MULTISTAGE INFERENCE MODELS FOR INTELLIGENCE ANALYSIS

Edgar M. Johnson, et al

Army Research Institute for the Behavioral and Social Sciences
Arlington, Virginia

1974
Intelligence analysis is an intellectually demanding task requiring both inductive and deductive reasoning, hypothesis generation and hypothesis testing, imagination and the ability to attend to meticulous details. The objective of intelligence analysis is to impose a consistent, coherent (and hopefully correct) interpretation upon apparently unrelated bits and pieces of information—to transform information into intelligence.

Virtually nothing has been written on the "how" of intelligence analysis, on the appropriate procedures to utilize in any given situation. With the tremendous increase in the quantity of battlefield information collected, increased demands are placed on the intelligence analyst. The Army has recognized the potential of computers to provide needed support for the analyst and the planning for such support has highlighted the need for a better understanding of the process of intelligence analysis. With the introduction of computer systems such as the Tactical Operations System, the power of intelligence analysis will be primarily limited by the availability of procedures which will allow maximum use of the system capabilities. This has stimulated an examination of both the current processes used in intelligence analysis and a search for new techniques.

Multistage inference models provide a potentially meaningful and useful framework for the analysis of current modes of intelligence processing as well as a basis for the development of processing aids. This paper describes the development and evaluation of a class of multistage Bayesian inference models. The derivation and nature of the formal models are described in the following section. In the
third section of the paper, the results of several experiments concerning human performance in multistage inference tasks and their implications are discussed. The final section provides a brief summary.

MULTISTAGE INFERENCE MODELS

When there is a large logical gap between the data considered by an inference maker and the hypothesis set of interest, the process can often be viewed as one of multistage Bayesian inference. That is, the inference consists of a series of single-stage Bayesian inferences in which the output at each stage becomes the input to the next stage and so on, until a terminal hypothesis set is reached. Inferences involving a staging or cascading of simple inference tasks are found in a wide variety of settings—oil and gas exploration; investment decisions; medical diagnosis; and intelligence analysis. An intelligence analyst evaluating enemy activity to diagnose enemy capabilities or alternative courses of action can be viewed as a multistage inference maker. He may evaluate enemy activity in the framework, for example, of the enemy’s striking force posture (stage 1) and then evaluate the enemy’s striking force posture in the framework of indications (stage 2), and as a final step revise his estimate of the relative likelihood of alternative enemy courses of action (stage 3).

The processing or inference task performed by the intelligence analyst is often formally analogous to a problem in statistical inference; items of evidence, data, are used to determine the relative likelihood of alternative hypotheses. An optimal strategy for processing data in single-stage tasks is Bayes' theorem, one form of which is:

\[ P(H_1/D) = \frac{P(D/H_1) P(H_1)}{\sum_i P(D/H_i) P(H_i)} \]

where \( P(H_1) \) is the prior probability of a particular hypothesis, \( H_1 \); \( P(D/H_1) \) is the probability of the occurrence of a particular item of data, \( D \), conditional upon the truth of \( H_1 \); \( P(H_1/D) \) is the posterior probability of \( H_1 \) conditional upon the occurrence of \( D \). Expressed in this way, the estimation of posterior probability is seen to involve two processes: first, the evaluation of the diagnostic impact of each datum, \( P(D/H_1) \); and second, the estimation of the posterior probability, \( P(H_1/D) \), on the basis of the observed data.

The traditional design philosophy for information processing systems identifies the role of man as one of both processing the
information and of making inferences; machines perform only unburdening functions such as display and information storage. Research comparing intuitive inference performance with Bayes' theorem has shown that men are conservative information processors—they consistently fail to extract from data as much information as the data contain. Men are conservative not because they fail to properly evaluate the diagnostic impact of each datum, but rather because they fail to aggregate the data properly. As noted some years ago, this suggests a novel design for probabilistic information processing systems: one in which men evaluate/judge the data and machines perform the aggregation.

In applying this design philosophy to tactical intelligence information processing, it is clear that the evaluation of tactical data is a more complex and difficult task than indicated by laboratory studies. Human evaluation of uncertain data requires an understanding of the probabilistic linkages between the data and the hypotheses; requires an understanding of the processes by which the data is generated. In the absence of this knowledge, data evaluation becomes a complex, difficult and error-prone task. If the cognitive complexity of data evaluation can be reduced through the development of structured decision procedures, then man's performance can be improved.

Basic Model. Higher order systems are based on a decomposition of the conditional relationship linking the hypothesis set and the data set, $P(D/H_i)$. The customary notation of Bayes' theorem becomes cumbersome for the purposes of decomposition and matrix notation will be used as an operator symbolism. This notation is illustrated using the single-stage model Eq. (1). We form a column vector of the input diagnostic assessments, and a diagonal matrix from the prior probabilities for the hypothesis set $H$.

$$V_{D,H} = \begin{bmatrix} P(D/H_1) \\ P(D/H_2) \\ \vdots \end{bmatrix} \quad (2) \quad M_H = \begin{bmatrix} P(H_1) \\ \vdots \\ P(H_2) \end{bmatrix} \quad (3)$$

The terms appearing in the numerator of Bayes' law are the elements of the vector $M_H V_{D,H}$, formed by premultiplying the vector $V_{D,H}$ by the matrix $M_H$. The posterior probabilities $V_{H,D}$ are obtained by

$\text{Note:}$ The present discussion assumes a single item of data or a data set whose elements are processed sequentially. Rather than introduce an additional subscript, $D$ is used to refer to the datum or data of interest; the interpretation will be clear from the context.
dividing each element of the resultant vector \( V_{D,H} \) by the sum of its elements. This normalization process may be indicated functionally by defining an operator \( N_{C}(\cdot) \), which divides the column elements of its argument by the corresponding column sum. Thus, the vector of posterior probabilities for a single-stage Bayesian system may be expressed as

\[
V_{H,D} = N_{C}(M_{H} V_{D,H})
\]  

(4)

A third-order system will be used to illustrate the decomposition process. The relationship \( P(D/H_i) \) is decomposed by identifying intermediate hypothesis sets \( A_i \) and \( B_k \), each consisting of mutually exclusive and exhaustive hypotheses. These intermediate hypothesis sets can be directly incorporated into the relationship \( P(D/H_i) \) by applying probability theory to yield:

\[
P(D/H_i) = \sum_j \sum_k P(A_j/H_i) P(B_k/A_j H_i) P(D/B_k A_j H_i)
\]  

(5)

Assume that the elements of the logic chain are pairwise independent, viz.,

\[
P(D/B_k A_j H_i) = P(D/B_k) \quad P(B_k/A_j H_i) = P(B_k /A_j)
\]  

(6) \( (7)\)

Under this assumption, Eq. (5) reduces to the following:

\[
P(D/H_i) = \sum_j \sum_k P(A_j/H_i) P(B_k/A_j) P(D/B_k)
\]  

(8)

In operator notation, this can be expressed as,

\[
V_{D,H} = M_{A,H} M_{B,A} V_{D,B}
\]  

(9)

where \( M_{A,H} = P(A_j/H_i) \) and \( M_{B,A} = P(B_k/A_j) \). If the right side of Eq. (9) is substituted into Eq. (4), the result is a third-order multistage Bayesian system,

\[
V_{H,D} = N_{C}(M_{H} M_{A,H} M_{B,A} V_{D,B})
\]  

(10)

This expression can be viewed as a logic chain, \( D \rightarrow B \rightarrow A \rightarrow H \) in which data, \( D \), impact in turn on the hypothesis sets \( B, A, \) and finally \( H \). The logic chain is analogous to a weighted linear filter in which the weights are determined by the probabilistic links in the chain.
The process shown in Eq. (10) can be conveniently diagrammed to illustrate the flow of information (Figure 1). When applied sequentially to a data set, the posterior probabilities, $V_{H/D}$, at each step become the prior probability to be revised on the basis of the impact of the subsequent data.

Thus, in a multistage inference the evaluation of data has been simplified to a judgment of its diagnostic impact relative to a hypothesis set which logically closer to the data than in the corresponding single stage inference. For example, an evaluation of data relative to striking force posture, $P(D_B/B_k)$, may be used to replace an evaluation of the same data relative to alternative courses of action, $P(D_B/H_i)$. However, this simplification in data evaluation requires prior estimation of the conditional links between stages of the inference process; that is, estimates are required in order to fill in each of the intermediate matrices $M_{A,H}$, $M_{A,B}$, $M_{B,A}$. In contrast to a single stage inference process, in multistage inference the decision maker must explicitly consider the impact of his initial data evaluation in terms of the diagnosticity of the intermediate stages, in order to revise the prior probabilities over the terminal hypothesis set.

The intermediate inference stages serve to reduce the diagnostic impact of the data on the terminal hypothesis set. As in any logic chain, confidence in a conclusion cannot exceed the confidence level of the weakest step in the logic. For example, from available information, an analyst may associate a high probability with one of a set of indicators. However, if the indicator has only a weak association with any of the potential alternative courses of action, the impact of the information on the final estimate will be slight. What at first appears to be highly diagnostic data may not, in fact, be very diagnostic. The impact of data on the final inference depends on the probabilities embedded in the logic chain.

Multilevel Data. The intent of decomposition is to construct an explicit logic chain relating the lowest level of data to be a terminal hypothesis set. In deriving the model of Eq. (12), it was explicitly assumed that all data could be directly related to the lowest level hypothesis set $B$ and that any impact on higher level hypotheses was mediated by the conditional links to hypothesis set $B$. A more general case is one in which each data element need not be related to hypothesis set $B$, but may instead be directly related to hypothesis set $A$ or to hypothesis set $H$. However, a real data stream may be, and probably would be, of the same hierarchical order as an appropriate multistage model.

The concept of data having differing levels of impact can be
formally incorporated into the multistage model. Assume that the
data set \( D \) is composed of equivalence classes \( D_H, D_A, D_B \), where the
subscript indicates the hypothesis set which mediates the data class.
Data elements in each class are assumed to be dependent only upon the
subscribed hypothesis set and at least pairwise independent of the
remaining hypothesis sets. Under these assumptions Eq. (10) becomes

\[
V_{H,D} = C (M_H^T M_D^T H M_A, H M_D^T A M_B^T A M_D^T B V_I)
\]

where \( M_H, M_D^T, M_A^T, M_B \) are diagonal matrices of \( P(D_H/H) \), \( P(D_A/A) \)
and \( P(D_B/B) \), respectively, \( M_A, H \) and \( M_B, A \) are transition matrices re-
ating hypothesis set \( A \) to \( H \) and hypothesis set \( B \) to \( A \) as before, and
\( V_I \) is a unit vector of rank \( K \). When a data class such as \( D_H \) is
empty, the convention is used that \( P(D_H/H) = 1 \). Thus, when the data
classes \( D_H \) and \( D_A \) are empty, Eq. (15) reduces to Eq. (9).

The effect of considering multilevel data in a multistage system
is that an internal symmetry is established, with each stage of the
system becoming essentially identical to every other stage. The
result is a structured decision procedure in which data can be eval-
uated relative to the hypothesis set to which it is most directly
related; the eventual impact of the data on the terminal hypothesis
set being mediated by the relationships between any intervening
hypothesis sets.

Intermediate Level Estimates. Multistage inference may appear
to be a series of single-stage inferences in which the output at each
stage becomes the input to the next stage and so on, until the termi-
nal hypothesis set is reached. Implicit in this view is the concept
that at each stage of the inference, a revised estimate of the pro-
babilities over an intermediate hypothesis set is obtained which then
serves as the input to the next stage. However, at each intermediate
stage of the inference process, what is obtained is a revised estimate
of the diagnostic impact of the data. In the previous example, start-
ing with \( P(D_B/B) \), one obtains in succession, \( P(D/A) \) and \( P(D/H) \).
The final step in the process is a revision of the probabilities over
the terminal hypothesis set. Estimates of the impact of the data on
the intermediate hypothesis sets are obtained by additional calcula-
tions through a "folding back" procedure (Figure 1).

The Decomposition Process. In common with other techniques of
decision analysis, the initial structuring of the problem is the
critical step. Decision analysis has not often been used in intelli-
gence and the absence of prior experience such as found in business
applications compounds the inherent difficulties of decomposition. Many inference tasks are intuitively hierarchical and a few such as collection planning have a close correspondence with formal multistage models. However, for most tasks of interest in tactical intelligence, the development of a formal multistage structure is both difficult and time-consuming. There are no formal algorithms to aid in identifying either the desired sequence of intermediate hypothesis sets or the dimensionality of the hypothesis sets. Several quasi-mathematical approaches exist for identifying a set of dimensions for a psychological space (e.g., factor analysis, multidimensional scaling). However, these techniques are cumbersome and time-consuming in application and are not well suited for the problem of constructing intermediate hypothesis sets. Thus, logical analysis requiring considerable time and effort is needed to develop a multistage inference structure for specific problems. The process of decomposition itself will aid in the improvement of intelligence by clarifying specific problems and enhancing the analyst's understanding of the elements of the problem. Although the details of decomposition must change to fit the situation, the problem structure will usually generalize across situations. An analysis of indicators within one scenario would simplify problem analysis and improve their use in other scenarios.

Advantages of a Multistage Model. The matrix format of the multistage Bayesian inference model (Figure 1) inherently lends itself to use in interactive, computer-based systems. Additional intermediate hypothesis sets may be readily incorporated into an online computer system by simply adding established subsystem computational modules. Moreover, the logical structure of the system implementation allows it to be used by individuals with little sophistication in mathematics.

The model appears to provide a number of desirable characteristics for solving complex decision problems. First, it facilitates the function of relating data to a primary hypothesis set by formally decomposing the process into a sequence of less complex steps. The analyst is allowed to build from the specific to the general in several gradual stages. A consequence is that information flows within the system follows a logical sequence in which decisions at one level become data for the next, more general level.

Second, the system facilitates the integration of historical data and/or expert opinions, which may be used as the prior conditional probability assessments within any of the various intermediate estimation matrices in the system. In this manner, an inexperienced user may be formally assisted by previous information, and will be able to
integrate these opinions directly into his own inference process.

Third, the system formalizes the bidirectional flow of information in an inference system. Data is related to the final hypothesis set through a sequence of intermediate hypothesis sets. A feedback loop relates the resulting system output to each intermediate hypothesis set. Thus, conditional and unconditional probability estimates reflecting the impact of data at each level of the system can be derived.

Alternative Structures. The multistage model developed here is only one of a family of multistage Bayesian models. The initial decomposition is a straightforward application of probability theory with the specific model arrived at dependent on the set of assumptions used to simplify the resulting decomposition. Thus, employing relatively simple mathematics, the result is a structured procedure which is both potentially useful and isomorphic to inference tasks found in intelligence analysis.

Bayesian models are not the only information processing structures which could be used as the basis for a multistage inference system. Any inference system is a set of decision rules and an algorithm for applying them. Hence, a similar decomposition can be accomplished in different ways. Any Bayesian structure can be represented in any one of several equivalent schemes: flow charts, decision trees, relevance trees, state-of-affairs models, etc. Further, Bayesian structures can be represented in analogous linear models such as regression models. All of the decomposition techniques relevant to a specific schema are primarily techniques for mapping a problem space into an information processing structure. The exact mapping or structure which is most useful will depend on the inference problem and the specific tasks to be accomplished. For example, flow charts and decision trees are likely to be most useful in the initial decomposition, whereas a Bayesian or regression model would be more useful for the aggregation of data.

Data Reliability. In the preceding development of multistage inference models, data have been implicitly assumed to be accurate. However, this is not a necessary assumption. Any procedure for incorporating source reliability into the inference process must differentiate between the actual occurrence of an event and the report of its occurrence. Data reliability can be incorporated into the multistage model by considering the links between events and reports of events as another stage in the inference process.
EXPERIMENTS CONCERNING MULTISTAGE INFERENCE

In the study of human inference performance, as in any study of human behavior, a critical problem is the establishment of criteria against which the performance can be evaluated. A useful approach to this problem is the development of appropriate prescriptive models which define optimal performance. The multistage inference model presented above provides normative performance standards and serves as the framework as well as the focus of the experiments described below. On the one hand, there is interest in the adequacy of the model as a description of human inference performance. On the other hand, the basic thrust of our research is to understand how man processes and utilizes information in specific classes of inference tasks and to develop techniques for improving and enhancing performance; the model provides a convenient framework for this study.

Three experiments recently conducted atARI will be used to illustrate the research approach and to provide insight into man's information processing and utilization in multistage inference tasks. Each of the experiments explores a different aspect of human performance and each provides information needed in the development of techniques for improving and enhancing performance.

Threat Diagnosis. An initial evaluation of a computer-based multistage inference system was provided by an experiment using a simple threat diagnosis task. The experiment and a pilot study involved over 20 intelligence officers. The experiment provided for data collection in a credible task and had three primary objectives: first, to guide the development of the multistage model; second, to assess the attitudes and opinions of intelligence officers toward the use of on-line computer-based inference aids; third, to compare aided inference performance with unaided inference performance.

The setting for the experiment was a scenario involving an Aggressor unit on fall maneuvers near a friendly border. A series of 50 messages provided the player with information which he used to estimate the probabilities that the Aggressor will Attack, Defend, or Withdraw. A simplified three-stage inference model served as an aid. Players were briefed on the structure of the model (Striking Force Posture (B)---Indicators (A)---Alternative Courses of Action (H)) the "current situation" and the type of judgment to be made in evaluating each message. Upon receiving each message, the player's task consisted of updating the situation map and entering an evaluation of the data through a CRT keyboard. Each player either evaluated the data relative to Striking Force Posture, P(D/B), or evaluated the data relative to Alternative Courses of Action, P(D/H). After each
JOHNSON & HALPIN

During the pilot stages of this experiment had a strong influence on the structure of the model presented earlier. This influence is apparent in the reactions of players to the system used in the experiment. All of these players felt they had an adequate understanding of the system and all provided reasonable explanations of system processing. All of the players felt that this type of system could significantly improve operations through more efficient utilization of information, increased speed of operation and greater acceptance of conclusions by the G3/S3. However, most of the players felt that the system would not increase accuracy.

In estimating the posterior probabilities of the alternative courses of action, players using the system tended to assign a higher probability to the course of action actually chosen by the aggressor, and to identify this alternative earlier in the course of play than unaided players. Performance was better for players receiving system feedback and for players using a multistage model rather than a single stage model. However, since the estimates in the intermediate stages were given to the players, further validation of these comparisons will be required to ensure that they do not reflect experimental biases.

These results indicate that a relatively simple man-computer dialogue and a multistage inference model could provide an acceptable inference aid for intelligence analysis. Although the results cannot guarantee that analysts would actually use such a system in an operational setting they do imply that such a system would be acceptable and used.

Unit Identification. Intuitive inference performance was examined in the context of a unit identification task. There were three objectives: first, to compare intuitive inference performance with a normative model; second, to analyze the effects of the number of stages in the inference process; and third, to determine the effects of differing levels of diagnosticity in the intermediate stages of the inference process and in the data. Thus the focus of this experiment was on the integration of information in multistage inference.

A series of 24 multistage "unit identification" problems were
presented in a test booklet to 18 enlisted men; the problems varied in the level of data input (company, battalion and division), and in the diagnosticity of the data and of the intervening stages. On each problem, subjects estimated the posterior probability that the unidentified division was of a certain type; all, armor and mechanized. The inference problem was decomposed into a three stage logic chain relating the probability of each company type to each battalion type (stage 1), the probability of each battalion type to each division type (stage 2) and ending with the prior probability of each division type (stage 3). The probabilities used in the two intermediate stages were chosen to be realistic, while the prior probability of each type of division was equilikely. Given information related to the type of company, battalion or division present, it was necessary to process this information through two, one or no intermediate inference stages, respectively in order to estimate the posterior probability of each division type. Estimates of the posterior probability of each battalion type and/or of each company type were also made on selected problems.

The results indicated that subjects failed to appreciate the degree to which a multistage inference is less diagnostic than its component single stage inferences. As the information given on a problem became logically distant from the probability of division type, as the number of intermediate inference stages increased, intuitive subjective estimates became more extreme relative to the optimal Bayesian estimates. Thus, subjects were less accurate in estimating the divisions identity, \( P(H_t/D) \), using information at either the company or battalion level, \( P(\tilde{H}_t/B_x) \) or \( P(D/\tilde{H}_t) \), respectively, than using information at the division level, \( P(D/H_t) \). Subjects' estimates were sensitive to the diagnosticity of the data and of the intervening inference stages; however, they used non-optimal inference strategies in integrating information into the inference process. This sub-optimal information integration suggests man-computer symbiosis as an approach for improving performance in multistage inference tasks. Man evaluates the data and the computer, using mathematical models, synthesizes human judgment into the inference process.

**Source Reliability.** Intuitive inferences based on data reports from sources of varying reliability were investigated in the context of a symmetric binary decision task. In addition to comparing performance with the normative model presented earlier, this experiment allowed a detailed analysis of subjective inference strategies.

Twenty-one subjects each received 60 problems consisting of all combinations of five levels of source reliability, four levels of data diagnosticity, and three levels of sample composition. On each pro-
blem, the subject received a report of a sample of data of known diagnosticity from a source of given reliability. He then indicated the most likely of the two hypotheses and his subjective odds favoring that hypothesis.

The subjects generally failed to extract as much certainty as possible from the data reports. Subjective odds were generally conservative (lower) with respect to the normative odds from the model. However, as reliability decreased, subjective odds decreased at a much slower rate than the normative odds (error decreased) until at the lowest level of reliability they were generally greater than normative odds.

Verbal protocols from the subjects and data analyses suggested two different, non-optimal, strategies subjects may have been using: a simple multiplicative rule and a derived multiplicative rule. To evaluate the fit of these two heuristic rules and the normative rule to subjects' responses a correlation analysis was performed. Product moment correlation coefficients between each subject's odds and the odds predicted by each rule were computed. The average coefficient and the average percentage of variance accounted for by the rules were (.31, 10.3%) for the normative, (.65, 44.4%) for the simple multiplicative, and (.80, 67.0%) for the derived multiplicative.

It is clear from these results that intuitive performance is far from optimal and the nature of this departure from optimality suggests several approaches for improving performance. One might be the traditional approach of training. Second, users could be required to consider other events which may have occurred, but which were not reported. A third strategy might be the use of computer aids based on the optimal multistage models.

SUMMARY

Multistage inference models provide a framework for the analysis of current modes of intelligence processing. Within this framework, two broad questions can be considered: "What is the man in the system doing with the information available to him?" and "What should he be doing with it?"

The first question raises a psychological issue which revolves around understanding how man processes and uses information. Experiments in intuitive inference in multistage tasks indicate that man's performance in processing information is sub-optimal and results from his use of heuristic strategies which are cognitively less complex than the optimal strategies. Man fails to properly integrate
JOHNSON & HALPIN

the information available to him.

The second question is more practical and involves the development of aids and methods to enable more efficient and effective information processing. Man-computer symbiosis, in which man evaluates the data and the computer, using mathematical models, synthesizes human judgments into the inference process, is a useful technique readily acceptable to intelligence officers for improving performance in multistage inference tasks. The introduction of tactical data systems such as the Tactical Operation System could provide the on-line computer support requisite to implementation of such aids.

REFERENCES


FIGURE 1. FLOW DIAGRAM FOR A THRESHOLD-LEVEL INFERENTIAL SYSTEM, SINGLE STAGE DATA