DIVERGENT AND CONVERGENT REASONING IN INTELLIGENCE ANALYSIS: A SELF-RECONCILING EVIDENTIAL DATABASE

Prepared by:

Marvin S. Cohen, Kathryn B. Laskey, Russell R. Vane, III,
James R. McIntyre, and Steven G. Sak

Decision Science Consortium, Inc.
1895 Preston White Drive
Reston, Virginia 22091
(703) 620-0660

Prepared for:

Michael Scott Zimmerman
P.O. Box 1752
Washington, D.C. 20013

Under Contract No. 87F370300

TECHNICAL REPORT 88-14

TABLE OF CONTENTS

LIST OF FIGURES

1.0 INTRODUCTION

1.1 The Problem

It is far easier to diagnose the reasons for an intelligence failure after the fact than it is to prevent one beforehand. Success or failure seems to hinge on analysis--noticing significant data in a background of noise, assessing their reliability, or finding a pattern that fills in gaps and resolves inconsistencies--as often as it does on the collection of data *per se* (cf., Laqueur, 1985; Burrows, 1986). Yet there are no easy prescriptions for these tasks:

- An analyst should avoid "biases" but must also draw effectively upon substantive knowledge of the topic and area. That knowledge (if it is useful) will certainly predispose the analyst toward some hypotheses and away from others.
- Almost any data may mean something other than what they seem, due to deception. Sensitivity to the possibility of deception, however, can lead to disregard of genuine evidence; consequently, excessive skepticism may be as misleading as excessive credulity.
- Involvement with policy makers may, on occasions, lead to interpretative errors--e.g., a "Cassandra" attitude (worst-case) or the opposite, "Pollyanna." Yet isolation from policy makers may lead to irrelevance and/or gaps in coverage.

The answer, it is easy to say, lies in *balance*: between attention to theory and respect for evidence; between extensive substantive knowledge and persistent questioning of the assumptions embedded in it; and finally, between divergent and convergent modes of thought--generating and taking seriously alternative possibilities, even comparing and contrasting alternative models and types of analysis, and yet in the end offering a reasonable (and reasonably definitive) conclusion.

The problem, of course, is <u>how</u> to achieve such balance in practical terms. Few would claim that currently available tools supply all the help that is needed. Specialized techniques (e.g., critical-indicators analysis, throw-weight analysis, "crate-ology") do not address the general problem of combining evidence and analyses of diverse types. Bayesian statistics does that, certainly, but to our mind does not provide a particularly natural representation of an inferential argument; perhaps more importantly Bayesian statistics responds inadequately both to the challenge of stimulating alternative points of view ("divergence") and to the requirement of resolving them in a meaningful fashion ("convergence"). Finally, general-purpose tools (e.g., database systems, spreadsheets, "Notecards"), though useful, have little to offer that bears explicitly on the distinctive problems of inference.

1.2 A New Approach: Overview

The present report describes a system that addresses these problems directly. SED (Self-Reconciling Evidential Database) brings together aspects of two approaches: (1) symbolic techniques for structuring arguments and for the adoption, utilization, and revision of assumptions; and (2) mathematical techniques for combining and propagating the impact of evidence. The result, we hope, is not just a hybrid, but a deeper synthesis: a system that is both compatible with the way analysts would naturally approach a problem and at the same time likely to yield improvements. In brief, SED has the following features:

Qualitative structuring. At the highest level, SED organizes information by Issues, i.e., Topics (e.g., "Krasnoyarsk radar"), Questions about those Topics (e.g., "Is it a violation of the ABM Treaty?"), and ~~its~~ a potential Answers. At the lowest level, SED organizes information by Reports, i.e., concrete pieces of evidence from satellites, informants, open sources, etc. Arguments, which link Reports to Issues, and Issues to other Issues, are the heart of SED. Each Argument supports a particular Position on an Issue. Finally, associated with each Issue is a Conclusion, reflecting the synthesis (and resolution) of one or more Positions in regard to the Issue.

Building arguments. A key feature of SED's approach is the phasing of Argument construction to fit the natural stages of an analyst's reasoning: i.e., a "first-blush" or "normal" reaction to the evidence (which we call a "Core Position") is followed by specification of a set of possible disrupting factors. For example, photographic evidence that there are no significant military bases or other assets near the Krasnoyarsk radar would *normally* suggest that its function is not to support a local ABM defense--*unless* (i) assets are planned, but not yet built, (ii) assets consist of natural resources or some other non-man-made feature, (iii) assets are camouflaged or buried, (iv) the function of existing structures has been concealed, (v) the photo analysis was badly done, etc. Any of these conditions (and no doubt others) could cause the argument based on photographic evidence to go wrong. Typically, these exception conditions are assumed false in the absence of direct evidence one way or the other, until and unless the "normal" interpretation of the evidence runs into trouble (i.e., conflicts with the conclusion of some other line of reasoning.) SED thus focuses attention on an evolving understanding of the qualitative meaning and reliability of evidence, as opposed to cut-and-dried numerical assessments of evidence strength.

Divergent Reasoning. In intelligence analysis, as in any inferential activity, there is sometimes a tendency to overlook potential weaknesses or sources of uncertainty in a favored hypothesis. In fact, experimental data, with experienced intelligence officers performing realistic intelligence tasks, suggest that apparently disconfirming evidence may be disregarded or even construed as supporting an initial hypothesis (Tolcott, et al., 1988). SED counteracts this tendency by focusing attention on the ways in which an argument could go wrong. SED encourages the analyst to suppose hypothetically, for each Core Argument, that the apparently supported Position is *known* to be false, and to ask himself how the obtained evidence could then be interpreted. This exercise continues by supposing in turn that each new interpretation is known to be false, and asking for another. The result is typically a long (and sometimes surprising) series of qualifications on the original argument: e.g., Report R means Position P *unless* qualification-1, qualification-2, etc. The analyst is thus prompted to act as his own Devil's Advocate, exposing hidden assumptions and exploring alternative points of view. At the same time, the net of his analysis is cast wider, to include any data that might bear on any of those assumptions.

Adopting Assumptions. While it is beneficial to make presuppositions *explicit*, it is not possible to do without them. Assumptions of some sort (e.g., about the reliability of a human source, the proper functioning of a sensor, continued accuracy of a dated observation, etc.) are necessary if definitive conclusions are ever to be arrived at. SED permits such assumptions to be adopted and utilized. However, it makes an important distinction (though only a matter of degree) between assumptions and beliefs supported by evidence. Assumptions are constrained by existing knowledge at the time they are made and are subject to retraction when and if they lead to trouble--i.e., when they conflict with new evidence or with lines of reasoning supported by other assumptions.

Conflict Resolution. When two pieces of evidence or lines of reasoning appear to have conflicting implications, standard normative models statistically aggregate the numerical measures of their strength (e.g., by Bayes' Rule, Dempster's Rule, fuzzy logic, etc.). For example, suppose an analyst has (1) the photographic evidence alluded to above (that no significant assets have been seen near Krasnoyarsk); and suppose he assigned a high degree of numerical strength based on this evidence to the position that Krasnoyarsk is not intended for local ABM defense. Now suppose (2) a covert human source, highly placed in the Soviet military hierarchy, reports that Krasnoyarsk *is* being built for purposes of local defense. Given his previous experience with this source, the analyst assigns the same high level of strength based on the new evidence to the conclusion that Krasnoyarsk *is* intended for local defense. In numerical systems, these two pieces of evidence will simply cancel one another out, leaving equal amounts of belief in both possibilities. An analyst, by contrast, is more likely to wonder why two highly regarded sources are telling different stories. He will look for an *explanation* of the conflict and, if he can, try to reduce it. SED supports that process. SED uses conflict as a symptom that something is wrong with one or more assumptions that led to the conflict (e.g., one or more sensors, models, human sources, etc. are not as reliable as supposed), and implements a process of higher-order reasoning that attempts to reduce conflict by reasoning about the assumptions or by collecting further data. Conflict, in short, is an opportunity to *learn* (e.g., are there possible undiscovered assets near Krasnoyarsk? Is there evidence of camouflage? How trustworthy is the informant? How credible are *his* sources? etc.)--not to blindly aggregate. The result may be valuable information for future use, and often, a more definitive picture of the problem at hand.

The Role of Numbers: SED can accomplish ~~each~~ most of its functions non-numerically--organizing evidence and hypotheses into arguments, exposing hidden sources of uncertainty, and distinguishing firm belief from assumption, ~~and~~

supporting well-reasoned resolutions of conflict. Numerical judgments, however, are of use at two different levels: in describing gradations of belief about hypotheses, and in guiding higher-level reasoning *about* those beliefs:

(1) Numerical assessments in SED are *constructed* through a more basic process of qualitative reasoning. The numerical impact of a piece of evidence is arrived at by exploring simple beliefs and assumptions regarding the disrupting factors (these beliefs may themselves be directly assessed, or arrived at through further simple Arguments). SED automatically computes the implications of these judgments in the form of a belief function (Shafer, 1976) or, as a special case, a Bayesian likelihood function. There is an affinity between SED's logical structures and Shafer-Dempster belief functions, since a belief function quantifies the chance that given evidence proves or fails to prove a hypothesis. Belief functions are, therefore, based on underlying (typically implicit) sets of judgments regarding the reliability of the link between evidence and hypothesis (Shafer, 1981b); SED requires that these judgments be made explicit, and in the process clarifies the meaning of a belief function representation and simplifies the required assessments.

(2) SED embeds belief function arguments within a higher-order reasoning process in which assumptions are adopted, evaluated, and revised. Extensions of belief function theory, in turn, provide tools that support these higher-order processes. Since belief functions measure the degree to which evidence fails either to prove *or* to disprove a conclusion, they define an area of ignorance within which the analyst is free to make assumptions; assumptions in SED go beyond evidence, but are constrained by it. In addition, measures can be defined both of the *degree* of conflict among arguments and of the *degree* of culpability of a given assumption for the conflict. As a result, conflict resolution becomes far more flexible and less *ad hoc* than in purely symbolic approaches (e.g., McDermott and Doyle, 1980).

Database Management. SED stores Reports, Arguments, Positions, and Conclusions in a standard relational database (dBase III). As a result, in structuring an analysis, it is not necessary for an analyst to explicitly stipulate linkages among diverse Arguments that bear on the same Question, among Arguments that bear on different Questions regarding the same Topic, or among Arguments that rely on a common Assumption. These connections are automatically established *via* the analyst's description of the Arguments themselves. In principle, the full power of standard relational access languages may be utilized by SED to manipulate and organize inferential structures.

1.3 Overview of the Report

The current SED prototype operates on an IBM PC/AT desktop computer. It utilizes an inference system called the Non-Monotonic Probabilist (Cohen 1986; Cohen, Laskey, and Ulvila, 1987), which combines aspects of both numerical and non-numerical approaches to uncertainty. NMP is implemented in Golden Common LISP by means of the Belief Maintenance System described by Laskey and Lehner (in press). A more extensive discussion of different concepts of uncertainty and a theoretical rationale for SED may be found in a previous report (Cohen, Laskey, and Ulvila, 1987).

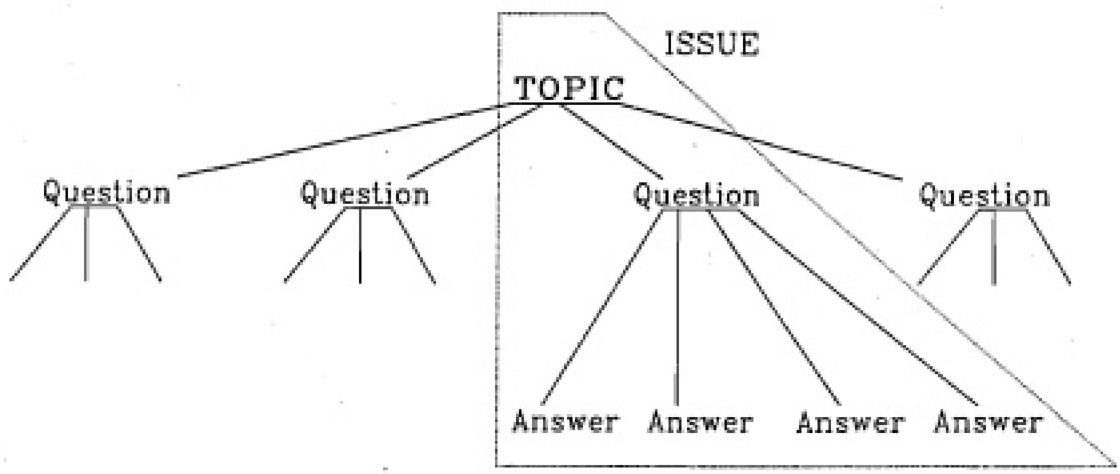
Section 2.0 gives an overview and a short example of how SED works from the point of view of a practicing analyst. Section 3.0 turns to a more detailed discussion of how SED is used and the underlying concepts. The focus of discussion is on how SED interacts with and supports the user's own problem-solving processes: how uncertain beliefs are elicited from users and represented in SED, the respective roles of qualitative and quantitative structures, and how SED deals with higher-order reasoning processes that adopt and revise assumptions. Occasionally we will refer to capabilities not yet fully implemented; but all described functions are operational unless explicitly noted otherwise. A more technical description of SED's reasoning mechanism is given in Appendix A. Appendix B describes the system architecture and relational schemas (implemented in dBase III) which are utilized by SED.

Section 4.0 summarizes SED's capabilities and explores some of the ways SED could be enhanced: (1) by a more powerful graphics interface; (2) in the technical details of its models and algorithms; and (3) by the addition of a permanent, modifiable knowledge base. Such a knowledge base would enable each analyst to construct his own "expert system" over time, to support Argument construction and to institutionalize accumulated analytical experience.

2.0 SED: A SHORT TOUR

2.1 SED's Semantics

An Issue in SED is a Topic, a Question about the Topic, and a set of possible Answers to the Question:



EXAMPLES

ISSUE	TOPIC	QUESTION	ANSWERS
#1	Krasnoyarsk radar	What is its function?	Local defense Early warning Space tracking Other non-ABM
#2	Krasnoyarsk radar	Will the Soviets agree to dismantle it?	Yes No
#3	Soviet supersonic aircraft	How many have been delivered to Latin America?	0 1-10 51-100 100>
#4	Columbian heroin	What will be its foreign exchange value (in current US $) in 5 years?	<$1M $2M - $10M $11M - $20M $20M - $100M $100M>

Often, the goal of an analysis is to arrive at a well-reasoned Conclusion, based on all the available evidence, regarding some Issue. Alternatively, the goal might be to see what Conclusions regarding what Issues are changed by a new item of evidence.

By selecting ISSUES from the main menu, the analyst can review current Conclusions for any Topic and Question in the data base. In the simplest case, a Conclusion is a specific Answer. For example, a possible Conclusion regarding Issue #1 above is:

	TOPIC	QUESTION
	Krasnoyarsk radar	What is its function?

ANSWERS

Local defense	
Early warning	
Space tracking	*
Other Non-ABM	

Support = 1.0

---

i.e., a 100% chance that the evidence shows the radar's function to be space tracking. In other cases, Conclusions may be less precise; e.g.,

---

	TOPIC	QUESTION
	Krasnoyarsk radar	What is its function?

ANSWERS

Local defense	
Early warning	*
Space tracking	*
Other non-ABM	*

Support = 1.0

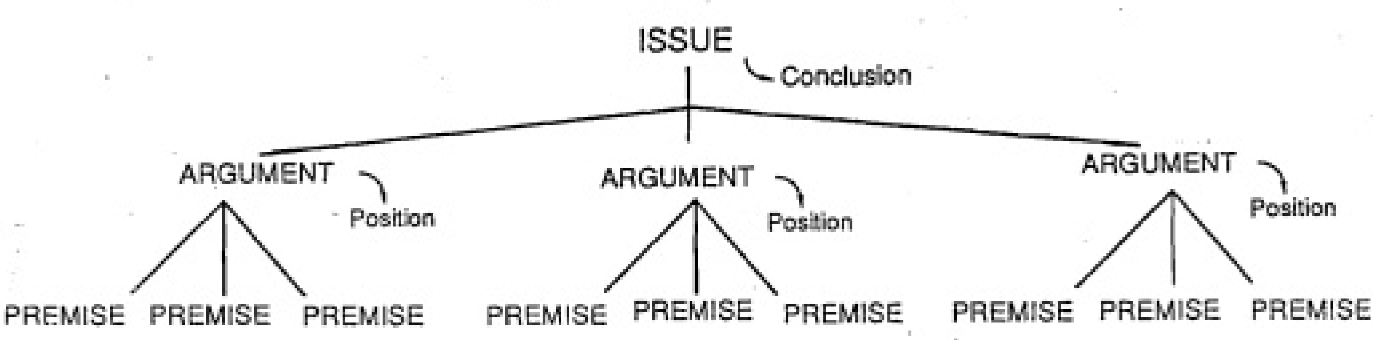
Here, the evidence demonstrates that the radar's function is *not* local defense, i.e., it is either early warning, space tracking, or some other non-ABM purpose; but the available evidence is unable to discriminate further among these possibilities. In still other cases, the Conclusion may involve uncertainty about what the evidence proves: e.g.,

TOPIC	QUESTION
Krasnoyarsk radar	What is its function?

ANSWERS		
Local defense		
Early warning		*
Space tracking	*	*
Other non-ABM		*
Support =	.3	.7

Here, we have a 70% chance that the evidence cannot discriminate among early warning, space tracking, or other non-ABM functions of the radar; but a 30% chance that the available evidence *is* sufficient to show that the function is space tracking. Note, however, that there is 100% (= 70% + 30%) belief that the function is not local defense.

SED helps the analyst arrive at Conclusions by means of Arguments. To build an Argument, the analyst selects ARGUMENTS from the main menu. An Argument is a set of Premises that implies a Position on the Focal Issue of the Argument. A Position has the same form as a Conclusion, except that it represents the implications of a *single* Argument and the evidence underlying it, rather than the entire set of Arguments and evidence. The following diagram shows the relationships among these basic concepts in SED:



In addition to a Conclusion, each Issue may have associated Assumptions and an associated degree of Conflict. Each Argument has a Core Position (the face-value interpretation of the evidence), a Final Position (which factors Assumptions into the Core Position), and a Revised Position (which takes into account possible exception conditions). Each Premise is associated with exception conditions and a description of their impact on the Core/Final Position. Premises are themselves Answers to Issues, and those Issues themselves may be the foci of other Arguments. These features are included in the more detailed conceptual model or "semantic map" of a SED problem in Figure 2-1.

A given Issue may figure as a Premise in multiple Arguments. Indeed, SED imposes no constraints in principle on the inferential connections that may be created among Issues by Arguments; cycles (e.g., smoke → fire → smoke) which cause trouble in other approaches (Pearl, 1986) are automatically treated in an appropriate manner.

2.2 SED's Screens

SED provides five modules, corresponding to main menu commands, each of which offers a different view, or slice, of the conceptual model of a problem:

ISSUES   ARGUMENTS   REVISED ARGUMENT   CONFLICT   REPORTS

ARGUMENTS permits users to build Arguments by specifying a Core Position, a set of Premises, a set of exception conditions, and Assumptions; it merges the Core Position and Assumptions into a Final Position.

REVISED ARGUMENT combines the Core or Final Position with the exception conditions to get a Revised Position.

ISSUES is a top-level summary of Conclusions, Assumptions, and unresolved Conflicts regarding Issues; it can be used to change Assumptions.

CONFLICT is a tool for diagnosing the causes of the Conflict associated with a selected Issue and for changing the Assumptions that prove responsible.

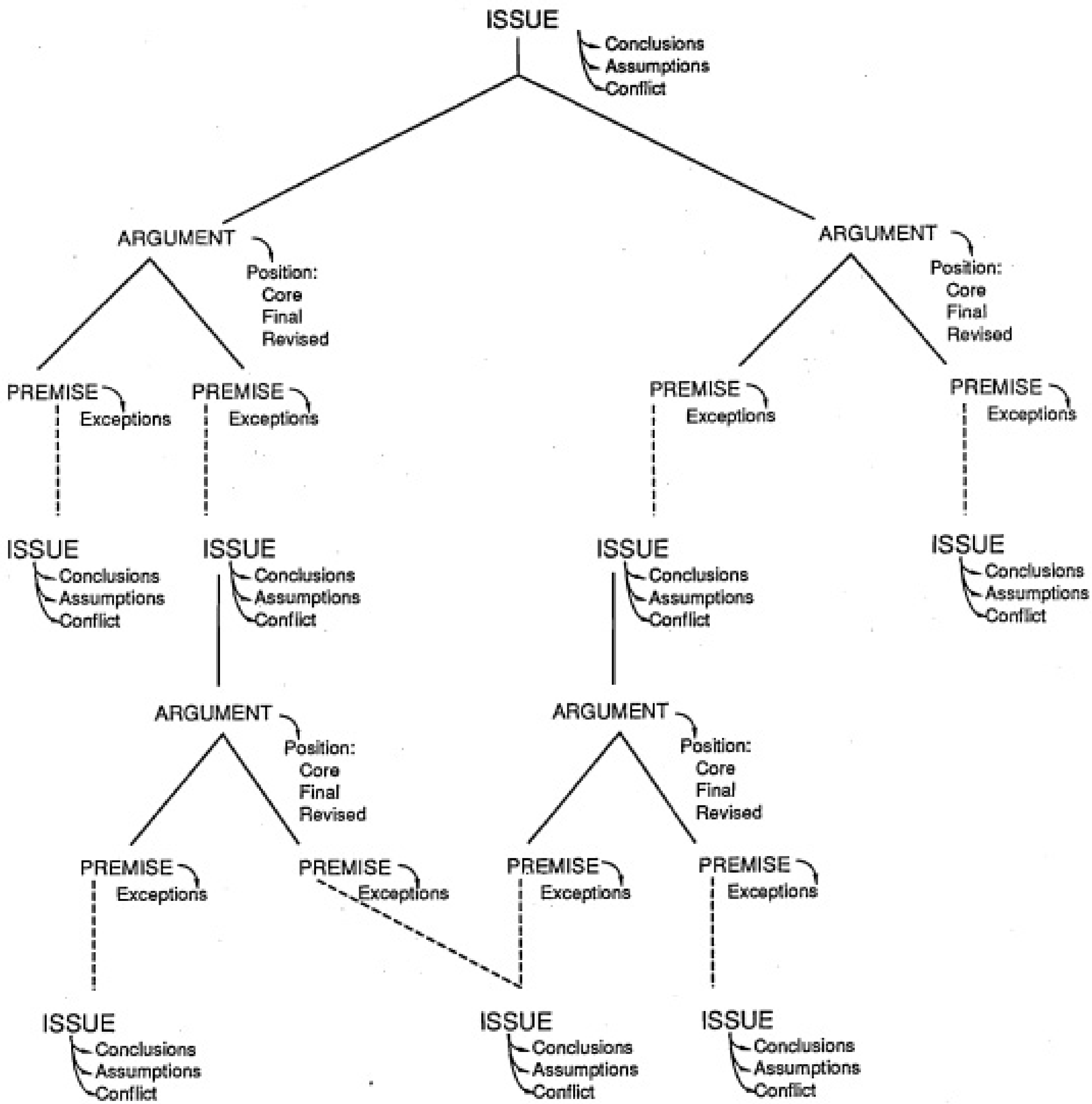


Figure 2-1. A Conceptual Model of SED.

The REPORTS screen (partially implemented) lets the analyst record an input to the analysis and a summary of its content; the occurence of the input becomes a Premise in a new Argument, and a proposition summarizing its content becomes the Core Position.

(In the text we will write the names of screens in capital letters and capitalize the initial letter of some terms like Argument, Position, and Conclusion when they are being used to describe inputs or outputs.)

SED does not utilize a deep hierarchy of operations, in which certain actions can be reached only after a lengthy series of other actions. Any of these screens can be accessed at any time from any other; any of the commands on a screen can be employed at any time; and any part of an inference structure can be worked on at any time. As a result, SED supports a variety of user problem-solving strategies. An analyst might work from the top down (e.g., start with an Issue of concern, construct or view Arguments bearing on it, construct or view Arguments bearing on the Premises of those Arguments, etc.), from the bottom up (e.g., start with a piece of data, construct or view Arguments with it as a Premise, construct or view Arguments with the results of these Arguments as Premises, etc.), or any mix of the two.

SED has two other (partially implemented) commands which, in conjunction with ARGUMENTS, enable the user to navigate freely through a linked network of Arguments:

GROUNDS CONSEQUENCES

CONSEQUENCES enables the user to move up an inference chain to examine Arguments in which the current focal Issue is a Premise. GROUNDS enables the user to move down the chain, to examine Arguments that bear on current Premises. ARGUMENTS supports lateral movement, i.e., the examination of other Arguments bearing on the same Issue. Each of these commands can be used from any screen.

In the remainder of this Section we introduce the use of SED with a simple but complete example, turning in Section 3.0 to a more detailed discussion of the system and its underlying principles.

2.3 Building Arguments with ARGUMENTS

In SED, the analyst is encouraged to state the *reasons* why a given conclusion might (or might not) follow from a particular piece of evidence--not simply a number measuring the degree to which the conclusion is associated with that evidence. Belief regarding an Issue is always determined by one or more Arguments.

The construction of an Argument in SED falls into natural phases:

- Step (1) is an initial face-value or "normal" interpretation of the evidence. It consists merely of specifying the evidence and a Core Position on the focal Issue that seems to follow from it. For example, the first-blush meaning of the satellite photograph showing no assets near Krasnoyarsk is that the function of the radar is not local defense:

```
____[POSITION]______CORE__________________________Argument 1 of 2___

              Krasnoyarsk radar        What is its function?

    Local defense
    Early warning  *                                                  I.
    Space tracking *
    Other non-ABM  *

    Core Support   1.0                                                II.
____[PREMISE]______________________________________________1 of 10___

              Satellite Photo          Are there assets near Krasnoyarsk?   III.
    Yes
    No             *
_____________________CORE____________________________________________
```

The ARGUMENTS screen, as illustrated here, is divided into three parts:

I. The focal Topic, Question, and Answers;

II. Numerical measures for subsets of Answers to the focal issue; and

III. Topic, Question and Answers for a Premise.

The user specifies the Core Position on the focal issue in I by placing asterisks next to the appropriate subset of Answers; indicates the degree of support for these subsets by placing a number (e.g., 1.0) under the appropriate subset in II; and places an asterisk in III next to the Answers that represent the Premise. Topics, Questions and Answers may be entered by the user in I and III or selected from the already existing database of Issues. If necessary, the evidence may be described by more than one Premise.

- Step (2), involves fleshing out the "Core Argument" with a set of background Premises. Background Premises are necessary for the normal linkage between the evidence and the Core Position, even though they may have little or no relevance to the focal Issue taken by themselves. For example, if there were a Soviet plan to build assets near Krasnoyarsk (e.g., a large military base), then the failure to observe current assets would lose its significance. Thus, we have as a Premise in the Core Argument, the proposition that no new assets are planned:

```
____[POSITION]______CORE___________________________Argument 1 of 2___

              Krasnoyarsk radar       What is its function?

        Local defense
        Early warning  *
        Space tracking *
        Other non-ABM  *

        Core Support  1.0
____[PREMISE]____________________________________________2 of 10___

              Krasnoyarsk radar       Are assets planned but not built?
        Yes
        No             *

____________________CORE___________________________________________
```

- Step (3) simply adds an account of what happens to the Position when a background Premise is false:

POSITION	CORE	EXCEPTIONS	Argument 1 of 2

	Krasnoyarsk radar	What is its function?
Local defense		0
Early warning	*	0
Space tracking	*	0
Other non-ABM	*	0

Core Support 1.0

PREMISE	2 of 10

	Krasnoyarsk radar	Are assets planned but not built?
Yes		*
No	*	
	CORE	EXCEPTIONS

The star next to "Yes" corresponds to the exception condition: new assets *are* planned. Above it in the same column, circles represent the impact of that exception on the Position of the Argument. In this case, circles are next to all four possible answers. If new assets were planned, the function of local defense could no longer be excluded, and the evidence could no longer discriminate among any of the hypotheses.

The impact of negating a background Premise may be a less drastic loss of precision. Moreover, background Premises need not be binary (i.e., yes/no) propositions. Thus, more than one exception condition may be specified for a given Premise, each associated with a different impact on the precision of the Argument. For example, suppose an engineering analysis of satellite photographs (e.g., showing that the radar is of type X) suggested that the function of the Krasnoyarsk radar was local defense. Among the Premises of that Argument might be assumptions about the state of Soviet technology and the choices that Soviet engineers would make to solve various problems. In particular, suppose one background Premise is to the effect that type X radar is not used for anything other than local defense:

CONCLUSION CORE EXCEPTIONS Argument 2 of 2

Krasnoyarsk radar		What is its function?		
Local defense	*	0	0	0
Early warning		0	0	0
Space tracking			0	0
Other non-ABM				0

Core Support 1.0

PREMISE 2 of 5

Type X radar		What is it used for?		
Local defense	*	*	*	*
Early warning		*	*	*
Space tracking			*	*
Other non-ABM				*

CORE EXCEPTIONS

If it turns out that radar X is also used for early warning, this Argument will be unable to discriminate between local defense and early warning, but space tracking and other non-ABM purposes will still be ruled out. If radar X is found to be used more widely, there is correspondingly greater dilution of the Argument.

What if one or more Answers to a Premise Issue are ignored by the analyst: i.e., they are neither part of the Premise itself nor part of any specified exception condition? In this case, SED fills the gap by making the neglected Answer or subset of Answers an exception condition and associating it with *total* loss of precision in the Position supported by the Argument. This facilitates rapid Argument construction by the user: the analyst need only specify the Core Position and the evidence (step (1)) and the background Premises (step (2)), and SED is ready to draw appropriate inferences; the user may later return and specify the impact of exception conditions more finely if he chooses (step (3)).

These three steps are the essence of Argument construction in SED. In Sections 3.1, 3.2, and 3.3, we will look at each of them in more depth.

2.4 Getting Results: Support, Assumptions, and Conflict

Assessments (other than 0 and 1.0) are not required in SED to build an Argument (or indeed an entire structure of interconnected Arguments, as in Figure 2-1). Ultimately, however, Issues may obtain varying degrees of Support via the Arguments that bear on them. For this to happen, simple numerical judgment is required only for those Issues that are at the "edge" of the inference network, i.e., Issues which serve as Premises in Arguments, but not as focal Issues for other Premises. The analyst need only provide a number between 0 and 1.0 to indicate where he believes the truth lies: e.g.,

POSITION ______ CORE ______ Argument 1 of 2

	Krasnoyarsk radar	Are assets planned but not built?
Yes		*
No	*	*
Core Support	.3(1)	.7(2)

(1) Assessed by analyst
(2) Supplied automatically by SED

The analyst in this example has assessed a 30% chance that his current knowledge proves there is no planned construction of assets in the vicinity of Krasnoyarsk. If the analyst provides no further inputs, SED will automatically allocate the remaining 70% Support to the set of all Answers, (Yes, No); i.e., there is a 70% chance that the analyst's knowledge is inconclusive on this Issue. This assessment is a very simple "belief function" (Shafer, 1976). A belief function is a measure of evidential support that assigns belief to *subsets* of Answers rather than (as in Bayesian probability theory) to the Answers themselves. As in probability theory, however, the sum of Support for all the subsets must equal 1. The principal advantage of belief functions is the representation of ignorance: assigning Support to subsets with more than one Answer means that the evidence fails to discriminate among those Answers. Support assigned to the subset containing all possible Answers (e.g., in this case, {Yes, No}) signifies the chance that the evidence is completely inconclusive. By contrast, a standard probability approach requires that all the probability somehow be allocated among the specific Answers.

If a Premise had more possible answers, the analyst could allocate different degrees of Support to many different subsets of Answers (so long as the total Support adds to no more than 1.0). The philosophy of SED, however, is to keep direct numerical assessments simple (e.g., assign support to only one subset in addition to the set of all Answers) and to build relatively more complex numerical models on their basis. Direct judgments of this sort are represented in SED by an Argument with no Premises.

Assessments with regard to Premises enable SED to generate the degree of Support implied by Arguments for the Issue of main concern. For example, if there were no other Premises in the Argument based on failure to observe nearby assets, the analyst would get the following Revised Position:

POSITION — REVISED — Argument 1 of 2

Krasnoyarsk radar — What is its function?

Local defense		*
Early warning	*	*
Space tracking	*	*
Other non-ABM	*	*
Support	.3(1)	.7(1)

(1) Supplied automatically by SED

The Core Position of the Argument (that the function of the radar is anything but local defense) is supported to the degree that the Premise (no assets planned) is supported. To the extent that the Premise is false or unknown, the Argument can no longer discriminate local defense from the other possibilities. (More complicated cases, with multiple Premises, are handled by rules discussed in Section 3.4.)

The analyst, however, may feel that this Argument should carry more weight. The 70% Support that remained uncommitted with respect to the Premise defines an area within which he is free to make assumptions. The analyst may allocate all or part of it, by assumption, either to Yes or to No, by specifying a number between 0 and 1.0 for "% Assumed": e.g.,

POSITION ______ CORE ______ Argument 1 of 2

	Krasnoyarsk radar	Are assests planned but not built?
Yes		*
No	*	*
Core Support	.3[1]	.7[2]
% Assumed	1.0[1]	
Final Support	1.0[2]	

(1) Assessed by analyst
(2) Supplied automatically by SED

In this example the analyst assumed no assets were planned. Final Support of 1.0 is equal to the Core Support of .3 plus 100% of .7. SED will now generate a more decisive Revised Position for the Augument:

POSITION ______ REVISED ______ Argument 1 of 2

Krasnoyarsk radar	What is its function?
Local defense	
Early warning	*
Space tracking	*
Other non-ABM	*
Support	1.0

Suppose that the analyst has constructed a second Argument that bears on this same Issue, based on an engineering analysis of the type of radar being used. And suppose the only Premise is the one concerning the likely uses of type X radar. The analyst makes the following judgments with regard to that Premise:

POSITION	CORE	Argument 2 of 2

Type X radar	What is it used for?			
Local defense	*	*	*	*
Early warning		*	*	*
Space tracking			*	*
Other non-ABM				*
Core Support	.5(1)	.3(1)	.2(1)	0
% Assumed	.8(1)			
Final Support	.9(2)	.06(2)	.04(2)	

(1) Assessed by analyst

(2) Automatically supplied by SED

Engineering evidence and past Soviet practice point to use of type X radar for local defense, but permits the possibility of modifying the radar to serve other functions that are technically similar. The analyst is prepared to assume in part, however, that the Soviets have not stretched the technology so as to include the other functions. Thus, he allocates to local defense 80% of the Core Support that includes local defense as a possibility. Final Support for local defense (.9) equals the original support (.5) plus 80% of the support for supersets (.8 x .3 + .8 x .2 + .8 x 0 = .4). The Revised Position for the Argument based on the engineering analysis now becomes:

POSITION	REVISED	Argument 2 of 2

Krasnoyarsk radar	What is its function?		
Local defense	*	*	*
Early warning		*	*
Space tracking			*
Other non-ABM			
Support =	.9	.06	.04

The analyst now has two Arguments regarding the function of the Krasnoyarsk radar: (1) that it is *not* for local defense, since no assets have been observed nearby; and (2) that it *is* for local defense, since radar of type X is used for the purpose of local defense.

	REVISED POSITION			
	Argument #1	Argument #2		
Local defense		*	*	*
Early warning	*		*	*
Space tracking	*			*
Other non-ABM	*			
	1.00	.9	.06	.04

SED combines these Arguments and displays the results on the ISSUES screen:

CONFLICT	TOPIC	QUESTION
.9	Krasnoyarsk radar	What is its function?

CONCLUSION

Local defense		
Early warning	*	*
Space Tracking		*
Other non- ABM		
Support =	.6	.4

The two Arguments taken together provide .6 support for early warning and .4 support for "early warning or space tracking." To arrive at this Conclusion, SED looks at the common ground between the two Arguments. SED finds all combinations of supported subsets from the two Revised Positions, ignores combinations where there are no shared elements, and assigns Support to subsets made up of the shared elements; Support is proportional to the product of the supports from the two Auguments. Thus the first Argument assigns suppport 1.0 to the subset (early warning, space tracking, other non-ABM). The second Argument supports three diffeent subsets, but only two of them overlap with the subset that is supported by the first Argument. Early warning is the only element shared by (local defense early warning, space tracking) from Argument #1 and the subset (local defense, early warning) from Argument #2; its Support is proportional to 1.0 x .06. The subset (local defense, early warning, space tracking) from Argument #2 has two elements in common with Argument #1's supported subset: viz., (early warning, space tracking); its Support is proportional to 1.0 x .04. .06 and .04 are converted to .6 and .4, respectively, by normalization. This Conclusion corresponds to Dempster's Rule. A comparable result would have been obtained by Bayes' Rule.

The analyst, however, may have some cause not to be satisfied with this Conclusion. The reasons are simple: the majority of the Support in each Argument went to subsets that did not overlap at all ("anything but local defense" in the first Argument and local defense in the second); and this dissonance in the evidence was simply ignored. Moreover, the Conclusion reflects strong support for subsets of Answers which the second Argument assigned very little possibility of being true.

SED encourages the analyst to question the results of statistical aggregation. It alerts him to potential problems by displaying a measure of the amount of Conflict associated with each Issue. In this case, Conflict is .9, the product of the support measures for the non-overlapping subsets from the two Arguments (.9 x 1.0). This reflects the chance that something is wrong in at least one of the two Arguments he has constructed.

SED helps the analyst find the causes of the problem by means of the CONFLICTS screen. CONFLICTS searches for Assumptions in the chains of Arguments leading to the Conflict, and prioritizes them in terms of their contribution to the Conflict:

CONFLICT	TOPIC	QUESTION
.9	Krasnoyarsk radar	What is its function?

SOURCES OF CONFLICT

.63	Krasnoyarsk radar	Are assets planned but not built?
.40	Type X radar	What is it used for?

Each Assumption appears with a measure of how much the Conflict would be reduced if that Assumption were retracted. For example, Conflict would become .3 x .9 = .27 if the Assumption that assets are not planned were dropped, a reduction of .63 from the current measure of .9. If the Assumption that radar X is used only for local defense were dropped, Conflict would become 1.0 x .5 = .5, a reduction of .40. (Note that these effects are not additive; dropping *both* Assumptions would still leave .3 x .5 = .15 Conflict.) The contribution of an Assumption to an inconsistency may properly be taken as indirect evidence of its falsity. The analyst may thus reevaluate his Assumption that the Soviets are not planning to build new assets near Krasnoyarsk; if he chooses, he may use the ARGUMENT or the ISSUES screen to retract it. If he does so, the ISSUES screen will display a new Conclusion:

CONFLICT	TOPIC	QUESTION
.27	Krasnoyarsk radar	What is its function?

CONCLUSION

Local defense	*		*		*
Early warning		*	*	*	*
Space tracking				*	*
Other non-ABM					
Support	.86	.03	.06	.02	.04

showing predominant support for local defense.

Finding the Assumptions responsible for a Conflict is trivial in this simple example. In more complex chains of reasoning, however, it is considerably less so. Figure 2-2 diagrams how the CONFLICT screen helps trace the causes of a problem.

Aside from Support, SED (in the REVISED ARGUMENT and ISSUES screens) displays two other useful numerical measures: Belief and Plausibility. For example, if the analyst wants to know how strongly the evidence implies an ABM function of *some* (unspecified) sort, he should look at the tendency of the evidence to support either one of the first two Answers: local defense or early warning. "Belief" in a particular subset of Answers is just the total Support for that subset plus all subsets contained within it. Thus, Belief in ABM function = Support for {local defense} + Support for {early warning} + Support for {local defense, early warning} = .86 + .03 + .06 = .95:

CONFLICT	TOPIC	QUESTION
.27	Krasnoyarsk radar	What is its function?

CONCLUSION

Local defense	*		*		*
Early warning		*	*	*	*
Space tracking				*	*
Other non-ABM					
Support	.86	.03	.06	.02	.04
Belief	.86	.03	.95[1]	.05[2]	1.00[3]
Plausibility	.95	.14	1.00	.14	1.00

(1) .86 + .03 + .06 = .95

(2) .03 + .02 = .05

(3) .86 + .03 + .06 + .02 + .04 = 1.00 (rounding error)

Belief summarizes the positive implications of the evidence for a particular subset of Answers. Plausibility summarizes the extent to which the evidence does not *exclude* a given subset; the Plausibility of a subset is 1 minus

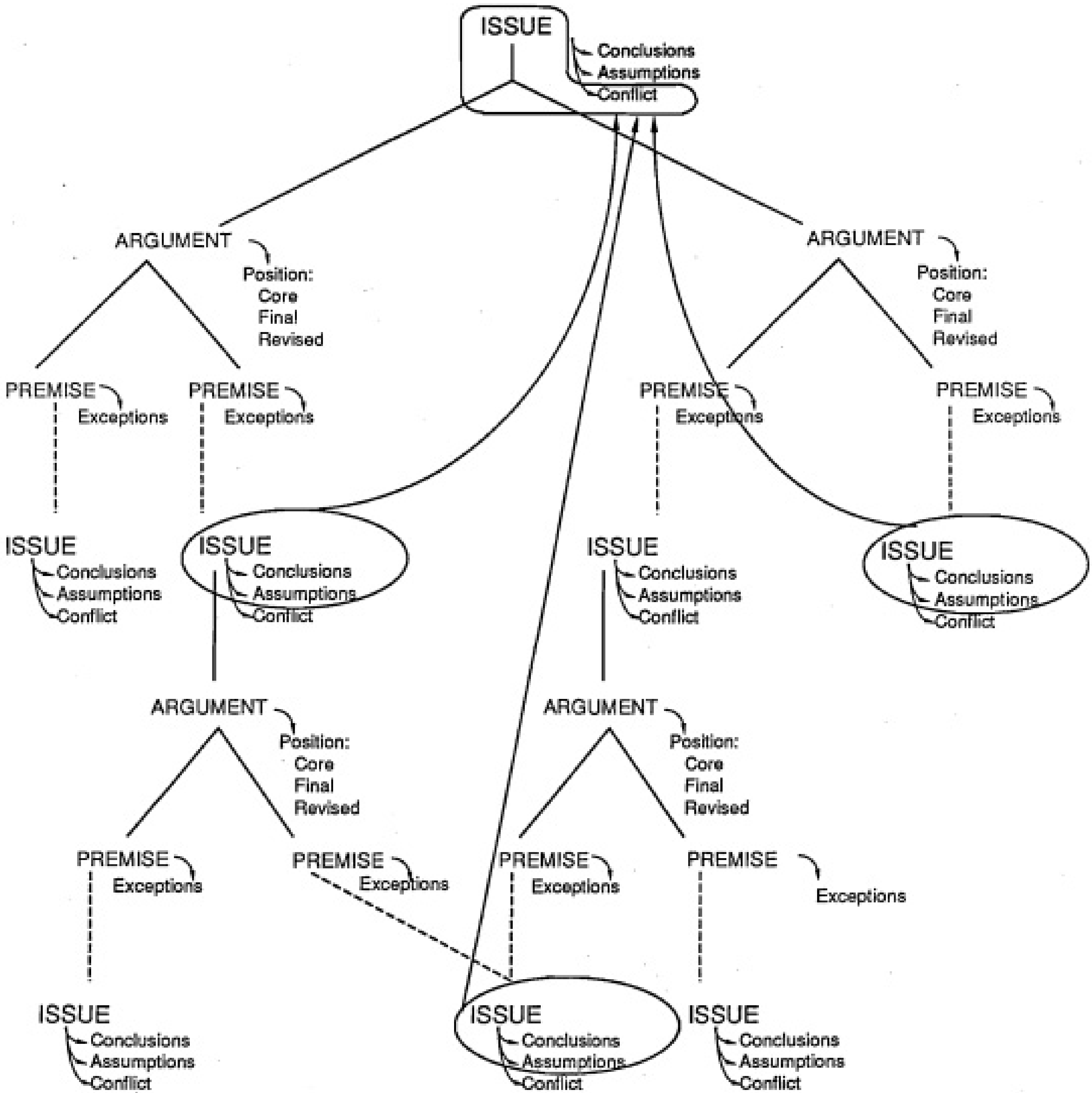


Figure 2-2. Sample View of the Conceptual Model that is Presented by the CONFLICT Screen.

Belief in its complement. For example, the complement of {local defense} is {early warning, space tracking, other}, and Belief in the latter equals Support for {early warning} + Support for {early warning, space tracking} = .03 + .02 = .05. Plausibility of {local defense} is thus 1 - .05 = .95, while Belief in {local defense} is .86. The gap between Belief and Plausibility reflects the failure of the available evidence (and, in this case, Assumptions) either to prove a hypothesis or to disprove it by proving its complement. It thus reflects the *completeness* of that evidence.

In Sections 3.4, 3.5, and 3.6, we will look in more detail at SED's approach to Support, Assumptions, and Conflict, respectively. First, however, we will turn back to the qualitative issues involved in building Arguments.

3.0 USING SED

3.1 Step (1): Origin of the Core Position

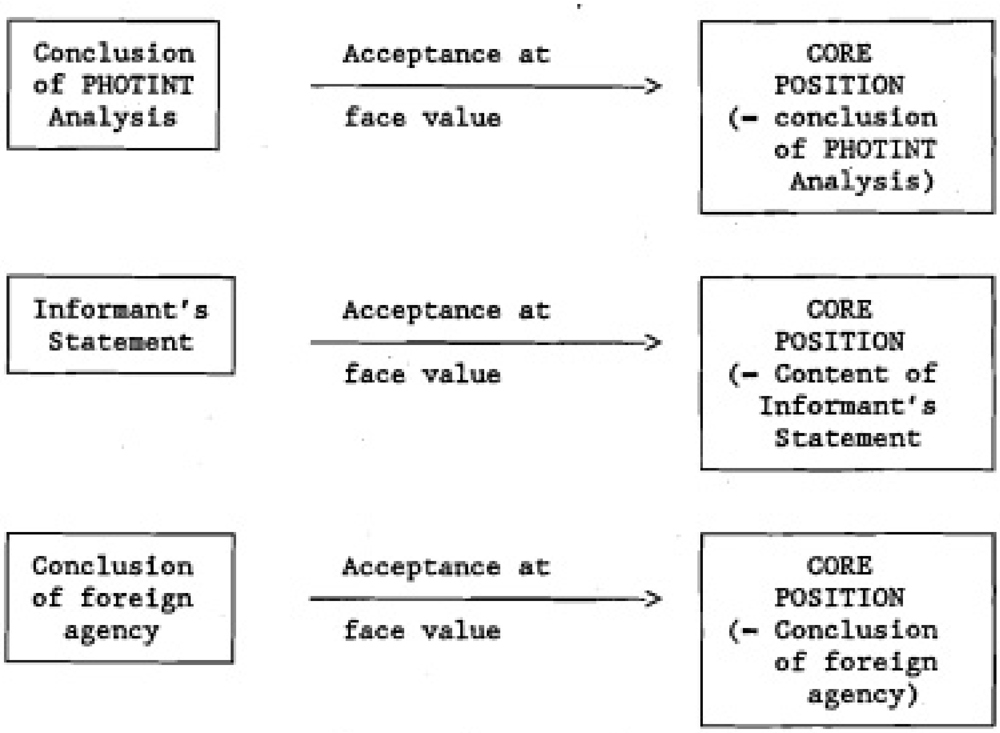
In Argument construction, SED begins by capturing the way an analyst or a model used by the analyst "naturally" reacts to a piece of evidence. Such evidence may include anything at all that serves as an input to the analyst's thought processes: e.g., raw SIGINT data, the results of prior SIGINT analysis, a HUMINT report, a satellite photograph, the results of a PHOTINT analysis, articles from foreign periodicals, raw economic data, the output of an econometric model, results of other Arguments, or even the conclusions of another analyst or agency. In cases where the data have not been previously analyzed, formal or informal models may be used by an analyst to generate the Core Position from the evidence: e.g.,

Raw SIGINT Data	SIGINT Processing ——————> and analysis	CORE POSITION (e.g., identity of radar type)
Raw Economic Data	Econometric ——————> modeling	CORE POSITION (e.g., projection of 1995 GNP of China)
Trucking routes, Bridge tolerances, Weight of trucks	Operations ——————> Research models	CORE POSITION (estimated weight of missile)

Modeling of this sort may take place either outside of SED or within it. In the latter case the analyst can represent the *model* itself in a rule-based format within SED, as a database of Arguments. Each Argument is in effect a rule, with evidence Premises reflecting model inputs or intermediate results. Prior to obtaining any evidence, the evidence Premises of each Argument would have Support = 0. When evidence was obtained, and Support > 0 entered for Issues representing the evidence, appropriate SED Arguments would be automatically triggered; the Conclusions of those Arguments could cause other Arguments to be triggered, and so forth--in exactly the way rules in an expert system are triggered by satisfaction of their antecedent conditions. The analyst can specify Background Premises to reflect ways in which the model could break down at each stage.

Desirable features not implemented in the present version of SED would be: (1) the ability to store generic model templates that could be instantiated for different uses of the same model-type; in these templates, evidence Premises, background Premises, and Core Position could be prespecified for sets of Arguments; (2) the ability to represent the Core Position as a numerical function of Premise Answers (e.g., GNP in year *n* as a function of GNP in years *n* - 5 to *n* - 1); and (3) the ability to handle and combine diverse representations of uncertainty (e.g., Bayesian probabilities for several possible radar types; a 95% confidence interval on projected GNP).

In other cases modeling *per se* takes place outside of SED, and SED may be used to synopsize the results and integrate them with other lines of reasoning. For example, for an analyst at the "all-source" level whose job is to integrate the results of other analyses or models, evidence Premises may reflect the outputs of these models:



In these instances, the Core Position will often (but not always) be a single Answer or a small subset of Answers: i.e., the *most* specific and *least* uncertain Position that follows, on the face of it, from the evidence. The analyst may then associate this Argument with a set of exception conditions that reflect his concerns about the reliability of the prior analysis. The Core Position is the first word, but certainly not the last, regarding the significance of the evidence.

Inputs to an analysis--whether raw data or the outputs of other analyses--may be entered into SED by means of either the ARGUMENTS screen or the REPORTS screen (only partially implemented). For example:

IMPLICATION — 1 of 1

Krasnoyarsk radar — What is its function?

Local defense
Early warning
Space tracking *
Other Non-ABM

Support = 1.0

DATE	SOURCE	NUMBER
5/28/86	Scaramouch	1

Soviet Foreign Minister Scaramouch in a statement in London today denied categorically that the radar installation built near the Soviet city of Krasnoyarsk has any role in defending the Soviet Union against ballistic missile attack. The statement came one day after President Reagan's assertion that the installation represents a "flagrant violation" of the 1972 Anti-Ballistic Missile Treaty. Mr. Scaramouch asserted that the purpose of the installation was to assist in the tracking of orbitting objects in space. He accused the American President of "poisoning the atmosphere" prior to the

In the REPORTS screen, inputs are indexed by Date, Source, and Number, and optionally, by a free text description. In addition, the analyst synops~~o~~izes the input by specifying Positions that it states or implies. In this example, a first-blush reaction to the Soviet spokesman's statement is that it supports X space tracking as the function of the Krasnoyarsk radar; and the analyst has indicated this at the top of the screen. The analyst may specify as many Implications as he likes in order to summarize the inferentially relevant contents of an input. Each Implication is automatically represented by SED as the Core Position of a new Argument; the Premise of each Argument is the same: a statement to the effect that a report on the Specified Topic/Question from the specified Source occurred on the specified Date. Of course, the analyst need not accept the face-value interpretation of this evidence. By selecting the ARGUMENTS screen, he may immediately indicate the exception conditions that reflect his concerns regarding a new Argument (e.g., Is it a deception?).

3.2 Step (2): How to Think Up Background Premises

Typically, the Core Position follows from the evidence only in the context of a large number of background beliefs and assumptions. SED prompts the analyst to make this background explicit. No matter what the basis or form of the Core Position, SED encourages a second (and a third...) look, and encapsulates the results in a set of background Premises. Elicitation of such Premises from the analyst is critical because:

- It exposes uncertainty where it might not at first be acknowledged.

- It identifies *reasons* for the uncertainty (rather than merely quantifying the amount).

- It brings out potentially subtle interconnections among different Arguments. Arguments that depend on the same Premises are appropriately treated by SED as non-independent.

- It points the way to possible additional data collection and analysis, in order to verify assumptions when there is significant Conflict with other Arguments.

- It provides for clearer justification and better understanding of Conclusions.

There are a variety of questions the analyst can ask himself to stimulate generation of background Premises. The simplest is: "Under what conditions would this Argument be valid?" or "What else must be the case for this Position to actually follow from this evidence?" A more powerful method for generating additional Premises is a technique that we call Conflict Resolution (Cohen, 1989; IPL/AMRD, 1982). The analyst forces himself to assume that the Core Position is *not* true and asks himself how that could be. An effective trick is to imagine that he has an infallible crystal ball that says the Core Position is false even though the evidence is true. Typically, the analyst will then be able to generate an explanation: e.g., the Core Position could be false even if the evidence is true, if Q_1 is the case. He now has a new Premise for the Argument, not-Q_1. The analyst consults the crystal ball once more; it tells him that the Core Position is false and the evidence is true, but Q_1 is also false! As a result, the analyst devises a new explanation, Q_2. Again, the crystal ball tells him Q_2 is false; and so on.

Experiments with this technique (in interviews with Army intelligence officers) show that it produces a rich harvest of unexpected information. It was not unusual, for example, to obtain numerous additional Premises by means of the "crystal ball" after more direct questioning of an analyst had run completely dry (Figure 3-1 gives an example). In one instance, after assessing the probability of a conclusion as 1.0, an analyst was able (by means of the crystal ball) to generate 8 different exception conditions with an average assessed probability of .31.

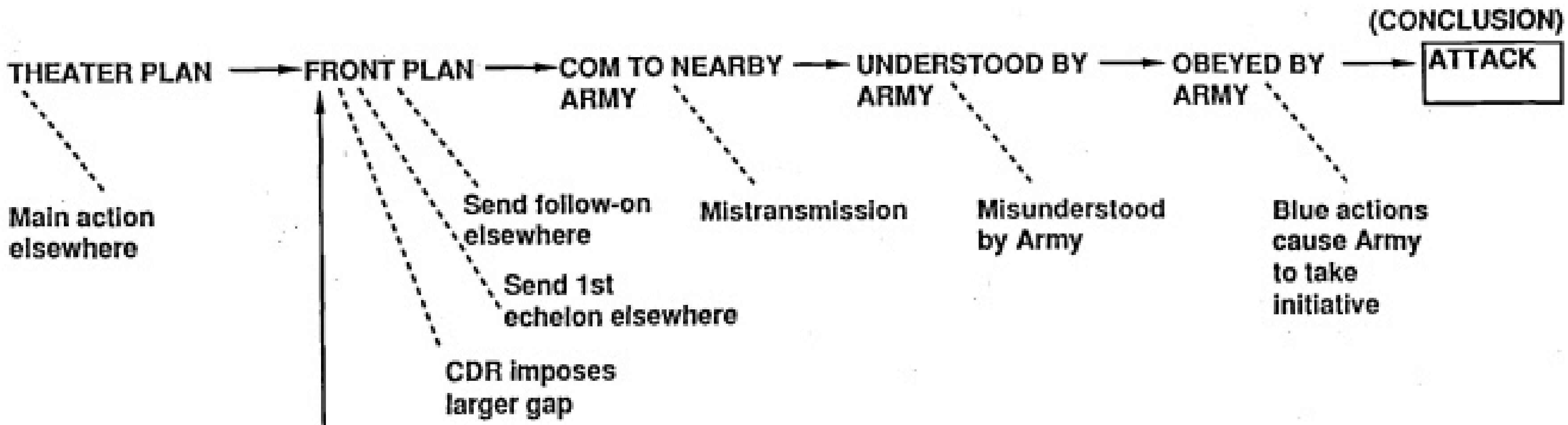
In generating background Premises, analysts must rely less and less on automatic reactions or rules of thumb and more on fundamental domain knowledge. They must become increasingly detailed in their examination of the causal or analytical processes that link evidence and conclusion, if they are to continue adding to the list of exception conditions in which those processes break down. A further stimulus in this process, therefore, is for the analyst to make explicit (perhaps in graphical form) the causal or analytical models underlying an Argument. Figure 3-2, which was developed during an interview with an Army analyst, illustrates a simple causal model underlying the rule in Figure 3-1. Charts such as this have at least two benefits:

IF FOLLOW-ON ARMY DISTANCE >72 HOURS

	<48 hrs	48-72 hrs	72> hrs	No attack
THEN ATTACK			*	
UNLESS				
Front CDR has misestimated distance of Armies	0	0	0	
Follow-on Army is intended for another sector	0	0	0	0
First-echelon Army is to be shifted to other sector			0	0
Front CDR decides to attack without follow-on support	0	0	0	
Theater plans main effort elsewhere	0	0	0	0
Mistransmission to Army	0	0	0	
Misunderstanding by Army	0	0	0	

Figure 3-1. Estimating Enemy Attack.

This example depicts an attempt to estimate the timing of an enemy attack in a Corps sector based on the distance of an enemy follow-on Army. The "first-blush" meaning of the observation that the Army is more than 72 hours away is that the attack is also more than 72 hours away. But there are a variety of exceptions that lead to less precise conclusions. For example, if the enemy Front Commander has received an erroneous estimate of the distance of the follow-on Army, the attack could come at any time.



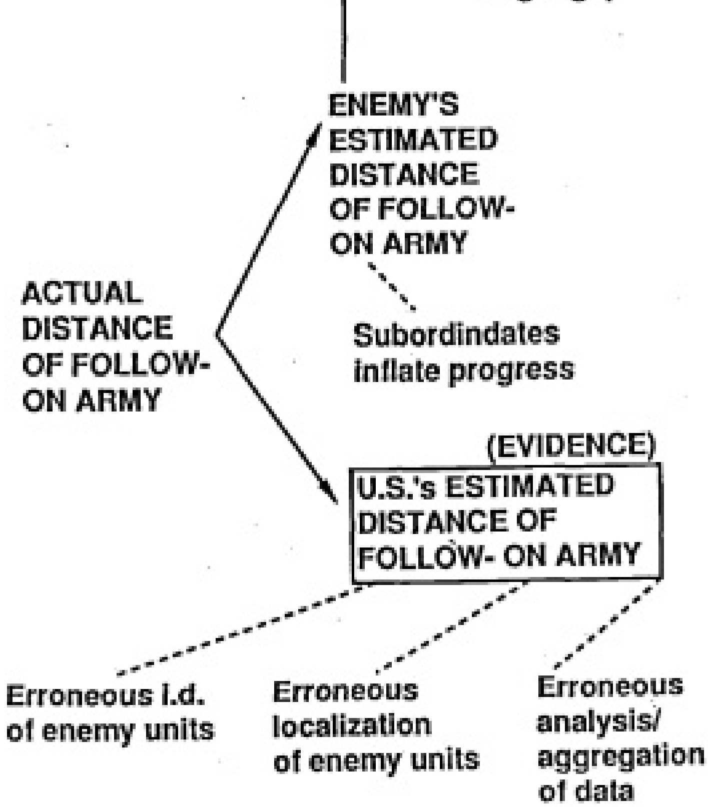


Figure 3-2. Causal Chain Linking Evidence and Conclusion.

The evidence is the U.S. unit's estimate of the distance of the enemy follow-on Army; the conclusion is the enemy time of attack. The causal link between these two is what permits inferences about the latter on the basis of the former. The causal link involves working back from the U.S. estimate to its cause, i.e., the actual distance of the Army, then forward to the enemy's estimate (which is also an effect of the actual distance). The enemy's estimate in turn influences the Front's plan, which is also influenced by the theater plan. The Front Plan must be communicated to the Army, understood by the Army, and obeyed by the Army, if the attack is to occur as predicted. Dotted lines and smaller type indicate alternative paths that can disrupt the inference at every stage.

- As noted, they stimulate generation of more background Premises. The crystal ball technique can be applied in turn to each *stage* of the causal process (e.g., the crystal ball says the problem is at this stage). In the same way, the crystal ball technique can be applied to each analytical step in any model that has been used to derive a Core Position.

- In the next section, we will discuss how causal diagrams can reduce the assessment burden for a set of interacting background Premises.

- Once a chart of this sort has been developed for one Argument, it can be used with appropriate modifications for other, related Arguments. For example, Figure 3-2 is in fact a generic causal model that underlies inferences regarding time of attack when the observed distance of the follow-on Army is 24 hours, 48 hours, 72 hours, etc. Moreover, a variant of the same model underlies inferences about time of attack based on the location of the Front Commander or the location of specialized units.

Note that the analyst could make each link in Figure 3-2 a separate Argument in a chain leading from evidence (U.S. estimate of distance between Armies) to ultimate conclusion (time of attack); or he can use it to generate a single Argument if the intermediate Issues (e.g., the content of the communication from Front to Army) are of no intrinsic interest.

Although the present version of SED does not provide automated support for it, the process of explicitly modeling the evidence-conclusion connection, and attaching appropriate exception conditions to each link, is a helpful adjunct to use of SED (and a promising possibility for automation in later versions). A more ambitious possibility is to let SED keep a store of generic schemas that could be instantiated and combined in particular problems. For example, Figure 3-2 may be thought of as a combination of instantiations of several highly general schemas: e.g., (1) for estimating a quantity (in this case, distance between the Armies is estimated twice, by us and by the enemy), (2) for planning in a hierarchical organization (plan → communicate → understand → obey → execute), and (3) a more domain-specific schema for enemy tactical spacing of Armies. Prestored schemas of this sort could serve as tools to help the analyst in the construction of new Arguments.

3.3 Step (3): Revising the Core Position

In a SED Argument, the Core Position is true if all the Premises are true. But what Position on the focal Issue is supported if one or more of the Premises are false? On the face of it, this would seem to place an inordinate assessment burden on the analyst. For an Argument with n binary Premises, there are on the order of 2^n combinations of truth and falsity of the Premises, for each of which a Position on the focal Issue would have to be specified. In a Bayesian model that conditions one variable on multiple other variables, a probability must be assessed for every value of the first variable conditional on every possible *combination* of values of the other variables (e.g., Pearl, 1986).

One way to reduce the assessment burden, of course, is to reduce the number of variables that are included. Indeed, the exponential growth in required assessments is perhaps a major reason why most approaches to inference do not actively encourage, as SED does, the process of making background variables or Premises explicit. As a result, however, the reasons for uncertainty are less well understood, and issues that may become crucial at a later point (e.g., to resolve conflict) are simply averaged out of the analysis.

A more promising approach is to look for a structure that insulates some variables from the influence of other variables, through conditional independence. For example, in the causal structure of Figure 3-2, once we know what higher-level command is obeyed (or not obeyed) by the nearby Army, the probability of attack is not influenced directly by the variables earlier in the causal chain; e.g., the impact of decisions at the Front is felt only *via* the actions of the nearby Army. Some problems here are: (i) significant economy is not guaranteed (e.g., a fairly large number of background Premises may pertain to the same causal stage); (ii) the analysis is complicated by the need to specify appropriate intermediate variables, which themselves may have no intrinsic interest; and (iii) the time required to develop structures of this sort may not always be available.

SED makes adding background Premises virtually painless, even in the absence of structures like Figure 3-2. It does so by exploiting the idea that negating a Premise has only *one* impact on a given Argument: reducing its precision; and that such impact can often be regarded as independent of the impact of negating other Premises. Thus, for each background Premise, the only requirment is to specify which Answers to the focal Issue could no longer be discriminated from one another if the Premise were false. This is done simply and qualitatively by placing *0's* next to the appropriate subset of Answers. This subset must contain at least one element that is also contained in the Core Position.

SED takes these assessments, together with the Core Position, and automatically calculates the Position supported by each combination of truth and falsity of the Premises. To do so, it simply takes the union of the Core Position with the subsets of Answers associated with the false Premises. For example, consider the Argument in Figure 3-1, and suppose two Premises were false: i.e., (1) the enemy Front Commander has misestimated the distance of the follow-on Army and (2) the first-echelon Army is to be shifted to another sector. The impact of falsifying the first of these by itself is:

```
____[POSITION]________CORE___EXCEPTIONS________________________________

              Attack                    What time will it occur?

          < 48 hours           0
          48 - 72 hours        0
          72 > hours     *     0
          No attack

               Support   1.0

____[PREMISE]___________________________________________________2 of 8____

              Front CDR                 Has he misestimated distance?

          Yes                  *
          No             *

________________________CORE___EXCEPTIONS________________________________
```

In other words, if the enemy has incorrect information on distances, we cannot use distances to discriminate different times of attack (although an attack itself is still expected). On the other hand, the impact of negating the second Premise by itself is:

POSITION CORE EXCEPTIONS

	Attack	What time will it occur?
< 48 hours		
48 - 72 hours		
72 > hours	*	0
No attack		0

Support 1.0

PREMISE

	1st Echelon Army	Will it be shifted to another sector?
Yes		*
No	*	

CORE EXCEPTIONS

If the first-echelon Army is to be shifted to another sector, we can no longer take the presence of the follow-on Army as an indicator of attack (but if there is an attack, it will come after 72 hours).

If *both* premises were known to be false, the Revised Position for this Argument would be:

POSITION REVISED

	Attack	What time will it occur?
< 48 hours	*	
48 - 72 hours	*	
72 > hours	*	
No attack	*	

Support 1.0

Since stars are next to all four possible Answers, evidence regarding the follow-on Army's location no longer tells us anything at all.

We have been supposing that the analyst always believes that a Premise is either true or false. But what if he believes that the truth lies in a subset of Answers that does not exactly match either the Premise itself or any of the specified exception conditions? For example:

POSITION | CORE | EXCEPTIONS

Attack — What time will it occur?

	CORE	EXCEPTIONS
< 48 hours		0
48 - 72 hours		0
72 > hours	*	0
No attack		
Support	1.0	

PREMISE

Front CDR — Has he misestimated distance?

	CORE	EXCEPTIONS	
Yes		*	*
No	*		*
Support			1.0

CORE | EXCEPTIONS

In this case, the analyst has no knowledge at all whether the Front Commander has misestimated the distance between the two Armies; so support for the subset containing both possibilities {Yes, No} is 1.0. Yet the Premise is {No}, and the exception condition is {Yes}.

SED handles such cases straightforwardly. The supported subset is, from the logical point of view, a disjunction: Yes or No. It follows from this disjunction that either the consequenes of Yes are true or the consequences of

No are true--viz., the union of the impact of Yes and the impact of No on the focal Issue. SED thus looks at each Answer in the supported subset (in this case, Yes and No) to see what exception condition(s) and/or Premise condition it is part of. The impact on precision by the supported subset is the combined impact of all the conditions it overlaps with. In the example, the supported subset {Yes, No} overlaps both with the exception condition {Yes} and with the Premise {No}; since the impact of the former is a superset of the latter, the impact of the supported subset is the same as if the Premise were simply false.

When there are n premises in an Argument, SED requires only n + 1 assessments: the Core Position plus an exception condition for each Premise. However, there are situations where more assessments will be desirable. The negation of a Premise, in addition to decreasing the precision of a given Argument, may provide positive support for a different Position on the same Issue. The analyst can express this by creating a new Argument, with a different Core Position on the same Issue (and, usually, a somewhat different set of background Premises). For example, evidence that the Theater level is planning a large-scale effort elsewhere might have the duel effects of (1) diluting the Argument based on the distance of the two Armies, and (2) providing positive support for the Position, No attack, *unless* a diversionary action is also planned, etc.

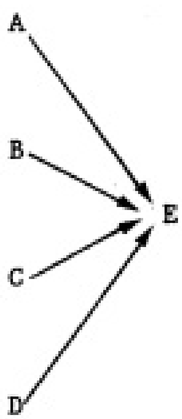
If all combinations of Answers to Premise Topic/Questions had specific significance, the analyst could use SED to create 2^n Arguments. But that is the worst case in SED, whereas it is the only case in traditional conditioning models, such as influence diagrams (Shachter, 1986), Bayesian causal nets (Pearl, 1986), and Bayesian hierarchical inference (Schum, 1980). The key difference is in the basic units of analysis. If a Topic/Question is a *variable*, the basic atom of analysis in SED is the relationship between specific *values* of variables: i.e., a concrete scenario or sequence of events. SED thus permits the user to focus on how an Answer (or subset of Answers) to one Topic/Question is related to the Answer (or subset of Answers) to another Topic/Question. It is this feature which (besides its psychological naturalness) enables SED to incorporate a large number of background Premises economically. Thus, suppose we have the following two Arguments:

			E_1	E_2
If A_1	then		*	
	unless	B_2	0	0
		C_2	0	0

			E_1	E_2
If A_2	then			*
	unless	B_2	0	0
		D_2	0	0

Two things should be noted: (1) the variable D is relevant to the inference of E_2 from A_2 but irrelevant to the inference of E_1 from A_1, while C is relevant to the inference of E_1 from A_1 but irrelevant to the inference of E_2 from A_2; and (2) the effect of B_2 and C_2 in the first Argument (and of B_2 and D_2 in the second Argument) is to disrupt the inference in specified ways independently of one another. These two features are, we think, quite common to evidential arguments that incorporate background variables; SED's representation scheme is tailored to exploit both of them: (1) the first Argument requires no mention of D, and the second Argument requires no mention of C; (2) all combinations of values of A, B, and C or A, B, and D need not be considered. Note in addition that SED automatically creates the logically equivalent Arguments in the reverse direction: e.g., if $\neg E_1$ then $\neg A_1$ unless B_2 and C_2; if $\neg E_2$ then $\neg A_2$, unless B_2 and D_2. Thus, the full inferential relationship among the five variables (A, B, C, D, and E) may be captured in this example by means of eight assessments: the Core Position for each Argument plus an exception condition for each Premise.

By contrast, the atom of analysis in traditional conditioning models is the relationship among *variables*, not *values* of variables. As a result, the *simplest* representation of the above example, in which four mutually independent variables have an impact on E, is:



				E_1	E_2
A_1	B_1	C_1	D_1	1.0	0
A_1	B_1	C_1	D_2	1.0	0
A_1	B_1	C_2	D_1	.5	.5
A_1	B_1	C_2	D_2	.5	.5
A_1	B_2	C_1	D_1	.5	.5
A_1	B_2	C_1	D_2	.5	.5
A_1	B_2	C_2	D_1	.5	.5
A_1	B_2	C_2	D_2	.5	.5
A_2	B_1	C_1	D_1	0	1.0
A_2	B_1	C_1	D_2	.5	.5
A_2	B_1	C_2	D_1	0	1.0
A_2	B_1	C_2	D_2	.5	.5
A_2	B_2	C_1	D_1	.5	.5
A_2	B_2	C_1	D_2	.5	.5
A_2	B_2	C_2	D_1	.5	.5
A_2	B_2	C_2	D_2	.5	.5

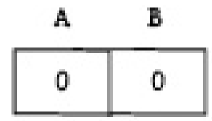
A probability assessment would be required for E given all combinations of values of A, B, C, and D. If all variables are binary, a total of $2^4 = 16$ assessments are required for a Bayesian model. Introduction of *one* new background variable would increase this to $2^5 = 32$, while it would add only a maximum of two assessments in SED for a total of $8 + 2 = 10$.

It is worth noting another implication of the difference in units of analysis between SED and traditional approaches. In a Bayesian model the relationship between two variables may be assessed in either of two directions: from cause to effect (e.g., P(smoke|fire), P(smoke|¬fire) or from effect to cause (e.g., P(fire|smoke), P(fire|¬smoke)). If the appropriate prior probabilities are also assessed, inferences may then proceed in either direction. But all assessments concerning the two variables must be made initially in the same direction. Notice, however, that while P(smoke|fire) is a natural causal judgment, P(smoke|¬fire) is not: we may find it hard to think of the *absence* of fire as the cause of anything. Because of its focus on Arguments that relate *values* of variables, SED provides more options. Some users might prefer to assess one Argument causally (e.g., fire → smoke) and another "diagnostically" (e.g., smoke → fire). Others might prefer to assess both Arguments causally (fire → smoke, ¬fire → ¬smoke) or both diagnostically (smoke → fire, ¬smoke → ¬fire).

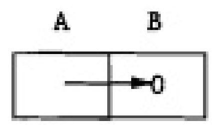
SED is attuned to particular features of background Premises that make it possible in many cases to economically represent large amounts of knowledge, and to do so in a psychologically natural and flexible way. In the *worst* case, if all combinations of Premises implied a different Position and thus needed to be assessed separately, SED could do *no worse* than traditional models. In most cases, it will do much better.

A possible extension. In the present implementation of SED, all exception conditions act directly on the Core Position, and their only effect is to reduce its precision. As we have seen, the result is often an enormous reduction in required assessments. Such economy is not realized, however, when the impact of an exception condition is more complex, or when the impact of one exception condition depends on the possible application of another. Under those circumstance, the present system requires the analyst to construct a separate Argument for each case.

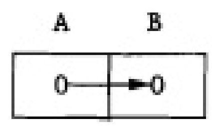
A rather simple generalization of the present approach would preserve the linear relationship of assessments to Premises in these cases as well. We can: (1) allow an exception condition to operate on the results of applying previous exception conditions in a temporal sequence; and (2) specify the impact of an exception condition more generally: instead of a subset of Answers within which discrimination can no longer take place, we can use a rule that substitutes one Answer or subset of Answers for another. Thus, as in the present system, we might have:



indicating that Answers A and B cannot be discriminated; whether the Core Position is A or B, the Revised Position thus becomes {A, B}. In a more general version, however, we could also have:



indicating that if A is in the Core Position, it is replaced by B; or



indicating that if A is in the Core Position, it is replaced by {A, B}.

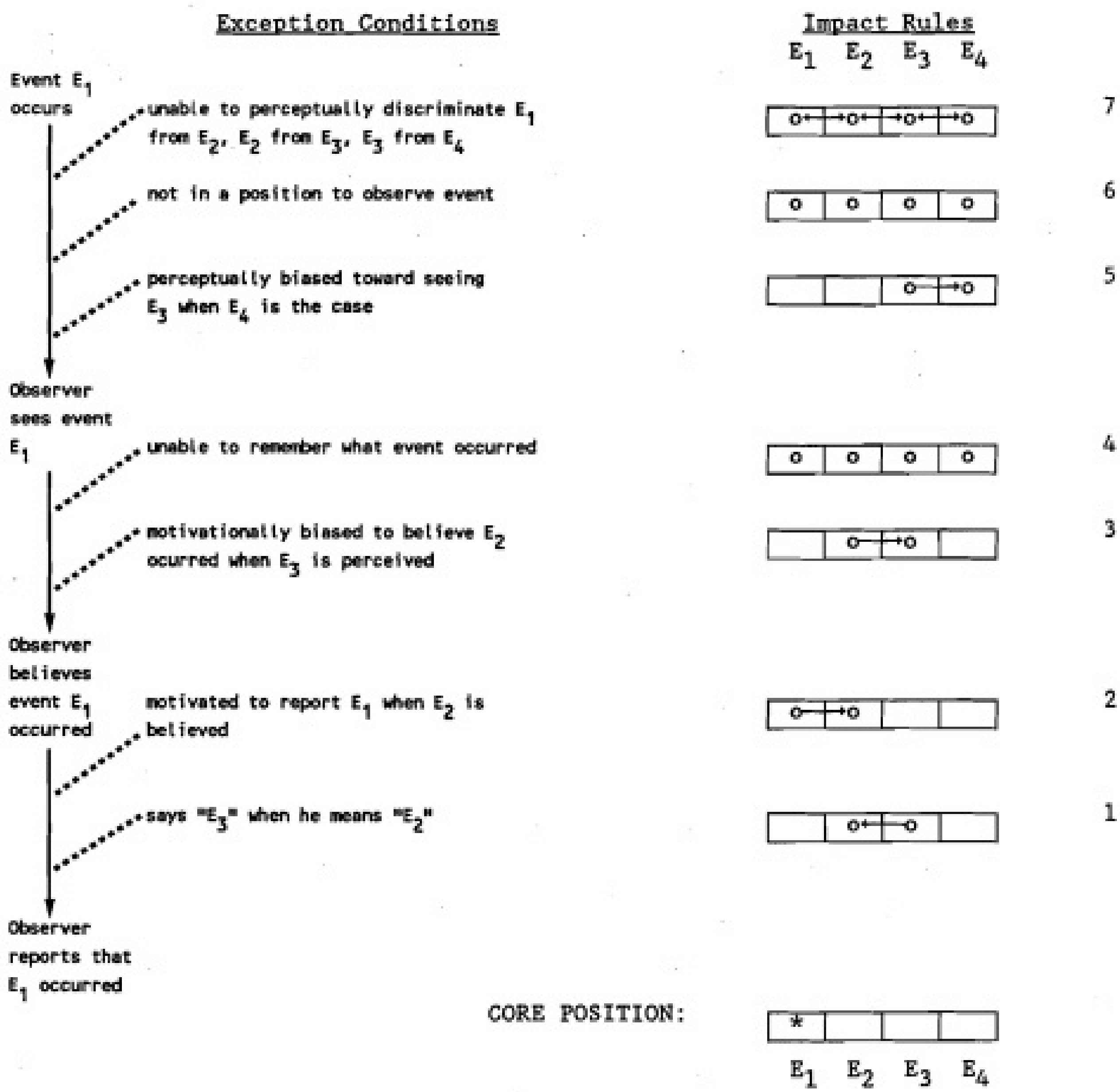
These extensions provide a very economical tool for representing certain quite general evidential Arguments. For example, a standard sequence of events involved in learning about an event or situation from a human source is the following (cf., Schum, 1989):

Event E_1 occurs → Perception by observer of event E_1 → Belief by observer that E_1 occured → Overt report by observer that E_1 occured

Examples include reports about the location and identity of enemy units in Figure 3-2, reports by covert sources regarding policies and decisions taken by foreign governments (discussed in Cohen, Schum, Freeling, and Chinnis, 1984), or reports by inspectors concerning the diversion of nuclear materials from processing plants (discussed in Cohen, Laskey, and Ulvila, 1987). A very similar sequence might occur when the event or situation is detected by technical or automated methods: e.g.,

Event E_1 occurs → Registration of E_1 by sensors → Analyzer model concludes that E_1 occured → Display or transmission of indicators that E_1 occurred

Arguments based on reports of either kind are subject to exception conditions at each stage in the sequence. For example (returning to the human case and borrowing somewhat from Schum's jurisprudence examples):



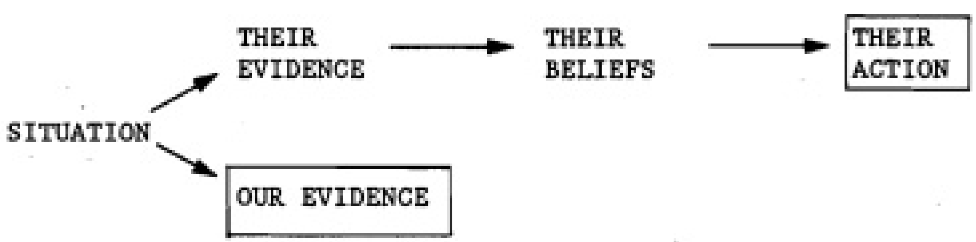
The Core Position, based on the report of E_1, is that E_1 occurred. But the observer may have misspoken (1) or be lying (2); he may honestly believe that he saw something different from what he actually saw, because of what he wishes had happened (3) or because he doesn't remember accurately (4); he may have misperceived the event due to perceptual biases (5), poor observational conditions (6), or limited perceptual capacities (7).

As noted by Schum (1989), each stage (perception, belief, testimony) is subject to both confusion and bias. Note also that the nature and direction of these errors can be different at different stages (e.g., the observer wishes to believe E_2 is true (3), but he wishes others to believe E_1 is the case (2)).

The interaction of exception conditions in examples of this sort can be effectively represented simply by ordering them in a temporal sequence. Suppose that exception condition (3) is in fact the case, i.e., the observer is biased to believe E_2 when his perceptual system's response is E_3. If no other exception conditions are the case, the Revised Position will be the same as the Core Position, i.e., E_1, since the observer's bias favoring belief in E_2 over E_3 is irrelevant. Now suppose that exception conditions (2) and (3) both apply (but no others): the observer is not only motivated to believe E_2 over E_3 (condition 3), but is motivated to claim E_1 is true if he believes E_2 is true (condition 2). Because of condition (2), the witness's testimony of E_1 could as easily mean that he believes E_2 as that he believes E_1; the (interim) Revised Position at this point is $\{E_1, E_2\}$. But condition (3) has become relevant because of condition (2): if in fact the observer lied and really believes E_2, condition (~~2~~ 3) says that belief in E_2 could be due to his perceiving E_2 *or* to his bias to believe E_2 when his perceptual response is really E_3. So the Revised Position after application of condition (~~2~~ 3) to $\{E_1, E_2\}$ is $\{E_1, E_2, E_3\}$. This is the set of "ground truth" situations implied by the evidence (the report of E_1) plus exception conditions (2) and (3).

The Revised Position corresponding to any other combination of exception conditions could be found in the same way: by working backward along the causal sequence from the evidence to the ground truth situation. The process starts with what has been reported (= the Core Position) and asks at each step how it could have been generated from the previous step in the causal sequence under the given set of exception conditions: e.g., (1) What beliefs could have led to the report? (2) What perceptions could have led to *those* beliefs? and (3) What true situations could have led to *those* perceptions? More generally, the process starts with Result = the Core Position and transforms Result at each step according to the appropriate exception condition rule at that step. When multiple exception conditions are temporally unordered (at the same step), Result becomes the union of their impacts and the previous Result. Result at the end of the sequence is the Revised Position for that combination of conditions.

The same method can be applied even more generally, to any causal structure of the sort depicted in Figure 3-2. As before, the process works from the evidence (where Result = Core Position) to the state of affairs that is the focus of the inference; but the direction may be backward in time (as in the reporting example above), forward in time (e.g., predicting what someone will do or say based on an observed state of affairs), or a combination of both. The latter occurs when evidence and combination are linked by virtue of being causally related to a third event. For example, we often predict what someone else will do or say based on our own *inference* of what the relevent aspects of the situation are and a presumption that they will act in their own self-interest:



Structuring an Argument in this way, with appropriate background Premises at each stage, may provide a safeguard against the danger of "mirror-imaging": their evidence may not be identical to ours; their conclusions about the situation may differ from ours; and their decisions about action may not match what we would do in such circumstances.

It can be speculated that any valid example of knowledge involves a causal connection of some sort between one's beliefs or evidence and the conclusions one wishes to draw (cf., Shope, 1983; Nozick, 1981). SED (in this extended version) exploits such causal connections in a very pragmatic way--to provide an economical representation of rather complex, evidential Arguments. n Premises can still be accomodated by n + 1 assessments, if we add a specification of their causal order.

3.4 Constructing Beliefs

SED enables an analyst to spend most of his time thinking in a qualitative and deterministic fashion: Step (1) What is the normal meaning of this evidence? Step (2) Under what conditions does the normal meaning hold? Step (3) What does it mean if each Premise is false?. In its non-numerical mode, i.e., when support is always 0 or 1, SED can serve an analyst as a source of insight into the structure of a problem. It may provide the final form of an analysis when the available evidence (or the willingness to make assumptions) is sufficient to warrant all-or-nothing conclusions. More often, however, the truth or falsity of Premises is neither known with certainty nor completely unknown; and the Arguments constructed on their basis are partially inconclusive. Thus, the analyst may wish to use SED to assess degrees of support for the Premises of an Argument. As a result, a single Argument may simultaneously support multiple Answers or subsets of Answers to various degrees.

Numerical measures may be added quite directly to SED's basic Argument structure. A natural choice for that purpose are Shafer-Dempster belief functions (Shafer, 1976). The reason is that there is a strong complementarity between SED's qualitative inference structure and the underlying semantics of belief functions. Belief functions make sense when we think of them as quantifying the chance that evidence proves a Position; that chance depends on considerations about the reliability of evidence that SED represents as background Premises. SED thus supports the explicit *construction* of belief functions from simpler and clearer judgments (in the spirit of Shafer, 1981a).

Suppose an observer testifies that he saw enemy artillery in a certain location. An analyst could, if he wishes, assess a belief function directly based on this evidence: e.g.,

POSITION CORE 1 of 1

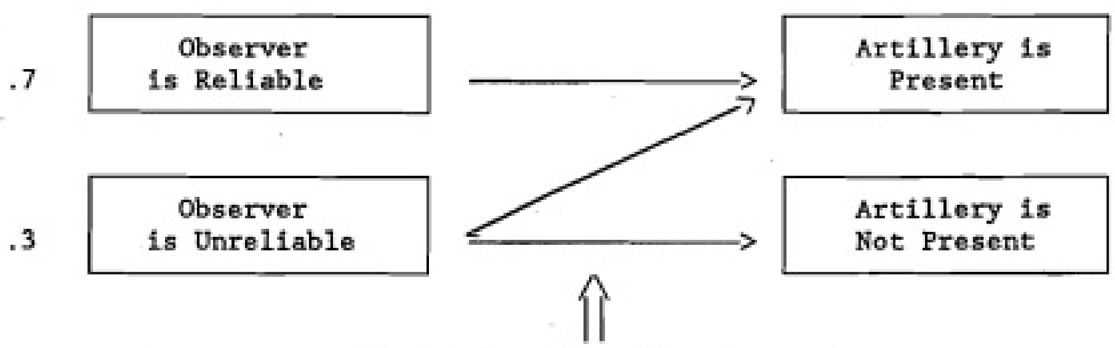
	Artillery	Is it present?
Yes	*	*
No		*
Support	.3	.7

This reflects a 30% chance that the observer's testimony proves artillery is present, and a 70% chance that it proves nothing at all. According to Shafer (in press), these numbers can be understood by reference to an implicit *background* set of hypotheses that is concerned with the reliability of the observer. In other words, the direct assessment above can be construed as resting on an implicit assessment of this sort:

POSITION CORE 1 of 1

	Observer	Is he reliable?
Yes	*	
No		*
Support	.3	.7

This Argument involves a standard probability distribution, i.e., an assignment of numbers adding to 1.0 to individual Answers. Moreover, there is a direct mapping from these Answers to Answers or subsets of Answers to the Question about the artillery:



Mapping based on the observer's testimony that artillery is present

These two features (a probability distribution on the background hypotheses and a one-to-many mapping to the hypotheses of interest) are all that is required conceptually to build a belief function. Reliability of the observer maps onto {artillery is present}; unreliability of the observer maps onto the set {artillery is present, artillery is not present}. The measure of support for a subset A of Answers regarding the artillery is just the probability for hypotheses about the observer that map onto A. (We have referred to this, somewhat loosely, as the probability that the evidence "means" or "proves" A; see Laskey, 1987; Cohen, Watson, and Barrett, 1985). Thus, in our example, Support{artillery is present} = .3; Support({present, not present}) = .7.

SED makes the reliance of belief functions on underlying hypotheses *explicit*. The mapping in the above diagram corresponds exactly to the representation of exception conditions by the ARGUMENTS screen in SED:

```
_____| POSITION |______CORE__________EXCEPTIONS__________________1 of 1____

              Artillery           Is it present?

      Yes          *         0
      No                     0

          Support   1.0
_____| PREMISE |_________________________________________________1 of 1____

              Observer            Is he reliable?

      Yes          *
      No                     *

          Support   .3        .7
_______________________CORE__________EXCEPTIONS___________________________
```

In SED, of course, the reliability of the observer can itself be assessed by a belief function that is not a probability distribution; e.g., support of .3 might be assigned to {observer reliable, observer unreliable}. In that case, SED would automatically construct a probability distribution that it could manipulate internally (with regard to the reliability of the evidence about the observer's reliability). But the user need not be concerned about this, since as we saw in the previous Section, the effect of Support for (observer reliable, ~~witness~~ observer unreliable) on the Question about the artillery is exactly the same as the effect of Support for (observer unreliable). The important points are: (1) SED's computational use of standard probabilities corresponds to the conceptual basis of belief functions, and (2) SED permits the construction of quite complex belief functions from simple assessments. X

To see how this works in a more interesting example, let us return to the Argument for time of attack >72 hours based on the distance of the follow-on Army. The background hypotheses in this case consist of all combinations of truth and falsity of the Premises listed in Figure 3-1. In SED the analyst need not explicitly specify the mapping from each of these combinations to a subset of Answers about time of attack. As we saw in the last Section, the analyst merely specifies a mapping for each false Premise, and SED computes the mapping for combinations.

Suppose, for example, that there is .4 Support for the Premise that the Front Commander has correctly estimated the distance of the follow-on Army; there is .3 Support for the Premise that the first-echelon Army is not to be shifted to another sector; and all other Premises are true. SED uses these numerical assessments to compute Support for different combinations of Answers to the background Issues, and uses the mapping rules laid out in the last section to calculate the impact of each of these combinations on the Core Position. The result is the following Revised Position:

POSITION — REVISED — 1 of 1

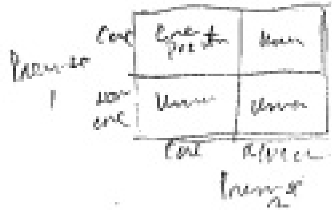
Attack		What time will it occur?	
< 48 hours			*	*
48 - 72 hours			*	*
72 > hours	*	*	*	*
No attack		*		*
Support	.12(1)	.28(2)	.18(3)	.42(4)

(1) Both premises true: .4 x.3

(2) 1st premise true, 2nd premise false or unknown: .4 x .7. Cannot rule out no attack.

(3) 1st premise false or unknown, 2nd premise true: .6 x .3. Cannot discriminate times of attack.

(4) Both premises false or unknown: .6 x .7. Can neither discriminate times of attack nor rule out no attack.



AND – both required for position

This belief function represents a 12% chance that the evidence regarding the distance of the Armies points to the Core Position (attack after 72 hours), 28% chance that it points to no attack or attack after 72 hours, 18% chance it points only to attack at some time, and 42% chance that it tells us nothing. It would be difficult, if not impossible, to make such assessments directly. In SED, a complex numerical assignment of belief, across subsets of Answers to the focal Issue, can be derived from a small number of simple and largely qualitative assessments.

Thus far, we have discussed special tools for the construction of a belief function Argument in SED. We turn now to the combination of different Arguments. According to Shafer (in press), a combination of Arguments can also be understood as resting on an implicit set of background hypotheses, a probability distribution over them, and a mapping from the background hypotheses to the hypotheses of interest.

To illustrate, let us return to our simple example (the observer's report of artilery), and suppose we receive a second report, i.e., based on satellite photography, that artillery is present in the area. We define a new belief function based on this report by specifying a set of background hypotheses (the satellite report is reliable, the satellite report is unreliable), and by assessing probabilities over them (e.g., .8 and .2, respectively). What is our new overall belief in the presence of artillery? The set of background hypotheses for the *combined* belief function includes all combinations of the background hypotheses of the individual Arguments:

<u>ARGUMENT 1</u>

Observer Reliable (.3)	{Artillery Present} .3 x .8 = .24	{Artillery Present} .3 x .2 = .06
Observer Not Reliable (.7)	{Artillery Present} .7 x .8 = .56	{Artillery Present, Artillery Not Present} .7 x .2 = .14
	Satellite Reliable (.8)	Satellite Not Reliable (.2)

<u>ARGUMENT 2</u>

OR —
either is a ??
for Conclusion

Each combination has a probability which is the product of the probabilities of the component hypotheses. There is a simple rule for mapping these combinations onto subsets of Answers for the Question about artillery: each combination is mapped onto the subset of hypotheses which is the intersection (or elements in common) of the mappings based on the individual Arguments.

As before, Support for a subset of Answers A is just the total probability of combinations that map onto A. Thus, according to this mapping (as shown by the labels in the combinations), support for the artillery being present equals the chance that either the observer or the satellite or both is reliable, i.e., .56 + .24 + .06 = .86. This is the result given by Dempster's rule. It is displayed to the analyst as a Conclusion by the ISSUES screen in SED:

CONFLICT	TOPIC	QUESTION
0	Artillery	Is is present?

CONCLUSION

Yes	*	*
No		*
Support	.86	.14

What if the satellite report contradicts, rather than confirms, the observer? That is, the satellite evidence suggests that artillery is not present in the specified location. In that case, the new set of background hypotheses appears as below. The only change is in the mapping of the combinations to subsets of Answers about artillery. It turns out that the combination corresponding to both sources being reliable does not map to *any* subset of Answrs: since the two reports have no common elements, both cannot be true. Thus, our knowledge of the two reports forces us to prune out the impossible combination. According to the mapping, support for artillery being present equals the chance that the observer is reliable and the satellite is unreliable, i.e., .06/(1 - .24) = .08, normalizing to remove the impossible case. Support for the artillery not being present equals the chance that the satellite is reliable and the observer is unreliable, i.e., .56/(1 - .24) = .74. Once again, these are the results of applying Dempster's rule. The weight assigned to non-overlapping subsets of Answers (.24) is a measure of the degree of Conflict between the two Arguments being combined; it is the probability that the two Arguments jointly imply a contradiction.

ARGUMENT 1

Observer Reliable (.3)	Conflict .3 x .8 = .24	(Artillery Present) .3 x .2 = .06
Observer Not Reliable (.7)	(Artillery Not Present) .7 x .8 = .56	(Artillery Present, Artillery Not Present) .7 x .2 = .14
	Satellite Reliable (.8)	Satellite Not Reliable (.2)

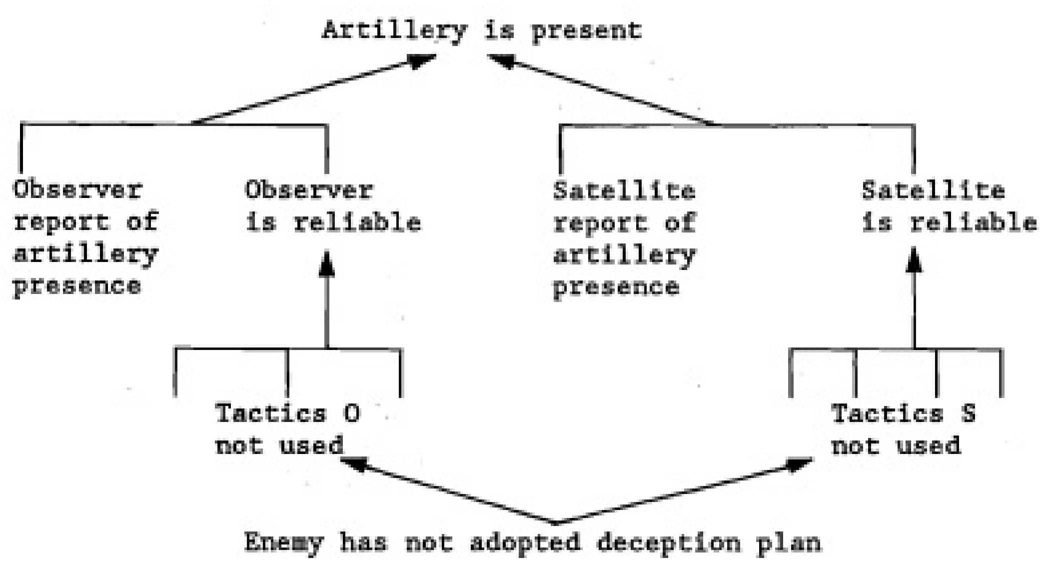
ARGUMENT 2

In the previous examples, the Argument being combined involved Premises which were themselves directly assessed; hence, they are associated by SED with internal underlying probability distributions. The manipulation of these probabilities by standard rules is what accounted for the essential features of belief functions on the hypotheses of interest. However, the analyst may also wish to combine Arguments that are higher up in an inferential chain, i.e., where the Premises are themselves the focal Issues of other Arguments. For example, the analyst might construct a second Argument regarding time of attack, based on the Premise that artillery is in forward positions, and combine it with the Argument we looked at earlier based on the distance of the follow-on Army. The Premise of the second Argument (that artillery is present in forward positions) was itself the subject of an Argument based on the report of an observer, with the background Premise that the observer was reliable. SED keeps track of the dependency of Premises on other Premises, and of those in turn on others, and so on back to the "edge" of the inference net where direct assessments must occur. Thus, a Premise in an Argument being combined (e.g., the artillery is present in forward positions) is replaced by the set of more fundamental Premises that imply its truth (viz., the observer is reliable). Since the latter Premises are associated with standard probability distributions, computations in SED always reflect the basic semantics of belief functions.

Even though the analyst has directly assessed support regarding an Issue, e.g., the reliability of the observer, it is quite possible that new evidence will later become available that bears on that Issue, and the analyst may then construct an additional Argument reflecting that evidence. SED treats direct assessments of Support as Arguments that are implicitly based on all the relevant evidence not covered by other Arguments. When the analyst subsequently constructs a new Argument, based on the new evidence, the two Arguments will be compared, examined for Conflict, and combined like any other Arguments.

Another possibility is that the analyst will wish to add qualifications, or background Premises, to a direct assessment--e.g., Support (the observer is reliable) = .3 *unless* it is raining, the enemy has placed mock-up artillery pieces in the area (to deceive us regarding the location and timing of attack), etc. The original assessment, Core Support = .3, will be modified by factoring the impact of the exception conditions into the Revised Position. In effect, .3 represents the observer's reliability even if all exception conditions are false. Support < 1.0 for the Core Position in any Argument represents residual uncertainty: it is equivalent to an additional Premise that says "this Argument works." SED associates such an assessment with an underlying probability distribution, and keeps track of the dependence of subsequent conclusions on it.

SED assumes that the fundamental probabilities utilized in its computations are independent. This assumption is what permits the multiplication of probabilities that occurs both in the derivation of a Revised Position (based on multiple false Premises) and the derivation of a Conclusion (based on combinations of Premises underlying different Arguments). The assumption of independence is not, however, restrictive. Any two Issues addressed in any inferential problem may be made dependent on one another; SED requires only that the *reasons* for the dependence be made explicit. In addition to the obvious case in which the Issues are linked *via* a chain of one or more Arguments, dependence between Issues is represented by constructing Arguments for them that have Premises in common (or if not, Arguments for their Premises that have Premises in common, etc.). As an example, the reliability of the observer and the reliability of the satellite in spotting the presence of artillery may both depend on the possibility of enemy deceptive tactics. If the methods for deceiving the satellite are the same as the methods for deceiving a human observer (e.g., a single type of mock-up could lead either one into "false positives"), we can express the correlation between satellite and observer reliability by qualifying both with the same Premise (that no such deceptive tactics have been employed). Suppose, however, that tactics O would be used to deceive a human observer and tactics S to deceive a satellite. Then the Argument for the observer's reliability says Support (observer is reliable) =.3 unless tactics O are used, and the Argument for the satellite's reliability says Support (satellite is reliable) = .8 unless tactics S are used. The non-indepdenence is then represented by creating two new Arguments: (1) If the enemy adopts a policy of deception regarding attack plans, then it will use tactics O unless..., and (2) If the enemy adopts a policy of deception regarding attack plans, then it will use tactics S unless... . If both observer and satellite have reported the presence of artillery, the structure of the inference would be the following:



Evidence for a policy of deception could then weaken the Conclusion that artillery is present (and any subsequent Conclusions regarding time or location of attack) by two different routes: *via* its impact on the reliability of the observer and *via* its impact on the reliability of the satellite.

3.5 Making Assumptions

Knowledge requires assumptions. An analyst will be justified in believing nothing at all unless he is prepared to act *as if* other things were true. Even in cases of reasonable certainty, e.g., when two reliable and independent sources confirm a Conclusion, there is the *possibility* of error (satellite photographic evidence can be fooled; a human informant may be misled). When sources do not agree, the dependence on assumptions merely becomes more salient. No analyst has the time or resources to rule out ahead of time all possible exceptions to a Conclusion (and exceptions to those exceptions, etc.). In short, although he may have knowledge or evidence regarding some of the Premises of an Argument, such knowledge will never be complete or completely certain.

Premises about which the analyst is unsure may thus play a crucial role both in his understanding of and reasoning about the problem, and in decisions regarding the collection of further information. For traditional Bayesians, knowledge about an uncertain event is fully revealed in a single choice or judgment and summarized by a single number. In such models, there are two principal ways to deal with variables about which one is ignorant: (1) omit them from the analysis altogether, or (2) make explicit probabilistic assessments. In both cases, assumptions are not so much avoided as swept under the rug. For example, the analyst might try to deal with ignorance by making the probabilities equal, e.g., judge that there is a 50% chance an unknown human source is reliable and a 50% chance he is not. As a result, his confidence in the conclusion will be cut in half. But there is no way to distinguish this case (where nothing is known about the source) from the case in which a large amount of evidence points equally in both directions. Moreover, the analyst could choose to represent the *same* state of ignorance by dividing up the possibilities differently before assigning equal probabilities: e.g., the source is accurate and honest, inaccurate and honest, accurate and dishonest, or both inaccurate and dishonest; in this case, his confidence will fall to 25%. The conclusion thus depends rather strongly on arbitrary (and unspecified) assumptions. Alternatively, the analyst may base the assessments on whatever knowledge he has; e.g., he may search his memory for experiences with human sources that resemble the present one in any way (e.g., same nationality, holding a similar government post, similar means of recruitment, similar family situation, same age, etc.). The result may well be a more definitive assessment--but the analyst will have to (implicitly or explicitly) make assumptions about the relevance of each aspect of similarity and dissimilarity, the independence of their effects, and the representativeness of the present case with respect to each.

The belief function model implemented in SED permits the representation of ignorance by assigning Support to subsets of Answers rather than individual Answers. Arbitrariness is removed since different partitions of the same possibilities do not require reallocations of Support; for example, if we know nothing at all about a source, we may set Support({source is reliable, source is unreliable}) = Support({source is accurate and honest, source is inaccurate and honest, source is accurate and dishonest, source is inaccurate and dishonest}) = 1.0. Other approaches to representing ignorance involve higher-order probabilities, convex sets of probability distributions, interval probabilities, and fuzzy probabilities.

SED combines a representation of ignorance with the ability to make assumptions. The concept of an assumption in SED means two things:

(1) Assumptions are beliefs that are constrained by, but go beyond what is more firmly known.

(2) Assumptions are beliefs that are subject to retraction when they conflict with other beliefs.

Could a rational decision maker get along without assumptions in this sense? To do so, he would have to deny (1) that any of his numerical judgments of belief are more firmly based than others, and (2) that he would ever retract such judgments in case of unexpected conflict with other lines of reasoning. In an ideal universe, where judgments reflect the totality of relevant knowledge, such claims may be plausible. In the real world, they are not.

The introduction of assumptions has important implications for how we think about uncertainty models. In any particular decision or judgment, a decision maker may well adopt assumptions that go beyond his more firmly based beliefs. In a different context, he may have reason to retract those assumptions and/or adopt different assumptions, thus making different choices or judgments, yet drawing on exactly the same base of knowledge. It is therefore necessary to distinguish the *manifest* (or behavioristic) meaning of a piece of evidence from its cognitive or *latent* meaning. The manifest meaning (e.g., the current impact of a piece of data in an Argument for a particular hypothesis) is revealed by a present decision or a present judgment, and it depends on both firm beliefs and a particular (possibly temporary) selection of assumptions. The latent meaning refers to all the *potential* impacts of the evidence on reasoning; and this can only be represented in a model structure that includes both firmly held beliefs and the set of possible assumptions from which the decision maker chooses on any given occasion of decision or judgment (cf., Loui, 1986). SED helps users build and manipulate such structures.

The two definitions of Assumption (going beyond firm belief, and subject to retraction in case of conflict) correspond to two methods for eliciting assumptions from users of SED; they provide a pair of converging operations whose agreement indicates that assessments are being made in a coherent manner.

The first method for assessing assumptions is "bottom-up": i.e., start with Arguments based on firm beliefs and make them more precise by Assumption. For example:

$\boxed{\text{POSITION}}$ CORE EXCEPTIONS

	Front CDR	Has he misestimated distance?
Yes		*
No	*	*
Support	.4[1]	.6[2]
% Assume	1.0[1]	
Final Support	1.0[2]	

(1) Assessed by analyst

(2) Supplied automatically by SED

The analyst has directly assessed Support of .4 for the Front Commander's not having misestimated the distance. This assessment may be based on the analyst's knowledge of the general capabilities of Soviet Commanders and their staffs, the time and information sources available under prevailing conditions, etc. This knowledge, of course, does not take him very far (e.g., it says nothing about this particular Commander). By his judgment, it leaves .6 Support uncommitted with respect to the possible Answers, Yes or No. This degree of uncertainty, however, would seriously cripple the analyst's Argument for time of attack based on the distance of the Armies. The Revised Position (even if all other Premises are known to be true) would be:

POSITION REVISED

	Attack	What time will it occur?
< 48 hours		*
48 - 72 hours		*
72 > hours	*	*
No attack		
Support	.4	.6

In short, 60% chance that no valid time of attack can be inferred. Let us suppose that such a result clashes with the analyst's judgment regarding the actual force of that Argument. In traditional systems, there is no way to reconcile these two judgments: (i) uncertainty about whether the Commander misestimated the distance and (ii) reasonable confidence in the Argument that time of attack will reflect the actual distance of the Armies. The analyst would be compelled either to exaggerate his knowledge about the former or to relinquish his confidence (and his ability to act) on the latter.

SED solves this problem by making a distinction between what is firmly known about a proposition and the impact it has on a current Argument. Its impact can be increased provisionally over what is strictly warranted by firm belief.

In SED the analyst is free to reallocate any or all of the uncommitted support by *Assumption*. More exactly, he can allocate Support that was committed to the set {Yes, No} to any proper subset, i.e., to {Yes} or to {No}. In this example, he has chosen to allocate 100% of the uncommitted Support to {No}. The result is Final Support for {No} equal to $.4 + (1.0)\ (.6) = 1.0$. The Argument for time of attack will thus proceed with the desired force--subject to eventual possible retraction of the Assumption(s) upon which it depends.

The second method for assessing Assumptions in SED (not yet implemented) is "top-down": i.e., specifying how much of a belief is firm and how much less precise the analyst would be willing to make it if it conflicted with other Arguments. For example,

POSITION CORE EXCEPTIONS

	Front CDR	Has he misestimated distance?
Yes		*
No	*	*
Support	.4[2]	.6[2]
% Firm	.4[1]	
Final Support	1.0[1]	

(1) Assessed by analyst
(2) Supplied automatically by SED

Here, the analyst begins with an assessment of Final Support for {No}, i.e., how much net impact that Answer will have in the current Arguments where it is a Premise. In this example, the analyst has chosen to act in these inferences *as if* he were certain about the Commander's estimate of the distance; thus, he has assessed Final Support as 1.0. However, available knowledge does not justify 100% certainty regarding that proposition. He now reflects on how much of the 1.0 Final Support is "firm" and how much is "soft"--i.e., how much of it he would be willing to give up in the limiting case of many other strong lines of Argument conflicting with it. In this example, he has judged that the Final Support is only 40% firm; he would be willing to transfer 60% of the Final Support for {No} to the superset {Yes, No} in case of conflict.

In both approaches, Assumptions are constrained by firm belief: in the first, by the requirement that Core Support for a set of possibilities can only be reallocated by Assumption to a proper subset of those possibilities; in the second, by the requirement that Final Support for a set can only be shifted (in case of conflict) to a proper superset. Support can be focused more sharply by Assumption (and widened by Conflict); but its direction cannot be altered.

The demarcation between knowledge and Assumption is not absolute and fixed. Firmness of knowledge is a matter of degree: Assumptions need not be entirely without evidential warrant; conversely, any belief might be retracted under some circumstances and thus have to be regarded as an Assumption. The location of the boundary between "firm belief" and "Assumption" is thus a matter of judgment for the problem at hand. Nevertheless, the distinction is a real one: there are beliefs the analyst is likely to hold onto come what may, and other beliefs that he is more likely to relinquish in the face of unanticipated conflict. The ability to draw such a boundary, even if it is itself a provisional one, is a powerful tool for capturing crucial aspects of reasoning about evidence.

3.6 Resolving Conflict

Two very different approaches to conflicting evidence have been adopted by students of inference. In logic-based systems, if it is possible to derive a contradition from a set of statements, then one or more of the statements must be false. Suppose, for example, we start from the following beliefs:

Argument #1. If Source A reports anything, it is true.
Source A reports R.
R implies S.

Argument #2. If Source B reports anything, it is true.
Source B reports Q.
Q implies $\neg$S.

From these two arguments, we could infer an impossibility: the truth of both S and $\neg$S. To remove the inconsistency, at least one of the beliefs responsible for it must be revised. We *know* we are wrong about at least one of the following: the credibility of Source A or B, what they reported, or the implications of what they reported for S, $\neg$S.

A quite different approach has been adopted in systems that quantify and combine *degrees* of belief, like probability theory, fuzzy logic, or Shafer-Dempster theory. Suppose we believed:

Argument #1. Support (If Source A reports anything, it is true) = .99
Support (Source A reports R) = .99
Support (R implies S) = .99

Argument #2. Support (If Source B reports anything, it is true) = .99
Support (Source B reports Q) = .99
Support (Q implies ¬S) = .99

Although it may follow that we have very strong evidence for S and very strong evidence for ¬S, there is no logical contradiction. Even strong evidence may be imperfectly correlated with hypotheses. Legitimate evidential arguments may, therefore, point in different directions as long as each argument falls short of conclusive proof. Thus, it is *conceivable* that all our original beliefs were correct: both Source A and Source B are highly credible; A reported R; B reported Q; the former is strong evidence for S; and the latter is strong evidence for ¬S. The more pertinent question is whether it is still *plausible*, in light of this conflict, to regard all these beliefs as true.

The first approach to conflicting evidence is *epistemic*: conflict is regarded as a symptom of faulty beliefs and is used as an opportunity to correct them--by explicitly identifying potentially erroneous steps in the conflicting arguments. The second approach may be loosely referred to as *stochastic*: conflict among imperfect arguments is *expected* to occur by chance some portion of the time, and it is dealt with not by changing the arguments, but by statistically aggregating them when they both apply.

Each approach has virtues: On the one hand, the "stochastic" view, unlike the epistemic, permits gradations of belief; moreover, belief revision in epistemic systems is often arbitrary since there is no principled way to select one culprit from among the many beliefs responsible for a contradiction (cf. McDermott and Doyle, 1980). On the other hand, the stochastic approach is likely to "resolve" conflict in ways that are unconvincing and that fail to extract permanent lessons that might improve future inferences. Resolutions of conflict by stochastic methods are typically either too bland or too definitive. In the example above, since Arguments #1 and #2 are equally strong in support of S and ¬S respectively, the Conclusion is equal Support for S and ¬S. If both Arguments had been 100% certain, there would have been no determinate answer at all. The stochastic approach is even more likely to produce overly definitive results, as in the following hypothetical cases:

- Argument #1 strongly supports hypothesis S_1 but allows a very small chance that S_2 is correct; Argument #2 strongly supports hypothesis S_3 but allows a very small chance that S_2 is correct. Statistical aggregation (Bayes' Rule, Dempster's Rule, etc.) results in 100% belief in S_2, which both sources regarded as highly unlikely (cf., Zadeh, 1984).

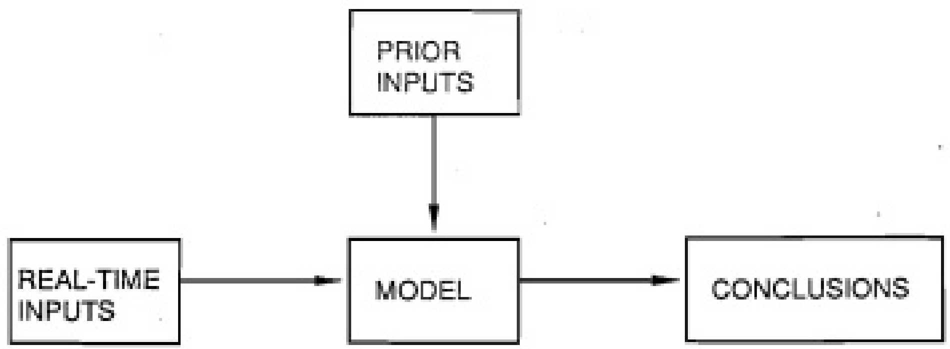
- Argument #1 strongly supports S and Argument #2 strongly supports ¬S, but the degrees of support are not quite symmetrical, e.g., 99 to 2 in favor of S for Argument #1, 99 to 4 in favor of ¬S for Argument #2. The result: a 2 to 1 preponderance of belief in favor of S.

- According to Argument #1, ¬S is impossible; according to Argument #2, ¬S is favored 10^{10} to 1. The result: 100% belief in S.

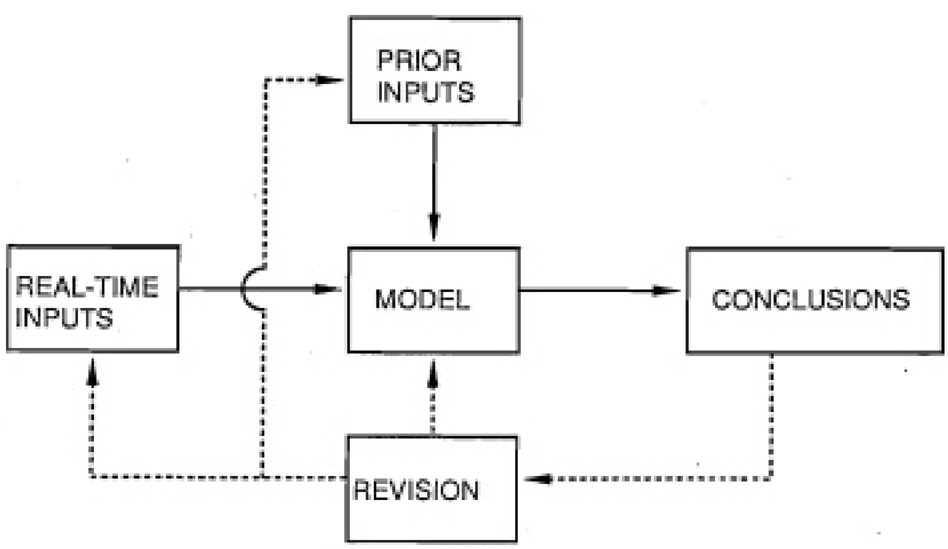
For most people, these conclusions will seem a bit premature. Not surprisingly, therefore, the initial response to conflicting Arguments is epistemic, rather than stochastic. Even when conflicting Arguments have been expressed numerically, people look for *reasons* for the conflict: Did I overestimate the accuracy or honesty of one or both sources? (e.g., Should I reduce my belief in Source A's credibility from .99 to something lower)? Was I wrong in my understanding of what they said? Do my conclusions really follow from my understanding of what they said? The result, hopefully, is both a more convincing resolution of the conflict and an enhanced store of permanent knowledge.

In SED, numerical measures and an epistemic response to conflict are complementary rather than mutually exclusive. Conflict resolution is carried out by higher-order processes that *reason about* quantitative uncertainty models; conversely, numerical measures from those models provide guidance for decisions about adopting and revising Assumptions. The result is a generalization of the epistemic approach, in which belief is graded, conflict is a matter of degree rather than all-or-none, and Assumption revision is intelligently directed at those beliefs that are most likely to be in error.

At the same time, SED represents a radical change in perspective on the stochastic approach. In most computerized aids that quantify uncertainty, inference is equated with an essentially linear process, in which a model or "knowledge base" is built, numerical inputs assessed, and outputs generated:



Such an approach may ensure consistency of inputs and outputs with respect to a set of axioms, e.g., probability theory; the problem is, more than one set of inputs and outputs, with vastly different implications for a decision, will be equally acceptable from a strictly *formal* point of view. Automation of uncertainty handling thus omits the *thinking* processes by means of which an analyst selects *one* consistent set of beliefs out of all those that are possible. Actual probabilistic reasoning is typically highly iterative: the results of one line of reasoning are compared with the results of other lines of reasoning (or with direct judgment); if there is a discrepancy, the inputs, parameters, and even the structure of the model or knowledge base may be revised:



Higher-order reasoning about the knowledge and assumptions underlying an uncertainty model is, thus, a required aspect of the "stochastic" approach. In SED, such reasoning is no longer hidden from view. SED provides direct support for the intelligent construction and modification of inference models in the light of experience with their application. In effect, SED redefines "reasoning": it is no longer the blind *application* of an uncertainty model, but its creation and maintenance.

Under what circumstances does the existence of Conflict justify changing one's beliefs? In making this decision, the analyst might consider: (1) the firmness with which he held the beliefs that led to the Conflict, (2) what the Conflict now tells him about the chance that each particular belief is wrong, and (3) the relative costs of retaining a belief if it is erroneous and rejecting it if it is correct. In an automated system, it would be possible to incorporate these factors into a formal algorithm. SED, however, adopts a more informal, interactive approach. SED supports (1) by making a relatively coarse distinction between Assumptions and other, more firmly held beliefs. It provides two measures that together support judgments about (2): the total amount of Conflict, and an estimate of the amount of Conflict attributable to a particular Assumption. The costs of different kinds of errors (3) are left for the analyst to weigh informally. [illegible] to belief

In SED, the total amount of Conflict between two Arguments is interpreted as evidence that beliefs contained in those Arguments are mistaken. This is a straightforward generalization of the logical strategy of showing a belief to be false by deriving a contradiction from it. Let T be the conjunction of beliefs in Arguments 1 and 2, e.g., the Revised Position of Argument 1 and the Revised Position of Argument 2. If T implies p and $\neg$p, then $\neg$T. In SED, T implies a quantitative weight on p and $\neg$p, corresponding to the *chance* that the beliefs in T (e.g., the two Revised Positions) imply a contradiction; that weight can be taken as the chance that Conflict proves at least one of the members of T to be false. If the Conflict measure is acceptably small, the two Arguments can be left as they are and Conflict resolved stochastically (in effect, by dropping the impossible states of affairs from the calculations). If the measure is large, however, it may be wiser to take a closer look at the contents of T.

There is an analogy here to hypothesis-testing in classical statistics. If the probability of an obtained sample is too low given the assumption of the null hypothesis, we reject the null hypothesis. In both cases, decision making is *heuristic*. Since prior beliefs and the costs of different kinds of errors are both left at least partly implicit, the choice of a "significance level," or threshold of acceptable conflict, is to some degree arbitrary.

The measure of total Conflict is an indicator of something wrong in the stochastic combination of two Arguments. The analyst, however, needs to focus his search for those Assumptions that seem to bear the most responsibility. The second measure relies on a rough decomposition of the Conflict into components that are attributable to separate Assumptions. These components may then be interpreted as approximately proportional to the chance that Conflict proves each particular Assumption to be false.

Let us return to the very simple example in which an observer reports that artillery is present but satellite photographic evidence suggests that it is not. Suppose the analyst has assumed that human observers of the relevant sort were pretty much reliable until proven otherwise: e.g.,

POSITION CORE 1 of 1

	Observer	Is he reliable?
Yes	*	*
No		*
Core Support	.3	.7
% Assume	.5	
Final Support	.65	.35

Although the available evidence regarding this observer's reliability is 70% inconclusive, the analyst assigns 50% of the uncommitted Support to the proposition that the observer is reliable, for a total Final Support of .3 + (.5) (.7) = .65.

Using ISSUES to combine this Argument with the Argument based on satellite evidence, we get:

CONFLICT	TOPIC	QUESTION
.52	Artillery	Is is present?

CONCLUSION

Yes		*	*	*
No			*	*
	Support	.27	.58	.15

The combined Argument favors {artillery not present}, since the satellite evidence is still regarded as superior to the observer (80% reliable). But there is a substantial amount of Conflict (52%). To see where these numbers come from, it is useful to depict the combined Arguments in terms of the relevant background hypotheses and probability distributions. In doing so, we separate out the contribution of Argument 1 to Support({observer reliable}) due to firm belief (.3) and to Assumption (.35):

3 measures:
(A) total conflict
(B) contrib to conflict
(C) contrib to Δ belief

ARGUMENT 1

	Satellite Reliable (.8)	Satellite Reliability Unknown (.2)
Observer Reliable (.3)	CONFLICT .3 x .8 = .24	{Artillery Present} .3 x .2 = .06
Assume: Observer Reliable (.35)	CONFLICT .35 x .8 = .28	{Artillery Present} .35 x .2 = .07
Observer Reliability Unknown (.35)	{Artillery Not Present} .35 x .8 = .28	{Artillery Present, Artillery Not Present} .35 x .2 = .07

ARGUMENT 2

Can't treat conflict as ...

① (.8)(.35)(.8)/.35 = .48

prob of ⊥ given assumed hyp. is true

contrib. to conflict

③ Δ conflict / Δ assumption ≈ .8

B = prob. assumption is false
(smaller when effect depends on other ...)

usefulness of assumption: increase in conflict vs total increase in ...

.28 of the total .52 Conflict is jointly due to the Assumption that the observer is reliable (.35) and the belief that the satellite is reliable (.8). *Given* that the latter belief is firm, the .28 can be taken as evidence specifically against the Assumption. Note that the contribution of an Assumption to Conflict is a joint function of its size (.35) and the magnitude of the beliefs it conflicts with (.8). also - what it depends on (chain of reasoning it is part of)

The situation is only a bit more complicated when more than one Assumption is involved. Suppose the analyst adopts a similar Assumption with regard to satellite photographic evidence, i.e., allocating 50% of the uncommitted support to the proposition that the satellite report is reliable. Looking at background hypotheses and probabilities, and again separating out Assumptions and firm beliefs, we get:

ARGUMENT 1

ARGUMENT 1 \ ARGUMENT 2	Satellite Reliable (.8)	Assume: Satellite Reliable (.1)	Satellite Unreliable (.1)
Observer Reliable (.3)	CONFLICT .3 x .8 = .24	CONFLICT .3 x .1 = .03	{Artillery Present} .3 x .1 = .03
Assume: Observer Reliable (.35)	CONFLICT .35 x .8 = .28	CONFLICT .35 x .1 = .035	{Artillery Present} .35 x .1 = .035
Observer Reliability Unknown (.35)	{Artillery Not Present} .35 x .8 = .28	{Artillery Not Present} .35 x .1 = .035	{Artillery Present, Artillery Not Present} .35 x .1 = .035

ARGUMENT 2

In this case, the total Conflict is .24 + .03 + .28 +.035 = .585. The contribution to the Conflict by the Assumption about the observer is .28 + .035 = .315; and the contribution to the Conflict by the Assumption about the satellite is .035 + .03 = .065. The CONFLICTS screen lists these Assumptions in order of their apparent "culpability":

$$\frac{prob(2|A)}{prob(2)} = \frac{.9}{.585}$$

CONFLICT	TOPIC	QUESTION
.59	Artillery	Is is present?

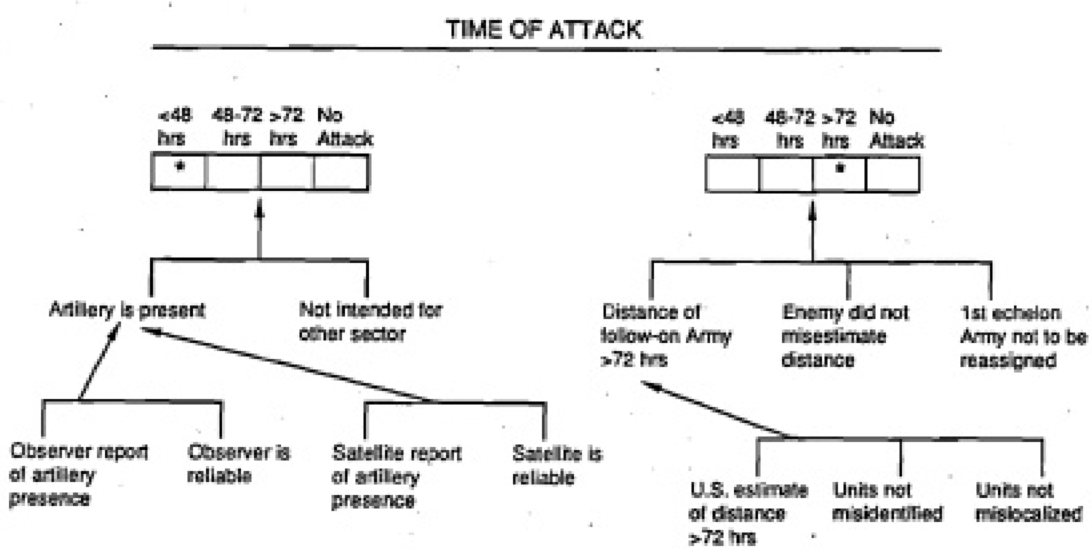
CONTRIBUTION TO CONFLICT	TOPIC	QUESTION
.32	Observer	Is he reliable?
.07	Satellite	Is it reliable?

These numbers might lead the analyst to drop the Assumption regarding the observer's reliability (thereby reducing total Conflict to .59 - .32 = .27). Even though dropping the Assumption about the observer will not significantly affect the Conclusion in this case ({artillery not present} will still be favored), it may well improve the accuracy of Arguments in the future that rely on that observer.

Notice, however, that the two Assumptions in the above example were not independent in their impact. After dropping the first Assumption, the contribution of the second Assumption to Conflict would be reduced from .065 to .03, since part of the total Conflict (.035) was jointly determined. Each measure of an Assumption's contribution to Conflict is thus a sort of upper bound, conditional on retaining both the other Assumptions and the firm beliefs that it clashes with. Assumptions have by definition a higher prior likelihood of being in error than firm beliefs and are thus more likely to be retracted. Hence the measure of an Assumption's contribution to Conflict is less ambiguous when the Assumption clashes only with firm beliefs; in this case, it is more readily interpretable as the chance that Conflict proves the Assumption wrong. Indeed, if firm beliefs could *never* be withdrawn, SED might focus exclusively on the Conflict attributable to Assumptions (in this example, .28 + .035 +.03 = .345). Only this portion of the Conflict would be treated epistemically; Conflict due to firm beliefs alone (.3 x .8 = .24) would always be handled stochastically.

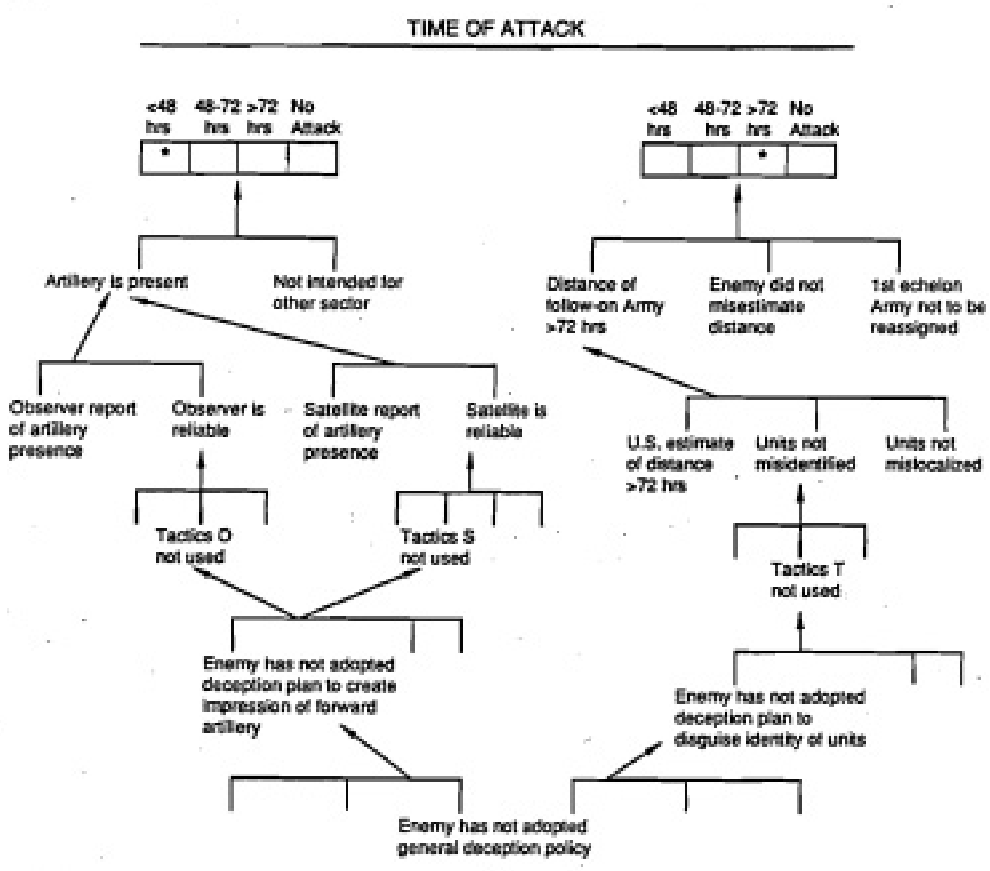
For SED, however, the boundary between Assumptions and firm beliefs is itself subject to review. A large measure of Conflict, if there were no Assumptions or no Assumptions clearly identifiable as culprits, might very appropriately lead an analyst to reexamine the relevant "firm beliefs" (e.g., using the GROUNDS and ARGUMENTS commands). He might then convert a firm belief into an Assumption by using the "% Firm" option in the ARGUMENTS screen. Alternatively, he might add exception conditions to the Argument expressing a firm belief (as with the "crystal ball" technique). He might then return to CONFLICTS to observe the potential effect on Conflict of dropping the newly defined Assumptions. Conflict resolution is thus an occasion for the continued elicitation and refinement of the analyst's beliefs.

Conflict can help an analyst search deeply through a network of beliefs for a potential culprit, and revisions may be made at any level. ~~In particular~~ An Example of reasoning with a fairly complex network is, conflict resolution may be a valuable tool for detecting deception, since evidence for detection is often available only in the form of evidential conflict. Suppose, for example, that we build an Argument for time of attack <48 hours, based on the observer's report of artillery in forward areas. Suppose that this Argument conflicts with other evidence, e.g., the Argument based on the distance of the follow-on Army, which suggests time of attack >72 hours. The observer's report is, however, confirmed by satellite observations of forward artillery. An inferential structure containing this evidence looks like the following diagram (each arrow points to the Core Position of an Argument, and items under the square brackets are Premises):



So far, deception has not been taken into account. The analyst, however, may construct an Argument that conditions the observer's reliability on a variety of factors, including that deceptive tactics O not be used by the enemy (e.g., setting up phony artillery pieces to fool human observers); he may also condition the reliability of the satellite on the enemy's not using deceptive tactics S (e.g., not creating mock-ups that could fool the satellite). These two Premises may not, however, be independent (Section 2.6). They may depend on Arguments that share a Premise: e.g., that the enemy has not adopted a deception plan in that sector (in order to mislead regarding the time and location of attack). Evidence for a *later* attack may also be tainted by the possibility of deception: e.g., tactics T that cause us to misidentify rear units as part of the follow-on force and fail to detect the true follow-on force that is already forward. Deception in this case is not *directly* related to the possibility of fooling the observer or the satellite. But there may be a more indirect connection: e.g., if we have learned from experience that the enemy's approach to battle generally includes deception, that knowledge supports the chance that both of the more specific kinds of deception might take place in appropriate circumstances.

What we end up with is a double hierarchy of Arguments: one hierarchy moving upwards toward time of attack, and the other moving downwards toward more general types of deception:



"Benefit [illegible] by size of Assumption?

Conflict among Arguments in regard to time of attack may now cause the analyst to reexamine his Assumptions about deception at any of these levels. The farther down an Assumption is, the more it depends on other beliefs or Assumptions for its impact, hence the more diluted is its contribution to Conflict. More specific and local types of deception will thus be looked at before more general and sweeping types of deception. This effect may be counterbalanced, however, if a more general Assumption contributes to Conflict in more than one way. For example, if there were considerable evidence against an early attack, the simplest resolution of conflict might be to drop the Assumption that there is no deception regarding forward artillery; this has the effect of weakening the Arguments for early attack based on *both* the observer's report and the satellite's report.

Conflict in one Arg may be counterbalanced by benefits in many others. ∴ harder to revise more "fundamental" beliefs.

Note that direct evidence regarding deception may also be obtained at any level: e.g., general experience of enemy deceptive tactics, overheard communications regarding a deception policy in a particular sector, and localized anomalous observations, e.g., no movement, personnel support, or electronic emissions associated with apparent artillery pieces. Such evidence can produce positive support *for* deception, thus limiting the scope of any possible Assumptions. Conflict resolution uses indirect evidence in a complementary fashion, by prompting the retraction of Assumptions *against* deception. The impact on subsequent Arguments (e.g., for time of attack) is the same, since ignorance of a Premise caused by retracting an Assumption is equivalent to denial of the Premise.

Interpret "conformist bias" as effect of this model

Fundamental beliefs play a role in more inferences.

Conflict resolution supports learning from experience. The paradigmatic case of learning, however, is one in which a general belief is shaped by observation of a large number of instances. For example, we might form a low opinion of the observer's reliability after many cases in which what he reported was found not to be true. Conflict resolution generalizes this paradigm in two important ways:

(1) An analyst may not have the luxury of learning over a long series of repeated instances. With SED, he can change his less firm beliefs about the observer based on just one occasion in which what the observer reports is false. Conflict resolution is a form of "explanation-based" learning (DeJong and Mooney, 1986), in which failure of a prediction in even one instance

prompts a search for an explanation of the failure. SED helps the analyst identify exception conditions whose truth might account for a Conflict; such "explanations" may be local to a particular inference on a particular occasion, or they might affect general beliefs about a domain that reappear in numerous subsequent inferences.

(2) No definitive feedback regarding the correctness of his beliefs may be available to the analyst. For example, he may never know for sure that the observer was wrong, only that other (uncertain) Arguments contradicted him. SED permits learning to take place even in the absence of "ground truth." If no one source can be thought of as infallible, SED calibrates competing sources against one another. The observer simultaneously provides feedback regarding the correctness of the Arguments that contradict *him*.

SED embeds numerical uncertainty representations within a process of higher-order reasoning about knowledge and Assumptions. Is such a higher-order process really necessary? Could the functions of conflict resolution be accomplished instead *within* a standard numerical calculus? The answer is: in principle, yes; in practice, no. Shafer (1981a) and others have argued against the notion that all change of belief can be characterized simply as conditionalizing one's current beliefs on new evidence. In the real world, decision makers cannot be expected to anticipate all possible evidence in their current beliefs. When evidence occurs that is not anticipated (or in *combinations* that are not anticipated), rote calculation (e.g., by Bayes' rule) is inadequate; a new set of beliefs must be *constructed*. Conflict resolution in SED is a method for *intelligently* performing that task, under the guidance of previous assessments about firmness of belief and automated calculations about contributions to Conflict.

To simulate the effect of conflict resolution within a numerical calculus, it would be necessary to explicitly represent all the situations in which conflict could arise and decide on a resolution ahead of time. We would need a vast number of exception conditions specifying which other sources and Arguments would override a given Argument, e.g.,

How to find culprit?
Likelihood of evidence given all premises = 1;
Likelihood of evidence given all but 1 premise = ?.

Assumption revision need not be all-or-none: subst multiple Assumptions.

In extreme: every Assumption involves some small amount of uncertainty [illegible]

get prob [illegible]

- Source A is reliable when he reports R *unless* source B reports Q and source C reports T and source D reports U... or source E reports V and source F reports W... or ...

In a numerical framework (e.g., Bayesian or Shaferian), a huge set of conditional assessments would be required, linking the elements of every actual line of reasoning to the elements of all other possible lines of reasoning. The price of such a strategy comes not only in the sheer quantity of inputs and computational intractability, but also in a loss of naturalness and modularity. Why should the analyst worry ahead of time, for example, about the relationship between the reliability of techniques for estimating the distance of two Armies and the reliability of a particular human observer? Linkage between beliefs about these two sources becomes relevant only when and if they happen to participate in conflicting lines of reasoning on the same topic, e.g., in regard to enemy time of attack.

In order to remain tractable, numerical inference models typically treat hypotheses about diverse information sources or lines of reasoning as if they were independent. The result is a stochastic approach to conflict that fails to extract its real significance when it occurs. SED achieves the best of both worlds: it retains economy and modularity of representation without in fact assuming that different lines of reasoning are independent. It does so by shifting the burden of dealing with conflict from the calculus itself to the processes that create and manage the calculus. Conflict among Arguments causes SED to reach inside of each of the Arguments, looking for the weakest links in each line of reasoning--even though the co-occurrence of the two Arguments was never anticipated. SED enables the analyst to bring to bear the conclusions of one Argument on the evaluation of the other without sacrificing the modularity of the different lines of reasoning.

3.7 Communicating Conclusions

In communicating conclusions based on incomplete, unreliable, and inconsistent data, the intelligence analyst faces a dilemma. If he reports only the uncontroversial elements of divergent views: e.g.,

Support ({time of attack 0 - 72 hours}) = 1.0
(i.e., it is nearly certain that attack will occur at some time or other),

confused decision ~~policy~~-makers are likely to object that he is too imprecise. If he provides an explicit account of competing, but precise possibilities: e.g., 70

Support ({time of attack <48 hours}) = .4
Support ({time of attack >72 hours}) = .6,

he may be accused of being too indecisive. Nevertheless, if he takes a precise and definite position, e.g.,

Support ({time of attack >72 hours}) = 1.0,

and it turns out to be wrong, the consequences may be even worse.

SED offers no magic solution to this analyst's dilemma. Perhaps SED's main contribution is to help the analyst organize and understand large quantities of data, and thus to increase the likelihood of reaching precise, definite, *and* correct Conclusions. Nevertheless, SED offers a flexibility in the way Conclusions are *reported* (even when they remain uncertain), that may also facilitate communication between analysts and intelligence consumers.

To summarize a Conclusion in SED, three things are, in principle, required:

(1) the Conclusion *per se*, i.e., an assignment of Belief to subsets of Answers on the Topic/Question of interest;

(2) the amount of unresolved Conflict; and

(3) the most important Asssumptions upon which the Conclusion depends.

(The analyst would also use SED to lay out the evidence and reasoning that underlies his results; but here we are focusing only on a top-level *summary*.) The same Conclusion has quite different significance depending on the amount of Conflict or the number of Assumptions it is associated with. Summaries that stop at item (1) are therefore problematic. An analysis that simply reports a probability for an event tells us nothing about the *quality* of that probability. Indications of possible error in an analysis may be either internal or external, corresponding to (2) and (3) respectively. Conflict reflects the chance of error somewhere within the analyst's reasoning. Assumptions reflect the chance of error due to what has been left out of the analysis: i.e., the data that would be required to confirm or disconfirm the Assumptions ~~(3)~~.

A further advantage of reporting all three elements is that it provides the analyst with more degrees of freedom in how he describes item (1). SED permits the analyst to explore a space of solution representations by imposing or rejecting Assumptions regarding elements of the analysis. Within the constraints of firm belief, a Conclusion might be reported in a very imprecise or non-definite form if few Assumptions are made; or it might be reported in a more precise and definite form, at the cost of adding Assumptions (and possibly increasing Conflict). No one of these representations is inherently better. Depending on the requirements of the decision at hand, one or the other (or conceivably both) might be preferred. Investigation of such a space of solutions may be utilized to develop the analyst's understanding of the problem, test the sensitivity of Conclusions, and select a representation that is both justified and suits the information requirements of intelligence consumers.

An illuminiating strategy may be to work backwards from a candidate Conclusion that is both precise and definite. The analyst asks how much revision of Assumptions and, possibly, firm beliefs would be required to arrive at that Conclusion; the analyst then goes on to consider another candidate Conclusion and asks the same question; and so on. The result of such an exercise is (1) an understanding of just how precise and definite a Conclusion the evidence will reasonably support, and (2) an informal evaluation of alternative precise and definite results in terms of "closeness of fit" to the evidence.

This process can be facilitated by use of conflict resolution in a what-if mode. The analyst hypothesizes a precise and definite candidate Conclusion and creates a temporary Argument (based on no evidence) that assigns it Core Support = 1.0. Examination of the CONFLICTS screen will reveal the primary

Assumptions that conflict with the proposed hypothesis; dropping them may reduce Conflict without appreciably reducing the plausibility of the analysis.

Once a reasonable and reasonably precise and definite Conclusion has been reached, conflict resolution can be used again in a what-if mode to retrieve the Assumptions that underlie it. The analyst simply creates a temporary Argument for the *complement* of the hypothesis, with Core Support = 1.0. The CONFLICTS screen will now display all Assumptions that conflict with this counter-hypothesis, and thus support the hypothesis of interest.

4.0 CONCLUSION

While the difficulties of collecting intelligence data are well understood, the difficulties in analyzing and interpreting those data are often overlooked. There is a growing awareness, however, that the success of the overall intelligence enterprise depends crucially on those processes which occur *after* the data have been collected. The present report has described a tool which is designed to make the intelligence analyst's task easier and more successful.

Uncertainty, of course, is at the heart of the intelligence analyst's job. While a variety of technologies for handling uncertainty have been introduced by statisticians and by expert system builders, they suffer from a variety of drawbacks:

- The meaning of numerical assessments is often unclear.

- A combinatorial explosion of assessments is required even in simple problems.

- Systems that automate the handling of uncertainty are not truly "intelligent" or robust in the range of situations that they address.

SED embodies promising technical solutions to all three of these problems. It clarifies the meaning of numerical assessments by emphasizing qualitative models of how evidence is linked to conclusions; it requires only simple numbers reflecting different ways that such evidential links could be broken. More complex numerical models are then automatically generated. SED wards off a combinatorial explosion of assessments by introducing a simple method for deriving the impact of multiple factors on a conclusion from assessments of their separate impacts, and by providing for non-independence of different lines of reasoning through a higher-order process of conflict resolution. As a result, SED encourages, rather than discourages, users to introduce new factors into an analysis: i.e., to make the reasons for uncertainty explicit. When there is no direct knowledge regarding such factors, it permits users to declare their ignorance and, if they wish, to adopt Assumptions. Finally, SED does not assume that a problem has been solved simply because a numerical model has been created. It focuses on the processes that intelligently create and revise such models. When two or more arguments point in different directions, SED does not sweep conflict under the rug by statistically aggregating them. It supports the analyst in a process of re-examining and modifying beliefs and Assumptions that contributed to the conflict. More generally, by working with the analyst at every stage of Argument construction, SED provides a framework which is compatible with the analyst's natural approach to his task while at the same time significantly improving on it.

In a variety of respects, however, SED is subject to improvement. Additional research might address some of the following issues:

Knowledge & Inference

- *Accumulating and using knowledge*. SED might store general knowledge about a domain in the form of generic models, which specify premises and conclusions for a set of interrelated Arguments. Such models might include econometric models, crate-ology models, or more qualitative sets of beliefs in such areas as political forecasting. In a new problem, the analyst would simply select a relevant model or set of models from a pre-existing library. These model templates would themselves be subject to augmentation, reevaluation, and revision in each new problem where they were applied.

- *Knowledge-prompting*. SED might interactively support the process of building Arguments by suggesting potentially relevant Premises (and even degrees of Support) based on pre-stored knowledge. Such pre-stored knowledge might pertain to the Topic, the Question, the source of information, the type of source, the age of the information, or any combination of the above. For example, SED would provide a set of candidate background Premises pertinent to the evaluation of a human source, a technical source, an open-literature source, etc. In some cases, SED might have pre-stored information regarding reliability of an individual source. In somewhat more complex cases, SED might store information about the reliability of a source-type or individual source with respect to a particular Topic and/or Question, or the reliability of a source-type on a Topic/Question when the information is of a certain age. Suggested Premises would be displayed for the user subject to his accepting, rejecting, or revising them.

- *Algorithms.* A useful enhancement would involve the ability to represent Arguments as numerical functions of evidence.

- *Alternative uncertainty calculi.* Another useful enhancement would involve the ability to utilize diverse representations of uncertainty, e.g., Bayesian probabilities and fuzzy logic, as well as Shaferian belief functions.

- *Causal modeling.* SED could permit explicit causal modeling of the link between evidence and conclusions, together with the potential disrupting factors at each stage. Such causal modeling would enhance the Argument construction process by stimulating the generation of background Premises. It would also permit representation of interdependencies among the impacts of different background Premises in calculating the Revised Position of an Argument.

- *Learning.* At present, SED "learns" by dropping Assumptions when they lead to conflict with other beliefs or Assumptions. This process could be supplemented by increasing the degree of firm belief when Assumptions are corroborated rather than contradicted by other lines of reasoning.

- *Information-collection options.* When an Assumption leads to conflict, it may often be possible to collect additional data to confirm or disconfirm the Assumption. SED could be augmented to help users evaluate such information collection options in terms of their potential costs and benefits.

Interface

- *User guides.* Users have complete freedom in how they utilize SED. Optional guides could be introduced, however, which direct them through specific sequences of SED operations associated with certain common functions: e.g., entering a new piece of data and looking for potential implications for previous conclusions; starting with a particular hypothesis and looking for evidence that supports or contradicts it; looking for all Arguments that depend on the credibility of a particular source; looking for all Arguments that involve a particular Topic, Question, or Assumption; etc.

- *Graphics*. At present, SED's user interface is largely alpha-numeric. Considerable enhancement in its usability could be achieved by introducing graphical representations. Such representations might include networks of Arguments showing the interdependencies among Issues and causal models linking evidence and conclusions in a particular Argument. More advanced graphical techniques could be explored to enable analyst's to navigate their way through an inferential space and to superimpose different Arguments upon one another in order to detect areas of agreement and conflict and to identify potential contributors to conflict.

Computational Efficiency

Current implementation evaluates all combinations (automatically created)

simplify by local ~~[illegible]~~ processing (Pearl; Shafer Shenoy)

supplemented by propagating assumptions only.

REFERENCES

Burrows, W.E. *Deep black*. New York, NY: Random House, 1986.

Cohen, M.S. An expert system framework for non-monotonic reasoning about probabilistic assumptions. In J.F. Lemmer and L.N. Kanal (Eds.) *Uncertainty in Artificial Intelligence* Amsterdam: North Holland Publishing Co., 1986.

Cohen, M.S. *Conflict resolution as a knowledge elicitation device*. Draft Technical Report. Reston, VA: Decision Science Consortium, Inc., 1989.

Cohen, M.S., Laskey, K.B., and Ulvila, J.W. *The management of uncertainty in intelligence data: A self-reconciling evidential database* (Technical Report 87-8). Falls Church, VA: Decision Science Consortium, Inc., June 1987.

Cohen, M.S., Schum, D.A., Freeling, A.N.S., and Chinnis, J.O., Jr. *On the art and science of hedging a conclusion: Alternative theories of uncertainty in intelligence analysis* (Technical Report 84-6). Falls Church, VA: Decision Science Consortium, Inc., August 1984.

Cohen, M.S., Watson, S.R., and Barrett, E. *Alternative theories of inference in expert systems for image analysis* (Technical Report 85-1). Falls Church, VA: Decision Science Consortium, Inc., January 1985.

DeJong, G., and Mooney, R. Explanation-based learning: An alternative view. *Machine Learning*, *1*, 145-176, 1986.

IPL/AMRD. Applications of inference theory and analysis of artificial intelligence to problems in intelligence analysis. *Intelligence Production Laboratory Conference Proceedings*. Rosslyn, VA: Analytic Methodology Research Division, Office of Research and Development, Central Intelligence Agency, November, 1982.

Laqueur, W. *A world of secrets*. New York, NY: Basic Books, Inc., 1985.

Laskey, K.B., and Lehner, P.E. Assumptions, beliefs and probabilities. *Artificial Intelligence*, in press.

Laskey, K.B. Belief in belief functions: An examination of Shafer's canonical examples, *Proceedings of the Third Workshop on Uncertainty in Artificial Intelligence*, July 1987.

Loui, R.P. Interval-based decisions for reasoning systems. In Kanal, L.N. and Lemmer, J.F. (Eds.), *Uncertainty in artificial intelligence*. The Netherlands: Elsevier Science Publishers B.V., 1986.

McDermott, D., and Doyle, J. Non-monotonic logic I. *Artificial Intelligence*, 1980, *13*, 41-72.

Nozick, R. *Philosophical explanations*. Cambridge, MA: Harvard University Press, 1981.

Pearl, J. Fusion, propagation and structuring in belief networks. *Artificial Intelligence*, 1986, *29*(3), 241-288.

Schum, D.A. Current developments in research on cascaded inference processes. In T.S. Wallsten (Ed.), *Cognitive process in choice and decision behavior*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1980.

Schum, D.A. Knowledge, probability, and credibility. *Journal of Behavioral Decision Making*, *2*, 39-62, 1989.

Shachter, R.D. Evaluating inference diagrams. *Operations Research* 34 (1986) 871-882.

Shafer, G. *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press, 1976.

Shafer, G. Constructive probability. *Synthese*, *48*, 1-60, 1981a.

Shafer, G. Jeffrey's rule of conditioning. *Philosophy of Science*, *48*, 337-362, 1981b.

Shope, R. *The analysis of knowing: A decade of research*. N.J., Princeton University Press, 1983.

Tolcott, M.A., Marvin, F.F., and Lehner, P.E. Decision making in evolving situations. *IEEE Transactions on Systems, Man and Cybernetics*, in press.

Zadeh, L.A. Review of Shafer's, A mathematical theory of evidence. *AI Magazine*, 1984, *5*(3), 81-83.

APPENDIX A: NON-MONOTONIC PROBABILIST

The non-monotonic probabilist (NMP) reasons with numerical degrees of belief, but in addition can represent the degree of shiftability of its own arguments in response to unexpected or conflicting evidence. NMP inference is built around a schema for representing an evidential argument. The argument schema makes explicit the background context within which an inference rule is valid. This supports the ability to call into question and revise background assumptions when they no longer appear to be valid.

NMP arguments are combined and chained together using a Shafer-Dempster belief calculus embedded within a process of default reasoning applied to the beliefs themselves. Nonindependencies due to shared premises are automatically accounted for in the belief calculations. Default reasoning serves to control the application of the belief calculus. Its role is to keep track of assumptions and to direct the process of belief revision when those assumptions lead to anomalous results.

A.1 NMP Argument Structure

Arguments in NMP are represented by an argument schema based on the one developed by Toulmin et al. (1984). In Toulmin's schema (Figure A-1), a claim, or conclusion whose merits we are seeking to establish, is supported by grounds, or evidence. The basis of this support is the existence of a warrant that states the general connection between grounds and claim. The warrant might for example be a general rule that this type of ground provides basis for this type of claim. The backing provides an explanation of why the warrant is regarded as reliable. That is, it provides theoretical or empirical evidence for the existence of an evidential relation or causal connection between grounds and claim. Modal qualifiers (e.g., probably; possibly) weaken or strengthen the validity of the claim. Possible rebuttals deactivate the link between grounds and claim by asserting conditions under which the warrant is invalid. A way of reading this structure is: Grounds, so Qualified Claim, unless Rebuttal, since Warrant, on account of Backing.

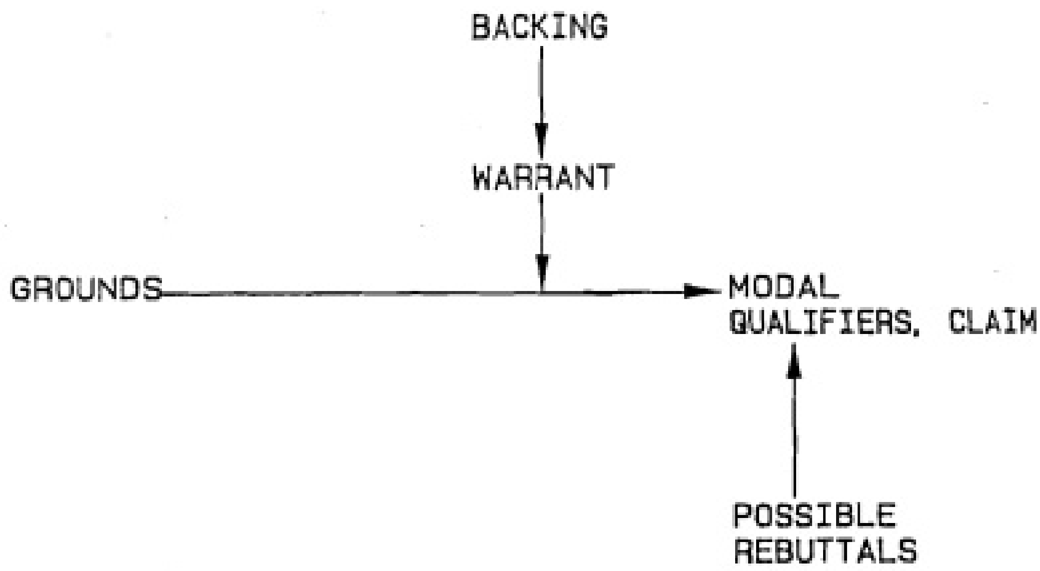


FIGURE A-1: TOULMIN'S STRUCTURE FOR AN EVIDENTIAL ARGUMENT

Figure A-2 shows how Toulmin's argument schema has been applied in the context of NMP. An argument from evidence to a conclusion is constructed using as a warrant a rule asserting that an evidential link exists between them. This rule may in turn be backed by a deeper theoretical or causal model, such as a general law or a statistical analysis. The evidential argument may be invalidated if any assumptions underlying the model do not hold.

A.2 The Belief Calculus

NMP arguments are combined and chained together using a Shafer-Dempster belief calculus embedded within a process of default reasoning applied to the beliefs themselves. Shafer's theory was chosen for our implementation because of several features that make it amenable to an intelligent control and belief revision capability:

- *Representing evidential incompleteness.* Usually in intelligence problems our evidence is incomplete. According to Shafer (Shafer and Tversky, 1985), the contrast between belief functions and probabilities focuses directly on this idea of incompleteness of evidence. While the probability of a hypothesis measures the chance that it is true conditional on given evidence, its Shafer-Dempster belief measures the degree to which the evidence *means* (or proves) that it is true (see also Pearl, 1988, chapter 9). By stressing the link between evidence and hypothesis, Shafer's theory is able to provide an explicit measure of the quality of evidence or degree of ignorance.

- *Diagnosis of conflict.* To the extent that two arguments support incompatible hypotheses, combining beliefs by Dempster's Rule creates support for the null set. This support is then removed by proportionately increasing support for all non-null sets. But null set support serves a useful function for NMP. It measures the degree to which propositions are inconsistent, and thus constitutes a natural measure of conflict in the evidence.

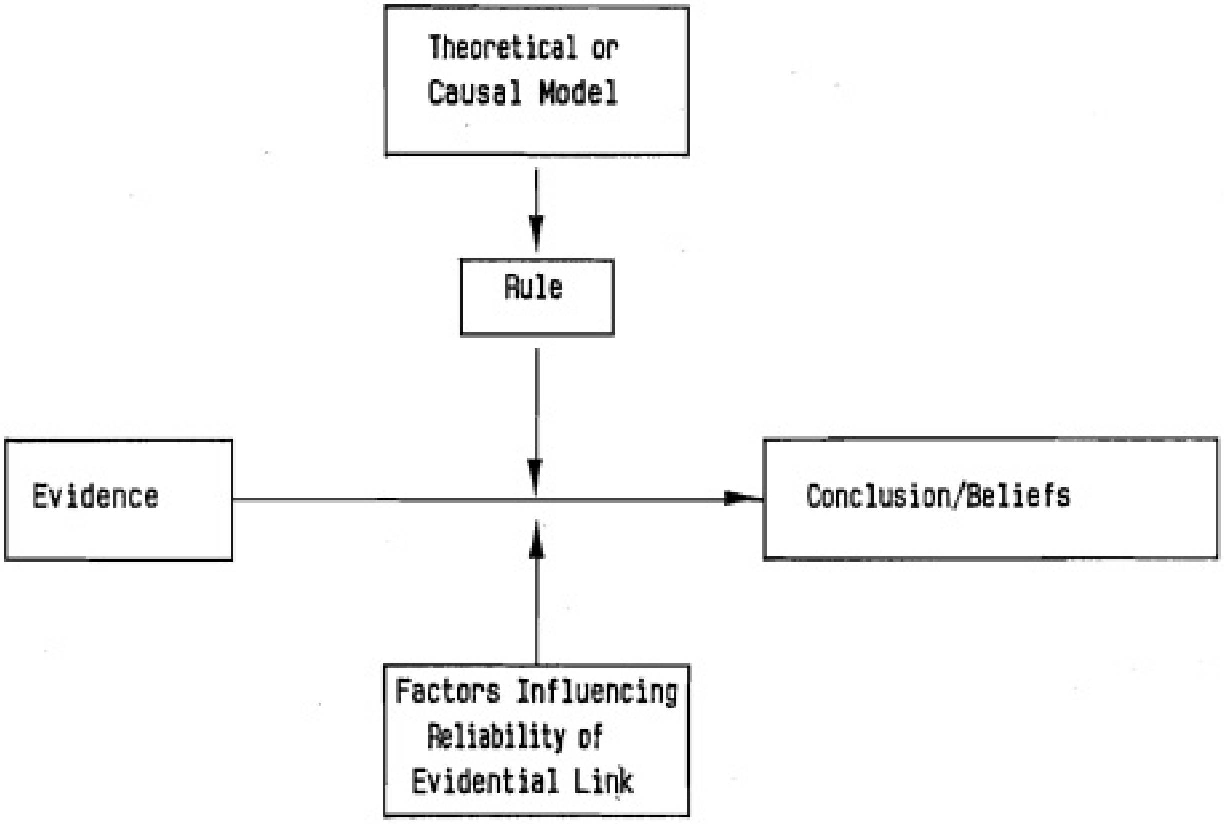


FIGURE A-2: ARGUMENT STRUCTURE IN NMP

- *Assumptions*. To the degree that current evidence is uncommitted with regard to the truth or falsity of a hypothesis, there is room for assumptions. An assumption could be naturally represented in Shafer's framework as a decision regarding the allocation of uncommitted belief. Such a decision, by definition, goes beyond the evidence, but remains within the constraints of the evidence.

- *Discrediting arguments*. The outcome of a process of conflict resolution may be the discrediting of one or more lines of reasoning that led to the conflict, by rejecting assumptions involved in those arguments. Partial or complete rejection of an assumption is represented by decreasing, possibly to zero, the degree to which uncommitted belief is allocated to the formerly assumed hypothesis.

Shafer himself does not address the notion of an assumption, as just outlined. Indeed, actions in response to conflict, such as re-examining source credibility, must occur outside the theoretical structure of belief functions. Non-monotonic probabilist embeds a belief function model within a qualitative assumption-based reasoning process. This qualitative reasoning process uses the tools implicit within Shafer's calculus to formalize and direct an iterative conflict resolution and assumption revision process.

A.3 Belief Propagation

Non-Monotonic Probabilist uses the belief maintenance system (BMS) (Laskey and Lehner, 1988) to compute beliefs and keep track of assumptions. The BMS represents belief functions as tokens attached to rules linking evidence and conclusions. Stored with each token is probabilistic information about the strength of the evidential link it represents. A probability calculus on belief tokens is formally equivalent to a Shafer-Dempster calculus or conclusions (Laskey and Lehner, 1988; in press). An explicit provision for making and revising assumptions has been added to complete the machinery necessary for implementing NMP.

Belief maintenance combines deKleer's (1986a,b,c) assumption-based truth maintenance system (ATMS) with a module for representing and reasoning with degrees of belief on symbolic tokens manipulated by the ATMS. DeKleer argues for explicit separation of reasoning into two functions, problem solving and truth maintenance. In NMP, the *belief* maintenance system performs a role analogous to the role deKleer proposes for *truth* maintenance. The BMS keeps account of assumptions, computes beliefs, determines the degree of conflict, and attributes that conflict to specific assumptions. It therefore supports a set of basic functions necessary for conflict resolution. The SED front end to NMP corresponds to deKleer's problem solver. The evidential reasoner constructs arguments symbolically, and passes to the BMS the task of computing beliefs and conflict.

Two features of the ATMS make it well-suited to its role as the substrate for belief maintenance. First, it is designed to be able to maintain belief simultaneously for multiple and possibly inconsistent propositions, a capability required for reasoning with numerical beliefs. Second, the design of the ATMS maintains an explicit separation between problem solving and truth maintenance. In our terms, this means that high-level reasoning about the application of the inference mechanism is explicitly separated from (although informed by) the process of keeping track of assumptions and computing beliefs. Belief maintenance is capable of representing the full generality of the Shafer-Dempster calculus. The ATMS automatically keeps account, in symbolic form, of the propagation of beliefs through chains of inference, nonindependencies created through shared premises, and inconsistent combinations of tokens. The belief computation module incorporates all this information to compute correct Shafer-Dempster beliefs when requested. Adding to this framework the capability to make and reason with default assumptions results in a fully integrated symbolic and numeric uncertainty management framework. This framework is well suited to qualitative reasoning about the application of a numeric uncertainty calculus.

A general formal presentation of how assumption-based truth maintenance can be used to encode and reason with belief functions is given in Laskey and Lehner (in press). In Laskey and Lehner (1988), this framework is extended to allow making and revising default assumptions.

An NMP rule has the general form:

```
(IF <antecedents> THEN <consequent>
                    PROVIDED <background antecedents>) .
```

Typically, the effect of the background context is summarized by a numerical belief value, representing the degree to which the evidence is taken to imply the conclusion. Consider, for example, a problem in which an analyst wishes to reason about the output of an image processing system. An NMP rule for this problem might be:

```
R38: (IF (Template_Matcher_Report Cluster Maneuver_Company))
          THEN (ID Cluster Maneuver_Company)
          PROVIDED (RULE-VALID-R38)),
```

which states that if the template matcher identifies a cluster as a maneuver company, then it is believed to be a maneuver company, provided RULE-VALID-R38. The symbol RULE-VALID-R38 represents a *belief token*, a special construct within the BMS that carries an attached probability. For example, if the assigned probability is .8, then the report will cause the ID of the cluster to be assigned .8 belief in Maneuver_Company, absent other evidence. The probability of the belief token RULE-VALID-R38 may be interpreted as the probability that the rule is "working". That is, this probability summarizes our belief that some condition disabling the rule has not occurred.

The ATMS propagates tokens, including belief tokens, through chains of argument. It maintains a *label* for each proposition in its database, which represents the token sets that are sufficient to prove the proposition. In the above example, after receiving the report (Template_Matcher_Report Cluster Maneuver_Company) the label of (ID cluster Maneuver-Company) would be:

```
(ID Cluster Maneuver_Company):  {RULE-VALID-R38} .
```

The ATMS can chain arguments together, and form multiple arguments for the same conclusion. For example, suppose we also had the label:

(In_Maneuver_Battalion Cluster): {RULE-VALID-R19},

as the result of firing another rule. Firing the rule

```
R30: (IF (In_Maneuver_Battalion Cluster) THEN (ID Cluster Manuever_Company)
                                         PROVIDED (RULE-VALID-R30))
```

changes the label of (ID cluster Maneuver_Company) to:

```
     (ID cluster Maneuver_Company):
          {RULE-VALID-R38}, {RULE-VALID-R19,RULE-VALID-R30} .
```

This means that the cluster can be proven to be a maneuver company if RULE-VALID-R38 is true (i.e., Rule 38 is "working"), or if RULE-VALID-R19 *and* RULE-VALID-R30 are both true.

The probability of a proposition's label is the probability that the proposition can be proven--that is, its Shafer-Dempster belief. In our example, to find the degree of belief in (ID Cluster Maneuver-Company), we need to find the probability of (RULE-VALID-R38 *or* (RULE-VALID-R19 *and* RULE-VALID-R30)). To do this, the probability calculator module of the BMS constructs a "truth table" representing all possible truth values of the belief tokens in the proposition's label. The probability of the label is then the probability of the rows in the truth table that imply the label (i.e. in which RULE-VALID-R38 is true or RULE-VALID-R19 and RULE-VALID-R30 are both true). Figure A-3 shows how this is done for this example, assuming beliefs .8, .7, and .9 for RULE-VALID-R38, RULE-VALID-R19, and RULE-VALID-R30, respectively.

A.4 Assumptions in NMP

As noted above, the ability to make and revise assumptions was an important design criterion for NMP. Often we wish the system to assume a high belief for a rule unless there is direct evidence to the contrary, even if this high belief is not directly justified by the evidence.

RULE-VALID-R38	RULE-VALID-R19	RULE-VALID-R30	Conclusion	Belief
T	T	T	T	.50
T	T	F	T	.06
T	F	T	T	.13
T	F	F	T	.01
F	T	T	T	.22
F	T	F	?	.02
F	F	T	?	.05
F	F	F	?	.01

RULE-VALID-R38: P(T) = .7; P(F) = .3
RULE-VALID-R19: P(T) = .8; P(F) = .2
RULE-VALID-R30: P(T) = .9; P(F) = .1

Belief in (ID Cluster Maneuver_Company) = .92

Figure A-3: Truth Table for (ID Cluster Maneuver_Company)

For example, consider the rule:

```
R39: (IF ((Template_Matcher_Report Cluster SAM)) THEN (ID Cluster SAM)
                                                 PROVIDED (RULE-VALID-R39)).
```

Suppose we felt justified in assigning belief of only .5 to this rule, but we wished to increase this belief by *assumption* to .9. We would then allocate 80% of the uncommitted belief of .5 to the hypothesis (ID Cluster SAM). Assumptions are implemented within NMP by assigning *default tokens*, or special tokens that are treated as if they had probability 1. To make such an assumption in NMP, we would use two rules. First would be Rule 39 above. Second, the system would encode the following rule:

```
R40: (IF ((Template_Matcher_Report Cluster SAM))
      THEN (ID Cluster SAM)
      PROVIDED (NOT-RULE-VALID-R39 RULE-VALID-R40 ASSUME-R40)) .
```

The token ASSUME-R40 is a *default token*, which is treated as if it has probability 1 until the system encounters evidence that makes it question its original assumption. The token NOT-RULE-VALID-R39 is the negation of RULE-VALID-R39 (and thus has belief .5). Its probability represents belief that the rule is not valid--i.e., there is no link between the template match and the ID of the cluster. The token RULE-VALID-R40 has belief .8.

When the template matcher report is logged and Rules 39 and 40 fire, we obtain the label:

```
    (ID Cluster SAM):  {RULE-VALID-R39}, {NOT-RULE-VALID-R39,
                       RULE-VALID-R40,ASSUME-R40}
```

The environment {NOT-RULE-VALID-R39,RULE-VALID-R40,ASSUME-R40} has belief $.5 \times .8 \times 1.0 = .4$, and so .4 is added to our belief in (ID Cluster SAM), as shown in Figure A-4.

RULE-VALID-R39	RULE-VALID-R40	(ID Cluster SAM)	Belief
T	T	T	.40
T	F	T	.10
F	T	T	.40
F	F	?	.10

RULE-VALID-R39: P(T) = .5; P(F) = .5
RULE-VALID-R40: P(T) = .8; P(F) = .2

Belief in (ID Cluster SAM) = .9

Figure A-4: Truth Table for Belief Allocation to SAM

A.5 Conflict Resolution in NMP

Representing assumptions explicitly is useful because the system can examine them and revise them when necessary. In NMP, assumptions may be revised in response to *conflict*. Conflict occurs when arguments support contradictory conclusions.

Let us consider an example. Suppose the system had a default rule stating the system's belief that no SAM emissions are emanating from an area if there is no specific evidence of emissions:

(In this case, *all* the belief has been allocated by assumption to the conclusion).

```
R03: (IF () THEN (not (SAM_Emissions_Near (Loc Cluster)))
            PROVIDED (ASSUME-R03))
```

This produces the label:

```
(not (SAM_Emissions_Near (Loc Cluster))):  {ASSUME-R03}
```

Now suppose we have another rule:

```
R25: (IF ((ID Cluster SAM)) THEN (SAM_Emissions_Near (Loc Cluster))
                              PROVIDED (RULE-VALID-R25 ASSUME-R25))
```

After firing R40 as described in Section 3.2.4, firing this rule results in the label:

```
(SAM_Emissions_Near (Loc Cluster)):
     {RULE-VALID-39,RULE-VALID-R25,ASSUME-R25},
     {NOT-RULE-VALID-R39,RULE-VALID-R40,RULE-VALID-R25,ASSUME-R40,
     ASSUME-R25}
```

Because the system knows that (SAM_Emissions_Near (Loc Cluster)) and its negation are inconsistent, it creates a *nogood* environment by combining their labels:

nogood {RULE-VALID-R39,RULE-VALID-R25,ASSUME-R25,ASSUME-R03}
{NOT-RULE-VALID-R39, RULE-VALID-R40, RULE-VALID-R25,
ASSUME-R40, ASSUME-R25, ASSUME-R03}.

Nogood environments are sets of assumptions that cannot all be true (note that if all the above tokens were true we could derive both (SAM_Emissions_Near (Loc Cluster)) and its negation).

Figure A-5 illustrates the belief computations for this example. Note the high degree of belief assigned to inconsistent sets, or the contradiction $\perp$. (Rows of the truth table are marked contradictory if, coupled with the current defaults, they are *nogood*). The degree of belief assigned to $\perp$ by the belief calculator algorithm is the *conflict* associated with the hypotheses (SAM_Emissions_Near (Loc Cluster)) and (not (SAM_Emissions_Near (Loc Cluster)). When this number gets large, the system examines the assumptions contributing to the conflict for possible revision.

In our example, revising any of the three assumptions (ASSUME-R40, ASSUME-R25, or ASSUME-R03) would reduce the conflict. The final beliefs the system is left with, however, depend critically on which is revised. Dropping either of the assumptions ASSUME-R40 or ASSUME-R25 would disrupt the chain of evidence leading to the conclusion (SAM_Emissions_Near (Loc Cluster)). Removing the assumption ASSUME-R03 removes the argument for (not (SAM_Emissions_Near (Loc Cluster)), leaving its belief equal to zero. Belief in (SAM_Emissions_Near (Loc Cluster)) is then given by the analysis in Figure A-6.

RULE-VALID-R39	RULE-VALID-R40	RULE-VALID-R25	(SAM_Emissions _Near (Loc Cluster))	Belief
T	T	T	⊥	.28
T	T	F	F	.12
T	F	T	⊥	.07
T	F	F	F	.03
F	T	T	⊥	.28
F	T	F	F	.12
F	F	T	F	.07
F	F	F	F	.03

RULE-VALID-R39: P(T) = .5; P(F) = .5
RULE-VALID-R40: P(T) = .8; P(F) = .2
RULE-VALID-R25: P(T) = .7; P(F) = .3

Belief in (SAM_Emissions_Near (LOC Cluster)) = 0
Belief in (not(SAM_Emissions_Near (LOC Cluster))) = .37

Conflict = .63

Figure A-5: Example of Beliefs When Evidence Conflicts

RULE-VALID-R39	RULE-VALID-R40	RULE-VALID-R25	(SAM_Emissions _Near (Loc Cluster))	Belief
T	T	T	T	.28
T	T	F	?	.12
T	F	T	T	.07
T	F	F	?	.03
F	T	T	T	.28
F	T	F	?	.12
F	F	T	?	.07
F	F	F	?	.03

RULE-VALID-R39: P(T) = .5; P(F) = .5
RULE-VALID-R40: P(T) = .8; P(F) = .2
RULE-VALID-R25: P(T) = .7; P(F) = .3

Belief in (SAM_Emissions_Near (LOC Cluster)) = .63
Belief in (not(SAM_Emissions_Near (LOC Cluster))) = 0

Conflict = 0

Figure A-6: Belief in (SAM_Emissions_Near (LOC Cluster))
After Dropping ASSUME-R03

APPENDIX B: SOFTWARE DESCRIPTION

B.1 Overview of SED Architecture

B.1.1 Software architecture. The high level architecture of SED is represented in the data flow diagram of Figure 1.

The software consists of two parts: the SED.EXE executable, written to use dBASE III+ files in the CLIPPER dialect of DBASEIII+; and the BMS (Belief Maintenance System), written in Gold Hill Common LISP. Due to the LISP severe environment requirements (it takes about 5 minutes to COLD BOOT), it was chosen as the underlying operating system framework. As a result, SED.EXE only exists when called, and had to be developed with an internal checkpointing system. The two interface files: bms_in.dat, LISP code generated by SED.EXE to feed the BMS; and bms_out.dat, a table of net supports by topic-question; are temporary and have meaning only while SED is running. Analysts interact with the SED.EXE portion of SED.

A more detailed view of SED is included in Figure 2. It consists of 10 primary modules:

SED	-	(re)starts executable and reads BMS_OUT.DAT
LOGIN	-	allows user to tag judgments
WORKLIST	-	allows the user to organize work to support analytical conclusions
REPORT	-	allows user to browse and annotate field reports
ARGLIST	-	allows user to build, modify, and view grounds and positions
CONFLICT	-	allows user to view CONFLICT and change assumptions
CALCULATE	-	constructs BMS_IN.DAT
CALLER	-	allows Function Key navigation between modules
EXIT	-	ends a session
CHECKPOINT	-	saves internal state for SED

B.1.2 Interface with BMS. As previously explained, the Belief Maintenance System (BMS) calculates the net belief, support and plausibility. It is covered in detail (REF) elsewhere in the report. It takes as input a file of stylized LISP code which identifies the hypotheses, arguments, and desired results of the calculations. The functions necessary to do this are:

```
(defhypset (...))
(defprobdist (...))

(defrule (...))
(compute-beliefs (...))
(compute conflict (...))
(assume (...))
(default-node (...))
```

and to terminate:

```
(end-execution)
```

It returns calculations of conclusion information for display in the following tabular form:

```
hyp1      bel1      pl1      confl1
hyp2      bel2      pl2      confl2
                .
                .
                .
hypn      beln      pln      confln

default-node-name1      conflict-attributed-to-it1
                    .
                    .
                    .
default-node-namek      conflict-attributed-to-itk
```

B.1.3 <u>System engineering requirements</u>. The following represents a minimal configuration.

1. <u>Host/CPU</u>: PC/AT or better machine capable of running Gold Hill Common LISP.

2. <u>Memory</u>: 8 MB-main memory
20 MB-hard disk

3. <u>Graphics</u>: Color Graphics Adapter (CGA) board and color Monitor

4. <u>Software</u>: Gold Hill Common LISP, Gold Hill Corporation
MS-DOS 3.3, Microsoft
SED.EXE*

   CLIPPER compiler, Nantucket Software, is needed to extend the basic system

5. <u>Database Files</u>: ARGUMENT.DBF
NETSUPPT.DBF
PREMISES.DBF
REPORTS.DBF
REPTOPS.DBF
TOPICS.DBF
USERIDS.DBF
REPORTS.DBT

B.2 Database Design

The database design consists of 8 SED.EXE system persistent files. These files represent all of the work done and should always exist with the system. The .DBF files are standard dBASE III+ files, and the .DBT is a memofield file (free text) associated with reports. The fields are shown in the next table. The relations are shown in Figures 3, 4, and 5. Lastly, the checkpoint file (CHECK.FIL) is shown.

This database is in *3rd normal form* and assumes that only 5 values may be assigned to any topic question, otherwise it is fully general. The 5 value decision puts a top on the set theoretic maximum of combinations.

B.2.1 dBASE files.

Database Structure:	Field:	Field Name:	Type:	Width:	Dec:
C:userids.dbf	1	USERID	Numeric	2	
	2	USERNAME	Character	10	
	3	FNAME	Character	10	
			Total	23	
C:argument.dbf	1	ARGID	Numeric	5	
	2	SEQUENCE	Numeric	2	
	3	CORE	Logical	1	
	4	USERID	Numeric	2	
	5	TQID	Numeric	5	
	6	IS_1	Logical	1	
	7	IS_2	Logical	1	
	8	IS_3	Logical	1	

Database Structure:	Field:	Field Name:	Type:	Width:	Dec:
C:argument.dbf	9	IS_4	Logical	1	
(cont'd)	10	IS_5	Logical	1	
	11	BASICSUPPT	Numeric	4	2
	12	BASICBLIEF	Numeric	4	2
	13	BASICPLAUS	Numeric	4	2
	14	CONFLICTAD	Numeric	4	2
	15	UNCOMMITTD	Numeric	4	2
	16	WELLFORMED	Logical	1	
			Total	42	
C:netsuppt.dbf	1	USERID	Numeric	5	
	2	TQID	Numeric	5	
	3	IS_1	Logical	1	
	4	IS_2	Logical	1	
	5	IS_3	Logical	1	
	6	IS_4	Logical	1	
	7	IS_5	Logical	1	
	8	NETSUPPORT	Numeric	4	2
	9	NETBELIEF	Numeric	4	2
	10	NETPLAUSIB	Numeric	4	2
	11	ASSUMEDPCT	Numeric	4	2
			Total	32	
C:premises.dbf	1	ARGID	Numeric	5	
	2	SEQUENCE	Numeric	2	
	3	TQID	Numeric	5	
	4	IS_1	Logical	1	
	5	IS_2	Logical	1	
	6	IS_3	Logical	1	
	7	IS_4	Logical	1	
	8	IS_5	Logical	1	
			Total	18	
c:topics.dbf	1	TQID	Numeric	5	
	2	TOPIC	Character	20	
	3	QUESTION	Character	50	
	4	CONFLICT	Numeric	4	2
	5	ANSWER1	Character	20	
	6	ANSWER2	Character	20	
	7	ANSWER3	Character	20	
	8	ANSWER4	Character	20	
	9	ANSWER5	Character	20	
			Total	180	
c:reports.dbf	1	REPORTID	Numeric	5	
	2	DTG	Date	8	
	3	SOURCE	Character	20	
	4	TEXT	Memo	10	
			Total	44	

Database Structure:	Field:	Field Name:	Type:	Width:	Dec:
c:reptops.dbf	1	REPORTID	Numeric	5	
	2	TQID	Numeric	5	
			Total	10	

B.2.2 Database diagrams.

Argument Relations (see Figure 3)

Net Support Relations (see Figure 4)

Report Relations (see Figure 5)

B.2.3 Ancillary files.

CHECK.FIL

An ASCII string: $a_1 \ldots a_7$. navigation

a_1 = USERID record number
a_2 = ARGUMENT record number
a_3 = NETSUPPORT record number
a_4 = PREMISES record number
a_5 = TOPICS record number
a_6 = REPORTS record number
a_7 = REPTOPS record number

navigation = name of module (e.g., "WORKLIST")

BMS_IN and BMS_OUT were reviewed in the previous section.

B.3 Alphabetical List of Software Units

B.3.1 Utilities:

clear_gets - clears pending screen reads.

ls - a short hand routine for LTRIM(STR()).

pad - a way to pad a string with blanks.

say_at - a way to display numbers with a value of 0.00 as " . ".

B.3.2 <u>SED specific routines</u>:

a_1, a_2, a_3, a_4, a_5 - a brute force answer checklist.

a_belief - gets belief values from 0.0 to 1.0 from BMS_OUT.DAT.

a_conf - gets conflict values from 0.0 to 1.0 from BMS_OUT.DAT.

a_plaus - gets plausibility values from 0.0 to 1.0 from BMS_OUT.DAT.
Calls: MEMOTRAN

a_star - tests if input is a "*".

a_tqid - parses incoming string, whether it is a token like tqtext makes.

add_it - signals the user's desire to add an argument, premise, or sequence number.

arglist - provides GROUNDS, ARGUMENT, AND CONSEQUENCES functions.
Calls: menustuf, getarglist, displaycyc, selecttop, make_net, showclause, menuselect, edittop, editanswrs, cleanup

assignblf - allows the user to modify basic support.
Calls: displaycyc, check, calcbelief, calcplaus, uncommittd

assume - allows the user to assign assumption to a tq answer set.

atext - constructs an answer suffix.

b_str - constructs a belief token.

bms_out - reads BMS_OUT.DAT.

browse5 - allows the user to browse TOPICS.DBF.

browse11 - allows the user to browse the REPORTS.DBF.
Calls: xbrowse

calc - modifies the navigation string (the routing address for caller).

calcbelief - sums the supports for all topic/question answer sets of an argument.

calcplaus - calculates the plausibility for a topic/question answer set.

calculate - constructs the entire BMS_IN.DAT instruction string to BMS.
Calls: stringset, +b_str, b_str, superset, propersub

caller - provides the SED software bus. Allows the high level modules to call each other.
Calls: login, help, worklist, reports, arglist, conflict, calc, exit

check - does range checking on belief entries.

checkpoint - does checkpointing.

chk_blf - sums belief to see if it exceeds 1.0.

cleanup - general cleanup routine, to keep data structures packed and clean.

conflict - displays conclusion conflict.
Calls: menustuf, say_heading, say_topic, menuselect

displaycalc - allows the user to cycle through (and view) supports, beliefs, and plausibility.

dont_show - keeps the user from using/seeing already used position TQs or premise TQs.

edit_tops - allows the user to add and update TQs.
Calls: menuselect, edittopic

editanswer - allows editing of answer text, but does not allow expansion of answer sets that have arguments.
Calls: showanswer

editanswrs - edits a topic/question's answer set.
Calls: a_star

edittext - allows the user to change the report message text.

edittopic - edit a topic/question and answer set in TOPICS.DBF.

finals - calculates the final support for a position.
Calls: uncommittd

finishcore - sets up a NETSUPPT.DBF record.

get_pic - gets "PICTURE" info about a field.

get_recs - uses db_ids array to retrieve record.
Calls: checkpoint, caller

getarglist - counts up the number of tqid's that match a search criterion and puts the arguments ids (argids's) into an array.

help - a stubbed version of a context sensitive help system.

hypotheses - constructs a hypotheses string for the BMS.
Calls: tqtext, atext

insert_str - inserts a tqtext string into another string (se tqtext).

login - allows the user into the system.

make_net - makes a new entry into NETSUPPT.DBF.

menuselect - allows the system to check function key presses against the current Menu.

menustuf - shows the function keys and menus.

newtopic - add a topic/question and answer set to TOPICS.DBF.

next_id - appends a blank record and gives it a sequential number.

propersub - returns true if one string is a propersub set of another.

reports - provides the REPORTS screen functionality.
Calls: menustuf, showtopics, showtext, brosell, edittext, edit_tops

save_recs - maintains the db_ids array.

say_answer - shows the user the answers for a TQ.

say_belief - shows BMS conclusions to the user.

say-deviat - shows the "circles" in arglist related to deviations.

say_final - shows complete information for an argument.

say_heading - provides labels for say_belief.

say_net - show the results of the BMS calculations.

say_topic - shows a TQ answer set.
Calls: say_answer

selectans - select the answers which are applicable to a specific position.
Calls: a_star

selecttop - allows the user to select a position or a premise topic/question and answer set.
Calls: dont_show, say_topic, displaycyc, menuselect, next_id, assignblf, make_net, selectans, newtopic, showclause

showanswer - shows the answer set and belief.
Calls: say_answer, say_heading, say_ansset, saybelief

showclause - the central display unit in arglist. It shows all TQ answer sets.
Calls: say_topic, say_ansset, say_deviat, say_final, say_net

showtext - show the free text of a report message.

showtopics - displays the TQs.

statline - shows record status infor for xbrowse.

stringset - constructs a string representing the T-Q answer set for BMS.
Calls: tqtext, atext

superset - constructs the UNION of two stringsets.

tqtext - constructs a tqtext symbol for the BMS.

uncommittd - calculates the sum of all support for proper supersets of a T-Q answer set.

sed.prg - main routine which establishes most of the globals and conducts checkpointing.

worklist - provides the ISSUES screen functionality.
Calls: showanswer, browse5, editanswer, assume

xbrowse - handles/traps keys for brows* routines
Calls: statline, menuselect

B.3.3 <u>Ancillary</u>:

structur.prg - lists the db file structure.