THE DIALECTICS OF INFORMATION A FRAMEWORK

Andrew BORDEN

Introduction

The dialectics of Information applies whenever there is a human conflict or competition in which information:

- Is a commodity that is not shared
- Is subject to attack

Three examples are:

- Military Command and control
- The game of bridge
- The banking industry

In the first example, attacks on the adversary's information are commonplace and the need to protect friendly information is recognized.

In the second example as well, a good player will convey as much information to his partner as possible while concealing it from the opponents. Consistent with the rules of the game, providing misleading information is even allowed by playing an unusual card (false carding).

In the third example, protection of friendly information against attacks is recognized as critical to business success, but the attacks themselves are usually unethical, if not illegal. Exploitation of a competitors information is equally important, so data privacy is a critical element in the banking industry.

The purpose of this paper is to present a generic, domain-independent framework for the information dialectic and to show how the framework can be applied to any selected domain. The program for doing so is straightforward:

- Define the generic tasks relevant to the development of information

- Identify the generic types of attack that can be mounted against these tasks (Identify Attack measures)
- Show how the performance of the tasks can be protected against these Attack Measures (Identity Protect Measures)

This framework will be developed in the context of an analog to the Shannon-Hartley channel capacity theorem.

The Shannon-Hartley Theorem

The Shannon-Hartley Theorem is one of the most elegant mathematical results of the twentieth century. In the proof, signal and noise are represented as an infinite dimensional vector in Hilbert Space. The proof also uses a simple, but little known fact from geometry…that the mass of an n-dimensional sphere migrates to the surface when the number of dimensions increases. The proof of the theorem relates these ideas from physics and geometry to the very abstract mathematical characterization of Information. The theorem is as follows:

$$C = W * LOG_2 (1 + S/N) \quad (1)$$

W is the bandwidth of a signal being transmitted over a noisy communications channel. S/N is the signal to noise ratio. C is the Channel Capacity, measured in bits per second. We are guaranteed that there exists a way to code information so that it can be transmitted at a rate arbitrarily close to the Channel Capacity over this noisy communications channel.[5]

Channel Capacity can be used as a unifying principle for Electronic Attack (EA) and Electronic Protect (EP) Measures in Electronic Warfare (EW). Every EA measure except Exploitation is an attempt to reduce the bandwidth of an adversary signal and/or to reduce the Signal to Noise Ratio. Every (EP) Measure (except Protect measures against Exploitation) is an attempt to increase bandwidth and/or increase Signal to Noise Ratio. For example, communications frequency hopping as an EP measure uses a large total bandwidth to protect against Jamming, but a small instantaneous bandwidth to protect against interception and Exploitation. The large total bandwidth in this case makes it difficult for the jammer to set on the transmission frequency, thus preventing a reduction in Signal to Noise Ratio.

For another example, repeater or gate stealing EA techniques must achieve a certain reduction of Signal to Noise ratio within the bandwidth of the victim signal to be effective. The corresponding EP technique might utilize a combination of guards and filters to recognize and eliminate the unwanted jamming signal, thereby protecting the signal to noise ratio.

Against Exploitation, a very large bandwidth with low average power might be used. The low average power reduces the probability of intercept, but the energy over the large bandwidth can be summed to extract the information from the signal. Therefore, the transmitter compensates for the low signal to noise ratio with increased bandwidth to transmit information at a fast enough rate. The jammer can only achieve high signal to noise ratios over small portions of the bandwidth.

The use of the Shannon-Hartley channel capacity formula as a unifying principle in EW is a useful device when teaching the subject. It gives the students a number of logical pegs on which to hang their collective hats. It is only useful however, because the Shannon Hartley theorem has provided such an elegant and simple way of determining the Channel Capacity.

The Capacity of a Decision Making System

It is tempting to think of decision making in the presence of uncertainty as analogous to attempting to send information through a noisy-communications channel. When doing a decision making problem, we are attempting to classify an event or object as one of a number of recognizable events or objects. There is an initial amount of Entropy or uncertainty based on the *a priori* probability distribution. If we look at one attribute of the object being studied and compare it to the data base, we may be able to reduce the Entropy. The percent reduction in the entropy is the Signal to Noise ratio for this attribute. The number of bits by which the Entropy is reduced divided by the time it took to evaluate the attribute is the decision making channel capacity for this attribute (bits/second).

Presumably, we have used the most efficient Entropy reducer first. If it doesn't solve the problem completely, we now have to evaluate the remaining entropy reducers (attributes) and use the one which is now most efficient. We continue recursively until the Entropy is reduced enough so that the probability of one event or classification exceeds the required confidence threshold. If we use all attributes and cannot reach the confidence threshold, the decision making system has failed for this object (event).

This process of designing an efficient decision making system can be summarized as follows:

> Compute the (information) Channel Capacities of all the available attributes, taking into account the current Entropy and the attribute values which are already known.
>
> Pick the available attribute with the best (information) Channel Capacity
>
> Measure this attribute

Compare the result to the data base. (Note: This changes the (information) channel capacities of all the remaining attributes since they are not independent)

If done, report

If not done, go back to the first step

For example, experience tells us that Pulse Recurrence Interval is the most efficient uncertainty reducer when attempting to identify a radar. It will be used first, then the remaining best Entropy reducer, probably frequency, will be identified and used. This procedure is repeated until the confidence level is reached or until the process fails.

This procedure is analogous to using one noisy channel, re-evaluating, using another noisy channel, etc. Each branch in the decision tree leading to solutions uses a different sequence of attribute measurements. Unfortunately, there is no analog of the Shannon-Hartley theorem to give us an elegant determination of the overall Channel Capacity for this disorderly situation.

The Analog of the Shannon-Hartley Theorem for Decision Making Systems

In the place of the Shannon-Hartley Theorem, we use a result from the mathematics of information [4]:

$$I(\text{Situation} \mid \text{Observations}) = H(\text{Situation}) - H(\text{Situation} \mid \text{Observations}) \qquad (2)$$

H(Situation) is the initial amount of uncertainty (bits) in the problem to be solved. H(Situation | Observations) is the amount of uncertainty (bits) remaining after the decision making system has been used. Therefore, I(Situation | Observations), the Mutual Information of the decision making system in this situation, is the expected value of the amount of uncertainty that has been removed by the decision making system. (See Reference 4 for a complete discussion).

Formula 2 seems a most unsatisfactory substitute for the Shannon-Hartley theorem. (data) bandwidth and Signal to Noise ratio do not appear explicitly in the formula. With no elegant computational formula, it seems impossible, practically speaking, to use it for real problems. The solution to this dilemma is to substitute computing power for elegance.

Quantifying Mutual Information

Finding the optimal Decision-Making System (DMS) under the circumstances we have described is an uncommonly difficult type of problem called NP-Complete. By accepting a slightly sub-optimal, but demonstrably good, design method, we can

develop a natural, formal method for building very efficient DMS's. The computations required are still formidable, but manageable if the problem isn't too large.

The computational difficulty is the down-side, but the corresponding advantage is that the computations enable the user to compute the (information) channel capacity as given in Formula 2. As when using the Shannon-Hartley formula, the designer specifies a confidence level and provides a data base and a capability to measure the values of a number of attributes. Presumably, each measurement has a cost in time. The noise (ambiguity) is inherent in the data base and is usually not under the control of the designer. For the attributes specified by the designer, the (information) Channel Capacity is a natural product of the design process. Response time and confidence level actually achieved are also available. The probably of a successful response and the conditional probability of a correct response are also provided.

The Mutual Information is a performance measurement very like the Shannon-Hartley channel capacity. It only has meaning however, if a standard, formal method is used to design the DMS. This method, and some experiments using it, were described in this journal.[1,2,3]

One of these experiments involved identifying a radar coming from a population of five radars. The Table contains the performance results for the best achievable DMS using the data base and measurement capabilities used in the problem.

Initial entropy (bits)	Final entropy (bits)	Entropy reduction (bits)	Mean time to classify (seconds)	(Information) Channel capacity (bits/second)
1.96	0.35	1.61	0.87	1.85

Based on the assumed *a priori* distribution of radars, the initial entropy is 1.96 bits. The conditional entropy (the final term in Formula 2) is 0.35 bits. Dividing the Entropy reduction by the mean time required gives the rate of entropy reduction in bits per second. This number means very little when taken out of context. However, a designer of Radar Classification algorithms would become very familiar with it and would have a pretty good idea of what (information) Channel Capacity would be good enough to meet operational requirements. If not good enough, the designer would attempt to improve it by increasing the (data) bandwidth (finding more parameters to measure and evaluate). Alternatively, the designer can state a requirement for higher quality data with less noise (ambiguity).

The Dialectics of Information

There are four tasks that must be performed on the information battlefield:

- Collect (data)
- Move (data)
- Store (data)
- Use (data) to perform situation assessment

The Collection task, for example, could be carried out by a search engine that looks for key words and retrieves text that might be relevant to the user's data requirements. Movement can be accomplished, by sending encrypted ASCII characters through a satellite link. Data can be Stored on paper in a filing cabinet or in compressed form on an optical disk.

If a standard method is used to design the situation assessment strategy, then we can give it a report card as shown above. Its performance will depend on the amount and quality of the data (the data bandwidth and the amount of noise). The data is vulnerable when it is being Collected, Moved and Stored. The Attack measures that can be taken against the data are the following:

- Degrade the data (delay or delete some data elements)
- Corrupt the data (add false data)
- Deny the data completely (usually by direct attack on the means of collecting, moving and storing)
- Exploit the data by listening, decoding and interpreting (usually when it is being moved)

An example of degradation against the Collection task would be the use of concealment. The use of dummies would be an example of Corruption against the Collection task. An example of Corruption against the Movement task would be intrusion and spoofing, that is transmitting false data that looks genuine. An example of Denial against the Storage task would be the introduction of computer viruses that damage operating systems, making the computer unusable for Situation Assessment purposes. The specific means for accomplishing Attack measures depend on the means being used to perform the information tasks

Exploitation is different from the other attack measures in that it does not affect the data in any way. It could be regarded as a part of the Collection task, rather than an Attack measure. As such, it would enhance (data) bandwidth and make adversary situation assessment more efficient.

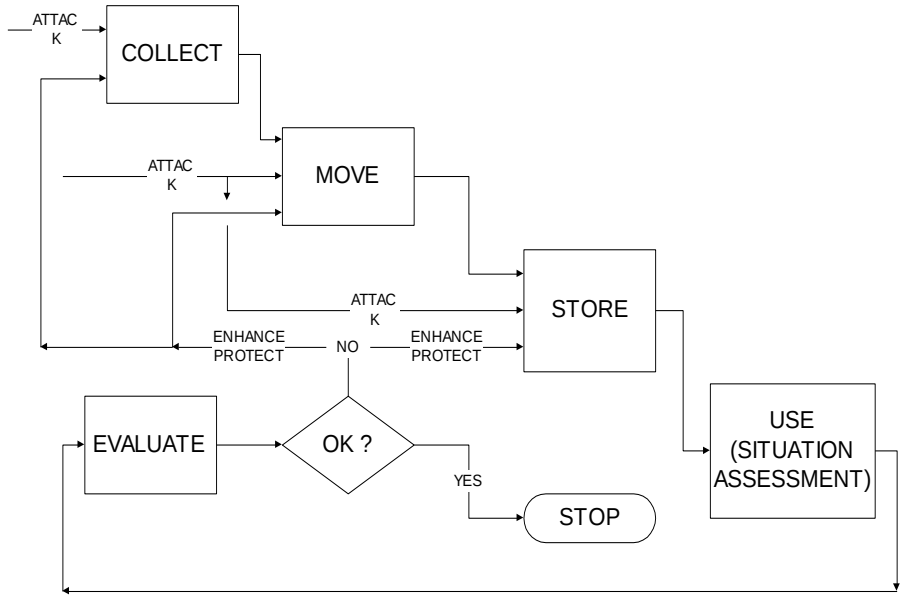

Figure 1. The Information Dialectic

The figure shows the framework for the information dialectic. The information tasks are shown: Collect, Move, Store and Use for Situation Assessment. Situation Assessment using the standard method of strategy design is evaluated for confidence and response time. This can only be done because the standard design method makes a large number of statistics available. If performance does not meet requirements, then the tasks of Collect, Move and Store must be enhanced. Either the (data) bandwidth must be increased by adding entirely new sources of data or the noise (ambiguity) must be minimized by protecting the three data tasks against attack.

The ability to measure the efficiency of Situation Assessment makes it possible to quantify Figures of Merit wherever the information dialectic applies. ..whether in war…in commerce or in contests of skill and ingenuity.

Conclusion

This domain independent framework for the dialectics of information is especially simple. It is only useful however, if the efficiency of the resulting Situation Assessment can be measured. The use of a standard design method for situation assessment strategies makes this evaluation possible.

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. Linda Elliott and Andrew Borden, "Human Intuition and Decision-Making Systems (II)," *Information & Security: An International Journal* 2, 1 (Winter 1999), 50 – 54.
4. Pierre LaFrance, *Fundamental Concepts in Communications* (Prentice Hall International Editions, 1990).
5. Claude Shannon, "A Mathematical Theory of Communications", *Bell Systems Technical Journal* 27 (1948), 379 – 423, 623 – 656.

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. His latest research interests are in comparing human and machine-based decision making. Mr. Borden is member of the Editorial Board of *Information & Security*. Address for correspondence: DRH Consulting, 1210 Scenic Knoll, San Antonio, TX 78258; Fax: (210) 497 4581. E-mail: borden@wireweb.net.