involved in maneuvering the aircraft.

(5) A body of propositions which, in effect, are the theory of aerodynamics and guidance. These, together with the representation of the aircraft, are used in the "effectance" calculations, i.e., in maneuvering the aircraft. These calculations are also used in checking whether the aircraft is functioning normally. Since automatic guidance of aircraft is standard practice, this portion of the system is not analyzed in any detail.

(C) Implementation:

Implementation of the interception mission may for purposes of description be broken down into two parts. The first is simply the gross procedures involved. The second is the way in which the principle of "maximum use-minimal reliance" is implemented. The latter is based on the notion that there are three loci of possible control of the aircraft, Y. These are the ground control station, V, the computer-based guidance system of the aircraft, Y, and the pilot, W. Under ideal conditions these three constitute an order of descending precision and expected effectiveness. The principle of maximum use would therefore prescribe reliance on ground control unless that became ineffective, then in that case reliance on the aircraft's own automatic control unless that became ineffective, and finally, manual control by the pilot as a last resort. Conversely, the three sources of control also represent a descending order of reliance on communication within the system.

Thus, implementation of a maximum use-minimal reliance principle takes the form of specifying a series of three levels of control (levels of reliance on communication) and a decision matrix which determines under what conditions control is transferred from one level to another. The SAIS capability of the aircraft is employed both in guidance operations and in the calculation of input to the decision matrix for determining the locus of control.

In section (D), below, a description is given of the gross procedures involved in guidance. Section (E) contains much the same information in a more computer-program-oriented format which shows the relation of the locus of control decision to the rest of the procedures. Section (F) contains a decision matrix which involves the calculation of six parameters. Each parameter is a state of affairs description of a particular kind, hence the contribution of the SAIS. Section (F) also shows what is required in order to calculate these descriptions.

(D) General Procedure:

Let Y, the aircraft, be the active element in question. References to the SAIS and its subsystems are references to the guidance and control system of Y itself. The "biographical register" referred to above now appears as a functor, "OTSH", in the FERS. This functor takes on as values the PAS schemas and stages which Y is engaged in at a given time. Corresponding to each PAS schema which may appear as a value in the OTSH is a subroutine which, when called into the "Historical Register" operates the guidance and control system of Y (subject to the locus of control decision). As noted above, the correspondence of the HR subroutine and the PAS schemas of the OTSH is such that the execution of the subroutine instructions is the performance of the role of Y in the PAS schema. (Note below that part of the surveillance capability of the SAIS is directed toward the diagnosis of structural-functional pathologies which are indicated when the observed results of executing subroutine instructions do not qualify as the performance of the role of Y in the PAS schema which is then in the OTSH. Such diagnoses are reflected in the locus of control decision matrix.)

We begin by entering the superordinate, PAS schema (e.g., defense of a given air space) in the OTSH and Historical Register, and selecting the Interception Schema (PAS-1) as the initial option selection at the first stage. In this way we bring it about that this mission is what this element is up to now. A sequence of sub-processes will now run off. Which sequences (options in the PAS-1 schema) are run off will depend on what PAS schemas are available and what means-ends alternatives can be generated within the SAIS:

(1) From the superordinate, transfer control to (PAS-1) (Stage 1).

(2) Enter in FERS the fact that Y is in (PAS-1) (Stage 1).

(3) Y refers to FERS for self-orientation. Note that it is the hierarchical structure of the PAS schemas which permits Y to move from one specific procedure to the next, since the completion of a given schema returns control to the next higher level PAS within which it was selected as an option within a stage and which thereupon selects an option within the next stage.

Thus Y enters (PAS-1) (Stage 1)
(Stage 2)
⋮
(Stage F)

This sequence is recorded in the OTSH functor of the FERS and in the OTS.

(4) An "effectance" procedure or sequence is a PAS transition which can

actually be accomplished. i.e., can be accomplished by executing an available subroutine or instruction in the Historical Register. The control and guidance of Y as it moves through Stages 1, 2 ---F of PAS-1 is formulated in the SAIS as the means-ends problem of finding an effectance procedure for each transition. Thus: For a given Stage N, where E_N is the event of completing Stage N, let E_{N+1} be entered as an End in the Means-Ends System and/or a more detailed description (via "decomposition" of Name A or expansion of Name AK). Let the means or process schemas be constrained to those for which an effectance sequence exists or could be composed. Let the execution of the effectance sequence for Stage N + 1 be the transfer to Stage N + 2. If a transition to the next stage cannot be effected, transfer back to the next higher level schema and look for a new alternative.

(5) In terms of PAS-1, the primary problem for which effectance procedures are sought is to decrease the distance between Y and X, which may be expected to engage in evasive or distractive maneuvers. It is in this connection that aerodynamic and guidance theory is employed. Success here will depend on the speed, maneuverability, and course of X as against the time required for the Z-V-Y link or for the SAIS computations of Y. These features can apparently be dealt with by straightforward numerical methods once a simulation framework is achieved.

(E) Calculations and Stage Directions

(1) Set OTSH to PAS-1 superordinate with PAS-1 as the selected option.

(2) Call HR subroutine PAS-1.

(3) Record (2) in FERS and OTS.

(4) Calling subroutine PAS-1 sets N = 1 for Stage N.

(5) PAS-1 is first executed by Y at Stage 6 of PAS-1.

(6) For the PAS-1 cycle 7a-12 (see the paradigmatic interception schema above) calculate STL(Y).

(7) Calculate control level via decision table.

(8) Identify E_n, the completion of a process which constitutes StageN.

(9) Express E_n as a function of E_{n-1}. In the present case E_n is a case of reaching a "go to" point. (The final "go to" point is the point of interception.) The point here is that E_n, rather than being expressed in absolute terms (being at STL (!)) is expressed as the outcome of a change in (or as a function of or a transformation of) Y's present (E_{n-1}) condition. Thus, it answers to the issue of how do you get there <u>from</u> <u>here</u>, hence it defines the relevant effectance sequence. (The "event descriptions" referred to in the description of PAS schemas in Section 8 are a prime resource for expressing E_n as a function of E_{n-1}.)

(10) Go through the following sequence in search of an effectance solution:

(a) PAS-1: Ideally, at some level of analysis, PAS-1 is directly effectable (of the level at which paradigm cases are represented).

(b) (PAS-1) (Stage N): A finer analysis of Stage N into a sequence of substages. That is, if some stage is not directly effectable, or not known to be so, break it down into components and see if each of these can be effected. If they can, then E_N can be effected.

(c) Means-Ends System; Locate E_N as an End and examine available means rank-ordered <u>as</u> means to this end. Examine in this order for effectability. If necessary, go to A-Space for possible means, using the substitutability principle referred to in discussing the A-Space in Section 8. (Note that this procedure may be used in conjunction with the breakdowns in (b) and not merely at the level of (a).

(11) If effectance is established, execute and set n = N = 1 and recycle.

(12) If no effectance, signal to decision matrix calculation and move up to next higher PAS schema for alternatives (a) for completion of mission or (b) what to do if mission cannot be completed.

(F) The Locus of Control Decision Matrix:

The decision matrix is a device for locating control of Y, the aircraft, in one of three places at any given time, i.e., at V, the ground control, at Y, the automatic capability of the aircraft, and at W, the pilot. These are designated as command levels 1, 2, and 3, respectively. V, Y, and W may be expected to make similar decisions, but under ideal conditions, V can operate more effectively than Y and Y more effectively than W. The principle implemented by the decision matrix is that control rests with V unless certain tests indicate that V's control is not effective. Under those conditions, control rests with Y unless certain tests indicate that Y's control is not effective, in which case control moves to W. The tests in question are performed by Y.

The decision matrix is a way of combining the information concerning six parameters, designated as P1 to P6. These are as follows. P1 is a "yes" or "no" answer to the question of whether the control decisions received from the lower level (1 lower than 2 lower than 3), are within a tolerable range of deviation of the control decisions of the next higher level. Since under ideal conditions approximately the same decisions will be made by all three, one of the bases for deciding that ideal conditions do not exist is that, e.g., the decision A_y by V does not lie within the range + or - E from the decision Q_y made by Y. The test question is "A_y E Q_y?" The value of E which defines the tolerance limits would differ for W vs. Y as contrasted with Y vs. V. These values would, of course, be among the relevant aspects of the system which would be varied systematically in the simulation method of technical capability appraisal. The progress check, P2 refers to stage 13 of the paradigmatic interception schema. It refers to the question of whether the distance d between X and Y has shown a decrease D over successive execution of instructions by Y. One of the primary grounds for deciding that a given locus of control is ineffective is that if gets Y no closer to X. The parameter P6 is simply a tally for P1 and P2. With this parameter in the picture we are able to decide to count the results of P1 or P2 as negative only if they are negative, e.g., n times in a row, or K percentage of the time in any N successive checks. The parameters P3, P4, P5 are diagnoses of abnormalities in the aircraft structures, the messages from V to Y and the computations by Y, respectively.

Table 3 shows the decision matrix in which level of control is a function of the values of parameters P1 to P5. In the matrix, "+" indicates success in passing tests P1 or P2 or normality in P3, P4, P5. The use of "θ" in the commentary is to indicate that no reference is being made to a given parameter. For example, "θ+++θ" indicates a pattern defined by + values for P2, P3, and P4 irrespective of the values of P1 and P5.

(g) Comments

The foregoing formulation illustrates the well-known fact that answering an apparently simple technical capability question may require an extensive set of experimental procedures. To add the complexity of SAIS procedures to an already complex task may appear to be simply asking for trouble. However, some technical capability questions are such that a useful answer is worth an exceptional amount of effort.

From the point of view of the SAIS as an implementation of the IA system, the present example illustrates how a reflexive, or self-representational, system may be achieved. This feature gives the SAIS the potential for contributing to the optimization of the communication network as well as to the appraisal of its technical capability.

Table 3. Locus of Control Decision Matrix

Computation by Y

Messages V to Y

Structure of Y

Progress check

P5	P4	P3	P2	P1	AEQ	Control Level
+	+	+	+	+		1
+	+	+	+	-		1 unless P2 is better for 2
+	+	+	-	+		2 unless 2 = 1, then go to 3
+	+	+	-	-		2
+	+	-	+	+		1
+	+	-	+	-		1 unless P2 is better for 2
+	+	-	-	+		2 unless 2 = 1, then go to 3
+	+	-	-	-		2
+	-	+	+	+		1 unless P2 is better for 2
+	-	+	+	-		1 unless P2 is better for 2
+	-	+	-	+		2 unless 2 = 1
+	-	+	-	-		2
+	-	-	+	+		1 unless P2 is better for 2
+	-	-	+	-		1 unless P2 is better for 2
+	-	-	-	+		2 unless 2 = 1
+	-	-	-	-		2
-	+	+	+	+		3
'						'
'						'
'						'
'						'

Table 4. Calculational Requirements

Parameter	Calculation
P1 AEQ	Receive A_y from V
	Parity check
	Calculate D_y
	STL (X) Last recorded location of X
	STL (YX) "Go to" point
	Calculate effectance
	Calculate $Q = D_y \pm E$
P2 Progress Check	Look up STL (X), STL (Y), latest known locations
	Calculate d (X,Y) present distance and record it.
	Look up last distance d'
	Calculate D = d - d' Is D positive or negative?
P3 Structures	Scan FERS N (normal characteristics) column
	Use FERS and Deductive system on observational feedback from effectance results; record results in FERS N column
P4 Messages	Calculate parity or redundancy (NB possibility of requests for repeat of messages) or infer from negative progress checks over time
P5 Computation	Parity check
	Infer from negative progress checks
	Receive special signal from V

11.0 SA SYSTEMS IN SCIENTIFIC-TECHNICAL DEVELOPMENT APPRAISAL

A third kind of task which falls within the scope of intelligence analysis is that of monitoring developments in certain scientific and technical areas in order to be able to summarize the state of the art at any given time and to trace and project the course of development in that area.

The phenomenon involved in this problem is that scientific theories are used in developing a variety of technologies, and technologies are developed which involve the convergence of a variety of scientific theories and subject matters. From the point of view of intelligence analysis, interest centers on a summary of technical developments for appraising current capability and on a summary of scientific development for appraising potential capability and on sequential analyses of both for appraising policies or direction of development.

The information available for reconstructing the phenomenon is in such forms as descriptions or samples of artifacts, descriptions of performances of instruments, vehicles, or systems, descriptions of manufacturing processes, reports of research, presentation of theories or hypotheses or applications of theories, surveys of states of the art in various fields, and so on. Thus, in many respects the problem of SAIS implementation is methodologically the same as the problem of surveillance discussed in Section 9. The technical problem which comes to the fore in scientific-technical appraisal though it need not appear at all in surveillance as such is the representation of theories and their use.

The use of scientific theories in technical development is here considered as a special case of "acting on the concept of X". More specifically, it follows the analysis of verbal behavior given by formulas (5) to (10). Technological developments exemplify "B" in formula (5), i.e., they are among the non-verbal behaviors which constitute acting on the concepts which are identified by the theoretical locutions. Scientific theories exemplify formula (5); descriptions of technology exemplify formula (7) and formula (9); and discussions of scientific theories exemplify formula (10) and formula (7). (These are heuristic and paradigmatic statements, not exceptionless generalizations.)

The most salient methodological problem connected with the representation of the use of a theory is the problem of the open-ended list. There is no way to specify explicitly all of the possible uses of a theory. In this connection we may recall the statement that the IA System was intractable to formalization because we generate an infinity of operations as well as of elements ; and further, the statement that performing operations in this system is simply and literally to engage in intentional action. The infinite set of possible behaviors remains infinite when we 'reduce' it to the set of behaviors which include a particular state of affairs concept (in the present case, the theory) in the value of the K parameter. As is well known, this is not a characteristic that is peculiar to theories. The same difficulty would arise if one tried to specify all the possible uses of a pencil, or a cup or an automobile or a stone. The problem is salient in connection with theories because it is in this connection that we are inclined to push the limits and actively seek to extend their applications.

It may be of some comfort to reflect that if "What is the set of uses to which theory X can be put" is a non-question it follows that nothing that people do or might do requires having that information, so that the non-existence of an answer does not prevent us from doing anything whatever. "The uses to which theory X could be put" loses its nonsensical character if it is thought of not as a description of some possible, though as yet unknown, state of affairs, but rather as a formula, or implicit prescription, for producing descriptions. As such, it is something which must be used by a person ("Every description is someone's description.") The formula is used by a person not in the abstract or in isolation, but in the context of a finite set of possibilites which that person is capable of considering.

Thus, we may, after all, use formula (5) as the paradigm for representing theories and their use. Several particular problems are encountered. The first is that of representing the theory in articulated form. After all, a scientific theory is an articulated system, not a single concept such as "red" or "pencil", and different parts of it may be used on different occasions or in different ways, and disagreements or developments will generally concern parts or aspects rather than a theory in toto. The second, which is more familiar, from Section 9, is the representation of the set of social practices in which the theory is employed in a technological way. Thirdly, of course, there is the problem of representing the relationship between the various practices and the aspects of the theory.

The most convenient beginning point appears to be the representation of technical behavior as a set of social practices via primarily the PAS schemas, supported by the other content subsystems. As usual, the hierarchical structure of the PAS schemas is indispensible. Further, as in the interception problem, it is important to anchor the hierarchical structures at a level of generality that is superordinate to the practices which are of most direct concern. This is because the general characteristic of technical developments is that they are new ways to achieve social practices in which they have a place. Knowing what a thing might be wanted for provides a primary clue as to what it might in fact be being used for.

Beginning with the representation of technological behavior as a set of hierarchically structured social practices leads fairly directly to the other two issues. The practices in question will have a variety of options or alternatives, and the selection of the actual procedures from among these options will be governed by a structure of contingencies. The relation of the theory to technology is that the theory provides a way of identifying states of affairs which serve as the contingencies which govern the selection of technical procedures. (Note that in this respect, scientific research involving theories has the status of being a particular sort of technology--which, of course, it is.)

The representation of the theory in articulated form is accomplished by considering the theory to be a set of verbal formulas ('propositions') which define a domain of possible states of affairs. One virtue of this procedure is that it does not require that the theory be "tight" or "rigorous" as in our ideal models of what a scientific theory ought to be. All that is required is some body of propositions. If they do not "hang together" very well, that will show up in the limited range of their uses.

To represent a theory as a set of formulas which defines a domain of possible states of affairs would in general be forbidding technical task. To do this for an appreciable range and variety of theories would compound the difficulty. It is of considerable strategic importance, therefore, to recall that the formal structure of the SAIS is designed to accomplish just that- the representation of a domain of possible facts. From this standpoint, the representation of theories appears to be entirely feasible, since the primary technical problems are already dealt with by the SAIS structure itself. To the extent that the SAIS can be neatly separated into pure format and content, the format can be used recursively and repeatedly for the representation of theories. In this connection the paradigm cases discussed in connection with PAS,PWS, and MES schemas also make a strategic contribution. In the context of the representation of theories, paradigm cases provide the functional equivalent of "preferred models" of a given theory. One of the principal contributions of paradigm cases is therefore to identify types of technical application and to distinguish significantly different types of application of the same theory. (Recall the use of paradigms in the pattern-matching procedures described earlier.)

In general, therefore, it appears that the SAIS technology makes it possible to represent a theory and its technical application by representing a hypothetical individual who has that theory and acts on it. Conflicting theories may be represented as disagreeing individuals, and the significance of the conflict may be appraised by examining what those hypothetical individuals do differently as a result of their theoretical preferences. Likewise, a field of technology and its state of the art may be represented by a hypothetical individual or set of individuals who know all the relevant theories and are therefore capable of engaging in practices that call upon that set of theories. Another point of technical economy is encountered when we deal with the problem of tracing technical and theoretical developments over time. In this case we have available the OTS chronologies of our hypothetical individuals. Further, for the purpose of interpreting scientific-technical development as an expression of policy which may be extrapolated to the future, the superordinate PAS schemas provide a framework for accomplishing that directly - it is by reference to these schemas that we are able to say what an individual might be up to in pushing certain kinds of technical efforts.

From the point of view of contributing to the FTD mission, the considerations discussed in connection with the surveillance problem are all applicable here. The major difference is that schema preparation now becomes a more complex task and one for which both kinds of expertise (the subject matter competence and the intelligence analysis competence) are indispensible. The problem of data preparation as such is the same in both cases; the volume of data requried for an experimental system would be more variable and difficult to estimate in general for the scientific-technical appraisal application.

12.0 CONCLUSIONS

Perhaps the most succinct summary conclusion to be drawn from all of the foregoing is that we appear to have reached the point where the problems of automatic analysis of intelligence data can be approached straightforwardly as technical problems. This is not to say that the absence of intractable conceptual and methodological problems is guaranteed. Rather (recall behavior descriptive Maxim 1) there does not now seem to be any in-principle impossibilities. The formulation of the IA System spanning behavior, language, persons, and world provides a coherent representation of the phenomenon of intelligence analysis and the tasks involved therein. The procedures described in the later sections illustrate at a reasonable level of detail the general form of a technical implementation of the conceptual formulation of automatic intelligence data analysis. Further illustrations of procedures for dealing with three major classes of intelligence analysis problems provide evidence of the flexibility of the basic SAIS approach and its potential for technical refinement. The interrelatedness of the three tasks of surveillance, technical capability appraisal, and scientific-technical development appraisal give weight and detail to the general thesis that the value of existing SAIS elements increases when new elements are added, and this is only part of the story of the potential for improvement.

On the other hand, problems are no less real for being technical rather than conceptual or methodological. What we may conclude is simply that the present state of affairs warrants the very substantial effort of setting up a SAIS as an experimental system in a "real world" setting. The analysis of the three applications of SAS procedures suggests that even a prototypical SAIS is capable of making a direct contribution to the FTD mission. Previous experimental validation of the C-Space and A-Space indexing and retrieval of documents may be regarded as providing some assurance that this contribution would be a substantial one.

REFERENCES

1. Auerbach Corporation
   Relational Data System Study (RADC-TR-70-180)
   Rome, N.Y.: Air Force Systems Command, 1970
   AD 720 263

2. Ortega y Gasset, Jose
   History as a System
   New York: W.W.Norton & Company, Inc., 1961

3. Ossorio, Peter G.
   Classification Space Analysis (RADC-TDR-64-287)
   Rome, N.Y.: Air Force Systems Command, 1964
   AD 608 034

4. Ossorio, Peter G.
   Dissemination Research (RADC-TR-65-314)
   Rome, N.Y.: Air Force Systems Command, 1965
   AD 625 905

5. Ossorio, Peter G.
   Attribute Space Development and Evaluation (RADC-TR-67-640)
   Rome, N.Y.: Air Force Systems Command, 1968
   AD 827 323

6. Ossorio, Peter G.
   Notes on Behavior Description (LRI Report No. 4B)
   Los Angeles: Linguistic Research Institute, 1969

7. Ossorio, Peter G.
   Meaning and Symbolism (LRI Report No. 15)
   Boulder: Linguistic Research Institute, 1978

8 Wisdom, J.O.
   Logical Construction
   <u>Mind</u>, 1931, Vol. XL