

HUMAN INTUITION AND DECISION-MAKING SYSTEMS (II)

Linda ELLIOTT and Andrew BORDEN

Background

One of the authors of this paper has developed a canonical design method for designing Situation Assessment strategies in the form of very efficient decision trees.^{1,2} This method uses the mathematical theory of information, developed by Claude Shannon in the 1940's, to reduce uncertainty (Entropy) most efficiently.³ The method applies when:

- There is a data base containing a statistical description of the objects in the Frame of Discernment;
- There is a capability to measure at least some of the parameters in the data base;
- There may be a differential cost, usually in time, for making parameter measurements;
- High confidence, on-time classifications are important.

Specifically, the program, called the Situation Assessment Evaluation Tool (SAET) produces a decision tree by selecting the most efficient entropy-reducing parameter to expand each node. When the decision tree is complete, the program provides a report card covering performance. The report card contains statistics which include reliability and response time. The program can also run a simulation and provide very detailed information about the performance of the run-time algorithm.

The motivation for developing the SAET was the task of designing an efficient Radar classification Algorithm for use in Radar Warning Receivers in combat aircraft. The Radar Warning Receiver makes measurements of radar parameters on demand. Each measurement has a cost in time. There is a strategy for fetching the measurements and comparing them to the data base in order to produce a confident, on-time

identification of the radar. Appropriate warning is given to the aircrew using an alpha-numeric cockpit display and audible tones.

This task is typical of many Situation Assessment tasks. Another such task is the design of Indications and Warning (I&W) programs. In I&W, parameters are collected and evaluated according to an efficient strategy in order to determine the probable activities and intentions of an adversary. In one of the previous studies,² the SAET was applied to several Situation Assessment tasks. A preliminary examination of the data bases suggested that the tasks were approximately comparable in difficulty. However, one of the assessments turned out to be two orders of magnitude more difficult than another. An I&W designer relying on intuition could easily be led astray if the perception of the relative difficulty of the design tasks was so far from being correct.

In the present task, airborne target threat assessment, a similar finding occurred. Data that (intuitively) should be helpful to a human decision maker actually degraded the capability to make confident, on-time decisions. In this paper, a description of the threat assessment study is given and the non-intuitive finding is explained.

The Target Classification Study

In a study being conducted by the Air Force Research Laboratory, Brooks AFB, Texas, the performance of human observers in learning to classify airborne threats is being investigated. In this study, a deterministic algorithm classifies targets using a point-scoring system. The lowest scores indicate that a target being seen by a radar is friendly without doubt. The highest scores mean that a target is hostile without doubt and represents a threat to friendly forces.

Nine parameters are used by the observers in making their assessments. In some cases, a partial presentation of five of these parameters is given. The parameters are visually coded so that the effects of shape, size and color of the presentation can be determined. The nine parameters are:

1. Type of IFF (transponder) return
2. Speed
3. Flight path relative to the radar
4. Size
5. Position relative to known airway corridors
6. Altitude
7. Range to the radar
8. ESM indications (type of airborne radar)
9. Change in altitude

Each parameter is scored as “Zero”, (Minimal threat) to “Two” (Maximal threat). Before being summed, the parameter scores are processed to reflect weight of importance and pairwise interactions. Since IFF is relatively important, its score is doubled, giving possible weights of “0”, “2” or “4”. The remaining eight parameters are grouped in pairs. The score for the pair is the product of the scores for the two parameters in the pair. For example, if a target received a score of “2” for speed and “0” for flight path, the score for the pair would be “0”. After processing the scores for IFF and for the remaining pairs of parameters, the numbers are added. A high score corresponds to “Defend” and a low score correspond to “Ignore”. Intermediate scores match the categories of “Monitor”, “Review”, “Warn”, “Ready” or “Lock-on”.

A large number of human trials were conducted. The subjects were told the meaning of the visual and other cues and were told generally that a high score corresponded to a more serious threat. They were not given the specific algorithm used for scoring, but they were told which parameters were to be considered pairwise. In some cases, they had only five of the nine parameters to consider. In other cases, time constraints were introduced.

The human responses have been analyzed and are currently being compared with the results produced by the ideal decision maker provided by the SAET. The remainder of this paper will consider the performance of the ideal decision maker, not the human subjects.

Results

The initial reaction of the investigator to the results of this experiment was that the SAET had worked hard to produce a poor result. Based on the relative frequency of outcomes in the Frame of Discernment, the initial Entropy is about 2.1 bits (The amount of Entropy associated with all the possible outcomes of two (three) tosses of a fair coin is two (three) bits). Each of the nine parameters, if used first, would reduce the Entropy by about 0.25 bits. After the optimal decision tree runs, about 1.47 bits of uncertainty remain.

The SAET was run to the 50 % confidence level, then forced to make a decision based on the highest probability, even if lower than 50 %. The correct classification was made 56 % of the time and the mean error was 0.58 categories. That is, a result of “4” would, with very high probability be between “3” and “5”.

“Max Nodes” are the possible combinations of parameters that could be encountered. The program only had to generate a small fraction of these combinations to design the decision tree. Generating more nodes would not have improved the result.

The second run (parameter dependencies considered) produced a slightly worse result. Each parameter, if used first, would reduce Entropy by only 0.14 to 0.19 bits. The remaining Entropy after running the optimal decision tree was 1.48 bits.

The forced decision was right 54 % of the time and the mean error was .62 categories. The investigator's reaction was that this result was not consistent with intuition and could be an error in the program. Since the data base for the second run was based on partial execution of the "ground truth" algorithm, the result should have been better.

Table 1. Performance of the optimal decision-maker compared to the "Real-World" (algorithmic) solution.

CONDITION	ENTROPY (INITIAL/FINAL)	PROBABILITY OF CORRECT DECISION	MEAN ERROR	MAX NODES	NODES USED
DEPENDENCY RULES NOT USED	2.1/1.47	0.56	0.58	19,683	353
DEPENDENCY RULES USED	2.1/1.48	0.54	0.62	1024	325

Discussion Of Results

The statistical distributions of the parameter measurements were almost identical. Only the first measurement (IFF) was slightly different from the others. This was true in both the nine parameter and the five parameter cases. There was a great deal of noise (overlap) in the probabilities for each parameter. In other SAET investigations, one or more parameters had minimum overlap, at least for several of the classifications being predicted. As a result, a classification of "4" for example, could be quite far from a result of "3" or "5". The program finds these productive parameters and uses them first to reduce uncertainty more efficiently. This property of the data base enables the SAET to prune the resulting decision tree very quickly.

In the present situation, results of "3", "4" and "5" are statistically quite close. It is difficult for the SAET to distinguish them with high confidence. Therefore, the performance of the SAET is reasonable in this very difficult decision-making task. The mean error of a classification was acceptably small. There is probably little operational importance to an error of one category. If a high threat target "7" were classified as friendly "1" that would be very serious, but this would occur rarely if at all.

The relatively poor performance of the combined (five) parameter case is also reasonable, but it seemed difficult to understand how partial execution of the scoring algorithm can actually be a handicap in predicting the threat category. The SAET is based on the assumption that any two parameters are either conditionally independent or correlated. The rule for combining parameter measurements in this investigation, however, is to multiply the scores for the two parameters and use the result. Thus, if one parameter was a “2” (high threat indication) and the other was a “0”, (low threat), the result would be “0” for the combined parameter. The impact of this method is that one parameter can negate another, thereby reducing the information bandwidth by discarding data that should produce real information. Parameters are neither conditionally independent nor correlated. This part of the investigation was studied very thoroughly. It appears correct that informing the human observers, in part, of how the scoring algorithm works actually degrades their decision-making ability.

Comment

It is the peculiar strength of the SAET that it surpasses human intuition and sometimes produces surprising, but correct results. The investigator noted a similar result in doing Indications and Warning experiments in which some problems turned out to be much harder to accomplish than other. Only a meticulous examination of the data base revealed why this was so. The current project seems to be another case in which the mathematical theory of information beats intuition.

References

1. Andrew Borden, “The Design And Evaluation of Situation Assessment Strategies”, *Information Security: An International Journal* 1, 1 (Summer 1998), 63 - 74.
2. Andrew Borden, “Human Intuition and Decision Making Systems,” *Information Security: An International Journal* 1, 2, (Fall-Winter 1998), 67 - 72.
3. Claude Shannon, “A Mathematical Theory of Communications,” *Bell Systems Technical Journal* 27 (1948), 379 - 423 and 623 - 656.

LINDA ELLIOTT received her Ph.D. in Management Sciences from the Michigan State University. She is a Senior Staff member with the Veridian Corporation, San Antonio, Texas.

ANDREW BORDEN is a retired USAF officer with a long background in developing systems that make decisions, especially in military avionics. His last active duty assignment was as Deputy Chief of Staff for Intelligence, (then) Electronic Security Command. He has worked in industry, in academia and for NATO as Principal Scientist for Electronic Warfare, SHAPE Technical Centre (now the NATO C3 Agency). Mr. Borden is the Chief Scientist of DRH Consulting, San Antonio, Texas. He has advanced degrees in mathematics from Kansas State University and The Ohio State University.