2018 Conference on Systems Engineering Research

Information Theory Applied to Decision Making Structures

Michael D. Watson[a]*

[a]*NASA Marshall Space Flight Center, Huntsville, AL 35812*

Abstract

Decision making structures, such as control boards, process information on many topics as they select options for the system and project. The decisions are based on the information known to the board members and presenters (subject matter experts) and shared at the board during discussion. This information flow through this process can be modelled using information theory. Information theory provides a mathematical basis to understand the flow of information through the decision making process and the information needed for a particular decision. Information theory also provides the mathematical relationships on which to base the optimal decision making structure for a specific system development and organizational structure. Since decision making bodies provide control for the system or project, control theory can be used to construct a decision making model. This provides a starting point for adding cognitive science models. Information processing by each individual board participant can be represented through cognitive processes which are integrated across the board participants through information theory relationship. The set theory view of information theory provides a structure in which to look at the relationships between the participants in a decision making structure.

Keywords: Cognitive Science, Control Theory, Decision Board, Decision Making; Information Theory; Set Theory;

Introduction

Information flow through an organization in the development or operation of a system is an important aspect of systems engineering. Systems engineering ensures that the correct information is provided to the correct engineers when or before it is needed. This is performed by understanding and managing the information about the system which resides in the design and in the organization. Information theory provides the tools to understand and manage this flow and the organizational decision structures which utilize this information.

Information theory has been applied to decision theory[i,ii] and provides a relevant structure to model a decision-making body (e.g., decision board). Webster's Dictionary defines information theory as "a theory that deals statistically with information and the measurement of its content in terms of its distinguishing essential characteristics or by the number of alternatives from which it makes a choice possible, and the efficiency of processes of communication between humans and machines." Expanding this definition to include human communication encompasses the organizational communications and hence decision making bodies. The decision-making body essentially operates as a communication system where information is presented and shared in an open forum. Fig. 1 illustrates a basic communication system model.

* Corresponding author. Tel.: 256-544-3186.
E-mail address: Michael.d.watson@nasa.gov

Nomenclature

C	Cognitive Function
f_n	Message or Understanding
H	Information Enthalpy and Information Uncertainty
I	Information
p_n	Probability that Message is transmitted
q_n	Uncertainty
X_n, Y_n	Information Contained in Memory

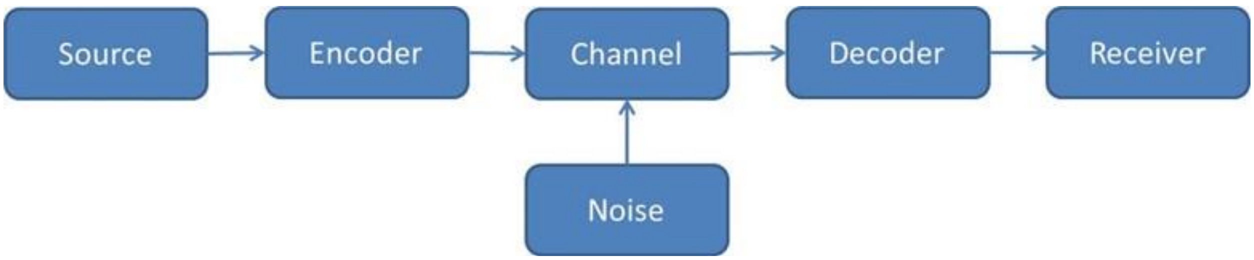

Fig. 1. Communication System Model

Information transmitted through this communication system model is a logarithmic function:

$$I = -\log p_n \tag{1}$$

Where p_n is the probability that message f_n was sent. Taking the average, this is the measure of uncertainty that the transmitted information represents a specific event,

$$\bar{I} = H = -\sum_n p_n \log p_n \tag{2}$$

$\bar{I}$, representing the uncertainty that an event occurred, is also defined as the Information Entropy, H, of the communication system.[iii]

1. Single Board Structures

In the context of a board structure each board member acts as a source and encoder, contributing information to the discussion. Each board member also acts as a decoder and receiver, receiving information and understanding (or interpreting) the meaning of the information. In this model, the board members include the board chair person. The chair person has the final decision authority in the board setting. In addition, subject matter experts (SME) often present information to the board or can be additional sources contributing information to the board discussions. The channel is the board meeting. Noise includes many factors including uncertainty in the information presented to the board, distractions (i.e., side conversations, board members working other issues on email, text, or side discussions), or physical noise in the room or on phone lines. Following this structure, a board can be modeled as illustrated in Fig. 2.

This model provides for the inclusion of the cognitive aspects of the board members. Each board member must present information in a clear and understandable manner. The extent of their skill in this is represented by the encoding of the knowledge that they possess. In addition, the decision to share or withhold information is a cognitive aspect of the board member. Similarly, the ability of each board member to understand what is being discussed is represented by the decoding of the information (understanding). Many cognitive factors influence the decoding (understanding) of the information including education background, experience, intuitive ability, etc. Cognitive science, then, can be used to establish the distribution functions for the knowledge, encoding, and decoding of each board member and SME.

In this simplest form, the board model assumes that all information needed for a decision is provided to the board and that the information is properly and completely understood. Therefore, the uncertainty in the decision is zero, and the information entropy $H = \bar{I} = 0$. This does not mean that no information is conveyed by the board, but that there is no uncertainty in the board decision. The information is fully sent and received with $p_n = 1$ such that the $\log_n p_n = 0$ leading to the average and uncertainty as zero. In this simple model, the uncertainty (or absence of) is absolute in the sense that the decision is fully understood and is not subjective. While, there are simple decisions in practice, most decisions in practice involve various types of uncertainty in the decision making. Understanding the decision outside the board is not addressed in this model and can lead to uncertainty in the larger context as well.

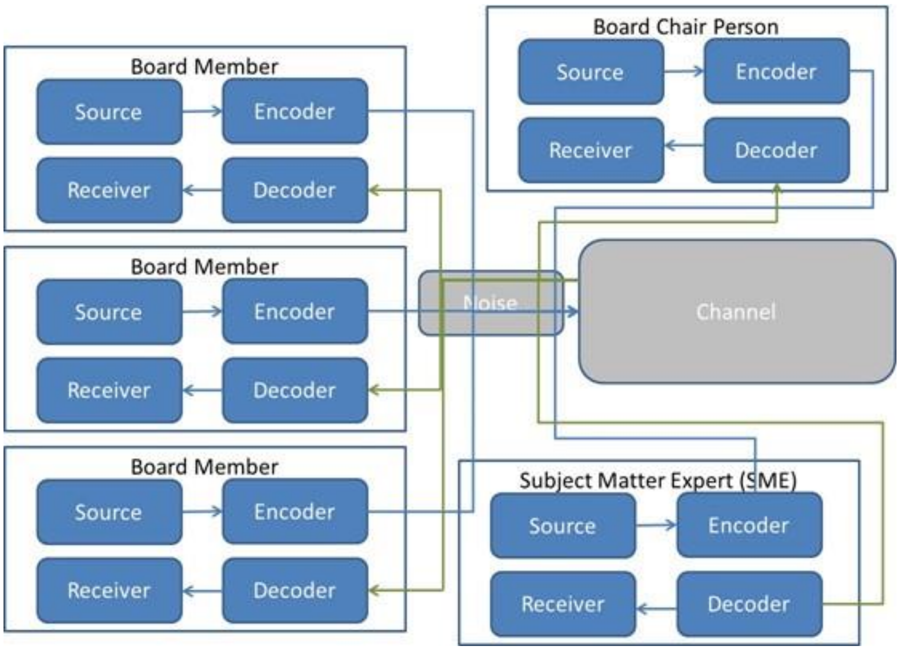

Fig. 2. Information Theory Board Model

There are many sources of uncertainty in board decisions. These include hidden (or withheld) information, cultural biases (creating blind spots on certain topics or ignoring factors), ignorance (not understanding aspects of the topic)[iv], and missing information in the board discussion.

Decision boards, as decision making bodies, are chartered with controlling a particular program, project, system, etc. As such, control theory applies to the basic functions of a board. Boards can be modeled as a Finite Impulse Response (FIR) system. Each board member comes to the board with information on a given topic. This information is cognitively processed forming preferences (i.e., weightings), relationships with other information, etc. These cognitive processing functions are quite complex. The board member then communicates with other board members during the board meeting and adds this information with their initial thoughts to create or modify their position. Thus, each board members thought processes can very simply be modeled as a cascade filter with feedback as shown in Fig. 3.

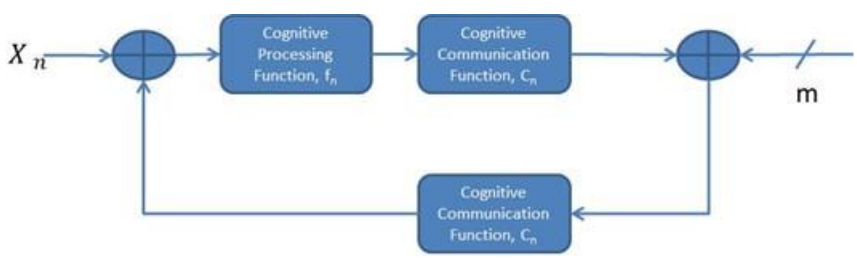

Fig. 3. Board Member Cognitive Processing Model

When all the board members and SMEs are combined, the board meeting then becomes a cascade filter model. Fig. 4 illustrates the control theory model of board operation. In this representation, the information theory model relationship is clearly seen, where the addition of the information of the board members and SME during discussion is the channel and noise is injected into the channel from external disturbances.

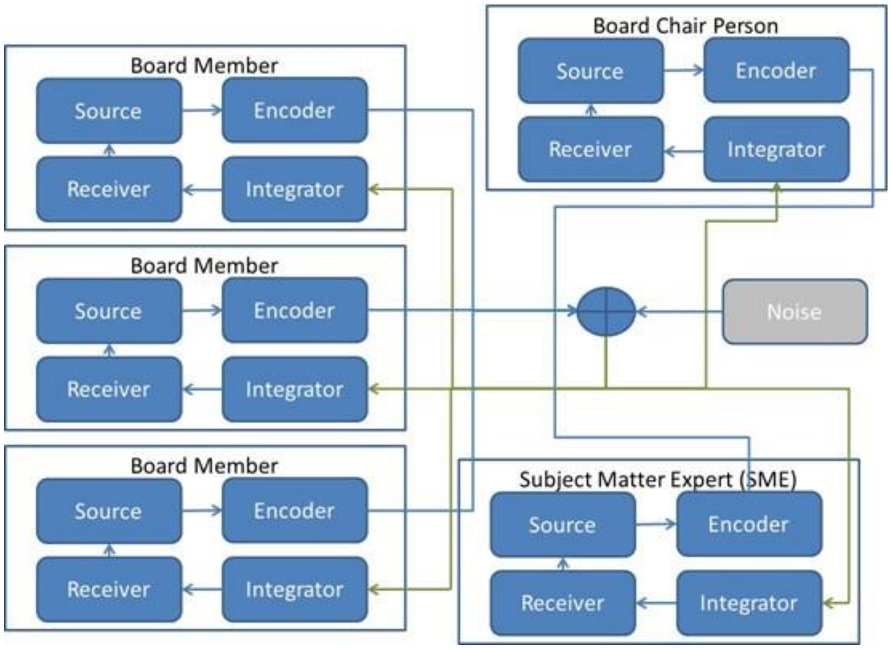

Fig. 4. Control Theory Board Model with Information Theory Board Member Representation

The board model can be updated with the board member decision making model where the Encoder is one form of Cognitive Communication Function and the Decoder is another form of Cognitive Communication Function. The Source and Receiver are combined as part of the Cognitive Processing Function, and X_n is contained in memory. Fig. 5 illustrates this model.

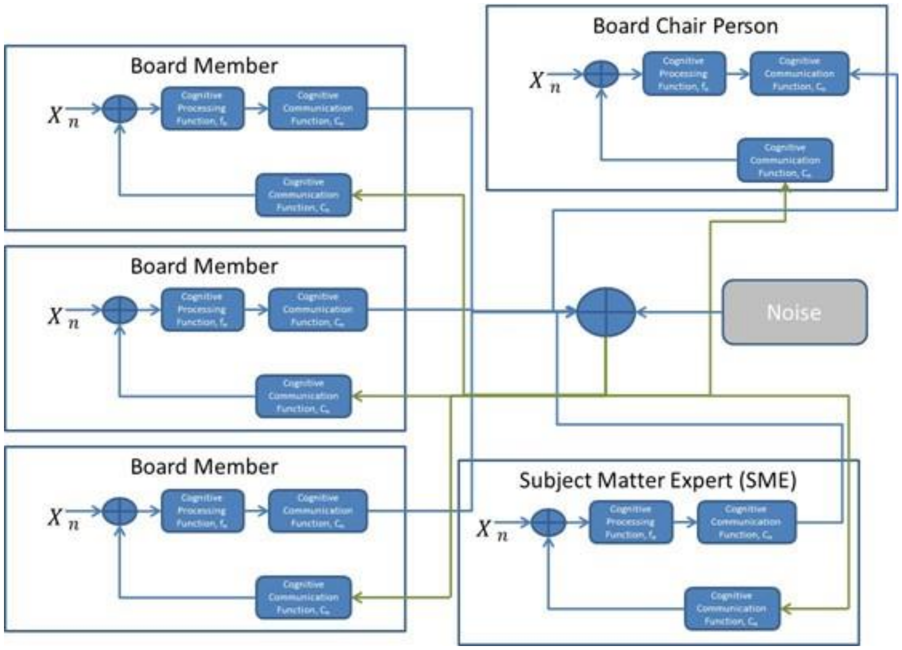

Fig. 5. Control Theory Board Model with Cognitive Functions

The equation represented in Fig. 5, can be written as:

$$Y_n = X_n + \sum_q f_{p,q+1}\{C_p[f_{p,q}(X_n) + \sum_{m.m\neq p} C_m[f_m(X_n)] + Noise]\} \quad (3)$$

where X_n represents a specific piece of information, the subscript, p, represents a specific board member or SME, q is the number of iterations in the board discussion, and the sum is over m, the total board members and SMEs participating in the decision. Equation 3 then represents the decision, Y_n, reached by the decision-making board member with inputs from the other board members and SMEs. This model assumes all board members and SMEs start with the same basic information, X_n. It allows understanding of the information to vary among the board members represented by the function, $f_{p,q}$. In this model, if a board member or SME has no knowledge of the topic (i.e., ignorance of the subject), $f_{p,1}(X_n)=0$. Similarly, if the board member withholds information on the topic, $C_p[f_{p,q}(X_n)] = 0$. Beyond this, the function, $f_{p,q}$ represents the level of understanding on a subject. Similarly, a decision to not share information is represented by this function as well. This function also encompasses preconceived ideas on the given information, preferences (personal or shared), intuition, deductive reasoning, and inductive reasoning. Clearly, the form of this function is complex. Some recent work in cognitive science[v,vi] may provide improvements in this representation.

Using this control theory based model, the transfer function of the board can be represented as the ratio of the initial understanding of the information divided by the final decision as shown in Equation 4:

$$T_n = \frac{X_n + f_{p,q}(X_n)}{X_n + \sum_q f_{p,q+1}\{C_p[f_{p,q}(X_n) + \sum_{m.m\neq p} C_m[f_m(X_n)] + Noise]\}} \quad (4)$$

This transfer function provides a model of the cognitive information processes as a beginning point to incorporate cognitive science models.

There are other information sources which can contribute to this model similar to SME inputs. These include text messages and emails to board members, personal side discussions (which also contributes to noise and affects the intake of other information). Since these inputs do not go to the whole board, but rather to individual members, and the external SME (particularly in electronic communication) may not be receiving all the board discussion, they have a function $C_s[f_s(X_n)]$, where s represents a specific SME and there is no iteration with the board discussion, q. The inputs are single events since the external SME is not part of the board discussion.

2. Multiple Board Structures

A question often asked, is what is the most efficient board structure? Will a single board suffice or are multiple boards more efficient? This has been a difficult question to answer. The set theory view of information theory provides the answer to this question.

A range that is too small (missing expertise) cannot be properly mapped to a decision domain for the intended outcome of the system. If this range can be mapped, then the missing expertise is not necessary for the decision in the context of the system. This immediately tells us that our board must have the right distribution of expertise for the system context and is therefore system specific. In addition, the information uncertainty increases as the information is distributed among separate boards as shown by the relation:

$$H(S,D,X,Y,Z) \leq H(S) + H(D) + H(X) + H(Y) + H(Z) \quad (5)$$

Information theory provides additional keys to understanding the board membership. Partitioning of information entropy, H, can only increase the uncertainty in the system by the relation,

$$H(p_1, p_2, \dots, p_n, q_1, q_2, \dots q_m) \geq H(p_1, p_2, \dots, p_n). \quad (6)$$

Thus, as more members are added to the board (additional members are represented by q), more uncertainty is created in the decision. This is balanced by range mapping being complete. Thus, the board structure needs to have only those members necessary for the system decisions (satisfying the mapping condition) and no more (minimizing H). This applies to a single board structure or to a multiple boards structure.

Within the set theory view of information theory, the board can be split (or delegated) if the information needed in one board is different than that needed in any other board. Then,

$$I_A \not\subset I_B \text{ and } I_B \not\subset I_A \quad (7)$$

so, there is no intersection of the information needed by the board and the board's domain (scope) can be different for each board.

When there is scope overlap, then, I_A ∩ I_B, and the boards cannot be separated. In this case

$$I_A \subset I_B \text{ and/or } I_B \subset I_A\ . \quad (8)$$

3. Statistical Properties of Boards

Splitting a board into multiple boards where there is significant overlap greatly increases the information uncertainty, H, in the board structure as shown in Equation 5. To examine this, we need to start with the characteristics of the uncertainty, or entropy, function itself. There are four (4) axioms the information entropy must meet:

3.1. Continuity

$$H(p_1, p_2, \dots, p_n) \quad (9)$$

is continuous in all p_n. Thus, there are no discontinuities in the information probabilities. This means, as noted earlier, that the range maps completely to the domain within the board. Discontinuities lead to highly uncertain, or in some cases blind, decisions. A robust board has all disciplines (i.e., affected or contributing parties) represented. This satisfies the range to domain mapping criteria and the related Continuity property.

3.2. Symmetry

$$H(p_1, p_2, \dots, p_n) = H(p_2, p_1, \dots, p_n). \quad (10)$$

Thus, the order of uncertainty does not contribute to the uncertainty in the decisions. This must be distinguished from temporal order of information sharing leading to a momentary information void on a subject until all aspects are explained for understanding. The process of understanding is always assumed to be complete in this model, and symmetry holds for a complete understanding of a subject. The order in which you discuss or think of a subject does not matter if you fully understand the subject.

3.3. Extrema

$$\boldsymbol{Max}[\boldsymbol{H}(\boldsymbol{p_1}, \boldsymbol{p_2}, \ldots, \boldsymbol{p_n})] = \boldsymbol{H}(\tfrac{\boldsymbol{1}}{\boldsymbol{n}}, \tfrac{\boldsymbol{1}}{\boldsymbol{n}}, \ldots, \tfrac{\boldsymbol{1}}{\boldsymbol{n}}). \quad (11)$$

The maximum uncertainty arrives when all decisions are equally uncertain. If any single decision can be distinguished from the others, then the uncertainty to choose or not choose that option is smaller. Similarly, if no options satisfy the decision criteria, then the board has no information on which to base a decision leading to

$$\boldsymbol{Min}[\boldsymbol{H}(\boldsymbol{p_1}, \boldsymbol{p_2}, \ldots, \boldsymbol{p_n})] = \boldsymbol{H}(\boldsymbol{0}, \boldsymbol{0}, \ldots, \boldsymbol{0}) = \boldsymbol{0}. \quad (12)$$

3.4. Additivity

If a probability of occurrence, p_n, can be subdivided into smaller segments, q_k, then the uncertainty can be represented as

$$H(p_1, p_2, \ldots, p_{n-1}, q_1, q_2, \ldots q_k) = H(p_1, p_2, \ldots, p_n) + p_n H\left(\frac{q_1}{p_n}, \frac{q_2}{p_n}, \ldots \frac{q_k}{p_n}\right). \quad (13)$$

3.5. Principles

Following from these 4 properties, information can be subdivided during discussions if all the information is presented (i.e., all q_k is present in the discussion) without affecting the uncertainty of the decision. Note that this requires all information to be present. Subdividing boards and segmenting the information does not meet this criteria and results in higher uncertainty.

These four properties and five equations (Equations (9) – (13)) from information theory provide guidance in the structuring of boards. These relationships indicate that impossible solutions do not affect the information entropy. These solutions do not fit in the domain of the solution and cannot be mapped from the range of the original decision question. In addition, the continuity of H requires all information be present for a decision. While a decision may be made with missing information, the decision is not actually addressing the original question. The question essentially changes, when all information is not present, and the decision addresses a different question than the one intended.

Information theory assumes a statistical basis of the information. Before we proceed further, we need to establish the statistical nature of boards, not that they are predictable, but that their underlying operations can be represented statistically.

There are four principles that establish the statistical nature of a decision board:

1. Uncertainty exists in complex decisions. In these cases, simplifying assumptions lead to a lower understanding of the decision intricacies and a higher uncertainty (not always recognized) in the decision process. Interactions among differing factors in complex decisions have dependencies that are not recognized (ignorance)[iv] or not well understood. Missing information is not always easily recognized. Factors not considered important in the decision can end up driving the system. Missing information comes from events (physical, chronological, or fiscal) not recognized as relating to the decision, unknown environments in which a system operates, unrecognized dependencies, and cultural biases (e.g., politics).
2. The uncertainty of which option is best collectively, and in some cases individually, leads to a statistical representation of which answer is best. In a board decision, the board vote is a statistical event with a distribution of yes and no positions. This is tied back to the cognitive functions. This statistical function is then combined with other statistical functions (i.e., other board members and SMEs) to produce a decision based on these functions.
3. The potential for misunderstanding (i.e., error) is also statistical. This includes miscommunication (not stating clearly what is meant or not understanding clearly what is stated (and therefore meant). These lead to unintended consequences in the decision-making process. These unintended consequences can be social, physical, chronological, fiscal, or environmental.
4. Cultural and Historical bias lead to sub-optimal decisions. Large social population actions form the basis for these biases and the effects on a person's cognitive information processing function, f_n, are statistical in nature.

Decisions can be represented statistically with various distribution functions depending on the individual preferences, biases, knowledge, and experience with the subject as discussed in the control theory model above. The cognitive processing functions, based on the properties of H, should fulfill continuity, symmetry, extrema, and additivity.

4. Information Bounds in the Board Context

In the board context, the board discussion forms the information channel as discussed above. The board members and SMEs are both information sources and sinks as modeled in Fig. 4. Information theory treats communication as the transmission of symbols. Natural language, where letters form words, words form sentences, and the order of the symbols and words are important in interpretation fits this model perfectly. And the board discussion is the channel where this information is transmitted between the board members and SMEs.

Information theory models the transfer of information through the board channel very well. A definition of terms is convenient at this point.

$H(X_n)$ is the average information shared by a single board member or SME as defined in Equation 2.

$H(Y_n)$ is the average information received by a single board member or SME also following the definition in Equation 2.

$H(X_n,Y_n)$ is the joint probability that what was shared by one member and heard by another (the average uncertainty in the total transmission through the board channel).

$H(Y_n|X_n)$ is the probability that one member actually heard what was stated by another. This brings in the effects of noise (and misunderstanding) in the channel. This focus is on the receiver of the information.

$H(X_n|Y_n)$is the equivocation probability that one member actually stated what was heard by another. This brings in the effects of recovery (or proper understanding) of the information sent and is a measure of how well the information is understood by the receiving member.

If the board discussion is clear, and no misunderstanding, is present, then the information provided by the speaker is accurately received by the listener (receiver). The information is perfectly transferred and Information Theory tells us that,

$$I(X;Y) = H(X,Y) = H(X) = H(Y). \tag{14}$$

Now, if there is complete confusion, then what is stated is not related to what is heard. This is the case where the received information is independent of the transmitted information and,

$$I(X;Y) = 0. \tag{15}$$

In this case, no information is transmitted through the channel (i.e., discussion). These two extremes, perfect transmission and no transmission, provide bounds on the information sharing in a board meeting. Typically, neither of these conditions is achieved and there is always some noise or misunderstanding during the discussion that limits the amount of information transferred among the board members.

5. Information Theory Representation of a Board

Set theory provides the mathematical basis for information theory which fits the board structure well. Information shared in a board discussion is the sum of all the information provided by the individual board members. This is illustrated in Fig. 6 for the example board structure used in Fig. 4.

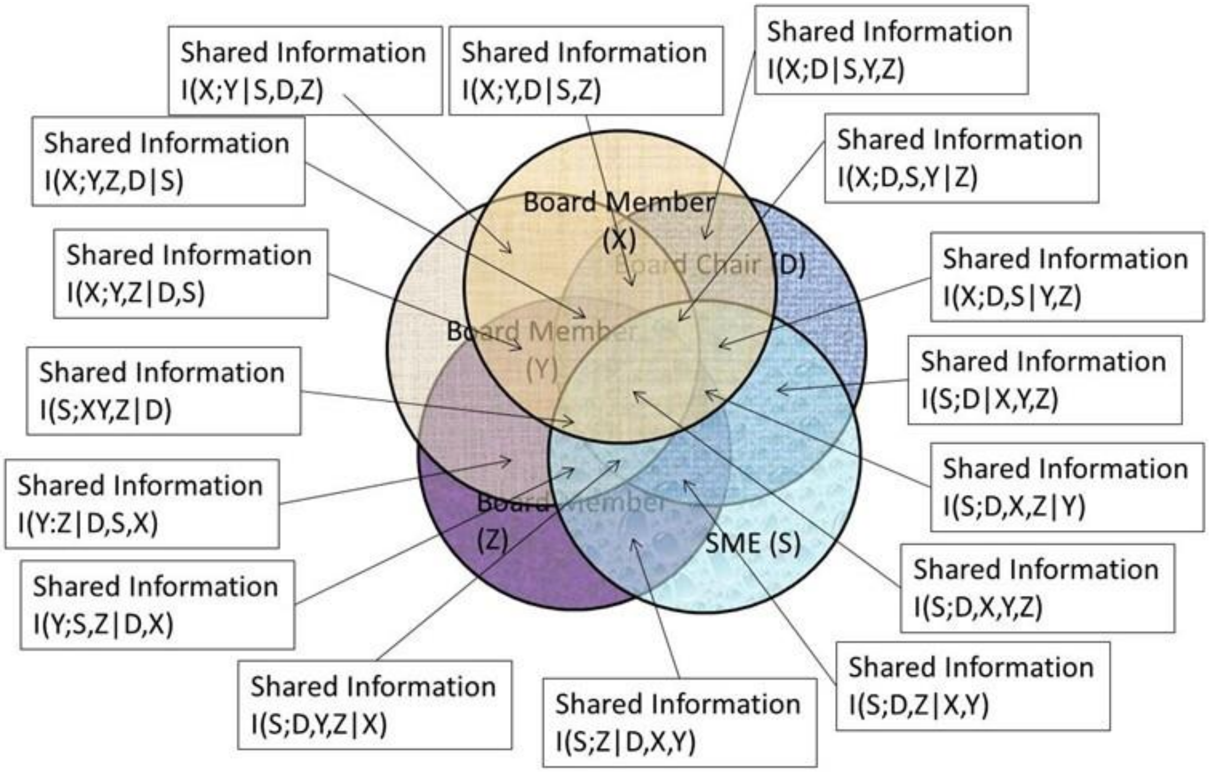

Fig. 6. Set Theory Representation of Board

This picture is somewhat complex in that there are many different areas of shared information. Note that the symbol, |, is read as "and not" so that I(S;D,X|Y,Z) is the information shared between the SME,S, the Board Decision Maker or Chair, D, Board Member, X, and not Board Members Y and Z. For a decision to be fully informed, the information for the decision must be contained in the center most ellipsoid, I(S;D,X,Y,Z). This represents the set of all information shared and received in the board discussion. Other information is shared based on the knowledge of individual board members and the SME, the ability of each to understand the information, and individual distractions. This can lead to board discussions which do not fully incorporate all board member knowledge. All permutations of this case are represented in the figure except for I(Y:D,Z|S,X), I(S:X,Z|D,Y), I(S;X,Y|D,Z), I(X:D,Z|Y,S), I(S;X,Z|Y,D), which is an artifact of the figure geometry (where non adjacent sets cannot be shown as excluded).

Information theory represents this as shown by Equation 5 above:

$$H(S,D,X,Y,Z) \leq H(S) + H(D) + H(X) + H(Y) + H(Z)$$

where H(S), H(D), H(X), H(Y), and H(Z) are how well the board members and SME communicate their information. This indicates that the sum of information can be no more than that provided by each of the members. Noise (distractions, misunderstanding, poorly stated (poor transmission)) and information not shared (intentional, unintentional, missing board member) invokes the inequality in the relationship.

Following the work of Reza[iii], set theory can relate the rules for information. This yields the following relationships:

$$I(X;Y) = f(X \cap Y) \tag{16}$$

which is the expected value of mutual information shared in the discussion. In set theory, this is a function of the intersection of the information held by X and Y.

$$H(X,Y) = f(X \cup Y) \tag{17}$$

which is the average uncertainty of the discussion. This is a function of the union of the information available.

$$H(X|Y) = f(XY') \tag{18}$$

which is the information received by X given the information that Y shared. This is the probability that the board understood the information shared by Y. Note, in set theory this is a function of the information X has that Y does not.

$$H(Y|X) = f(YX') \tag{19}$$

which is the information shared by Y given the information that X heard. This is the probability that the board understanding is what was shared by Y. Note, in set theory this is a function of the information Y has that X does not.

From these relationships, then, perfect understanding occurs when f(X) = f(Y) and both parties understand the information fully. When there is, no information shared $I(X;Y) = f(X \cap Y) = 0$. Thus, there is no intersection of the information sets and no common understanding. In the board example used above $\mathrm{I(S;D,X,Y,Z)}. = f(S \cap D \cap X \cap Y \cap Z)$ and the shared information is represented in Fig. 6 by the intersection of the 5 circles representing the knowledge to share for each decision.

In these representations, H(X), etc. represents the uncertainty in the information shared by board member X. This uncertainty stems from the board members understanding (or knowledge) of the decision requested and the associated decision factors, cultural bias (which indicates if information will be shared or withheld), and personal comfort in sharing specific information or engaging in debate about the information.

Channel capacity (i.e., board capacity) in information theory is:

$$C = max\big(I(X;Y)\big) = max(f(X \cap Y)). \tag{20}$$

Thus, channel capacity (i.e., the board capacity) for a decision is defined by the mutual information, or the intersection of information, shared in the board discussion. The maximum board capacity then is based on the intersection of knowledge held by each board participant. The intersection represents the integration of individual board participant's knowledge to form a decision. Note, this indicates that if a board is segmented, and required knowledge for a decision is not present, then the board does not have the information necessary to decide. A decision can be made, but the scope of the decision does not address the actual question being considered. This results in unintended consequences for the decision because the board does not have all the facts.

One implication of this capacity is that a board with a missing member(s) will have a lower capacity since mutual information for the topic will be reduced. Similarly, adding a member that has largely overlapping knowledge can create disjoint relationships where the two members approach the topic differently (based on their differing cognitive functions), do not overlap in their understanding, and I(S;D,X,Y,Z) = 0 blocking the board decision.

6. Summary

Information theory provides a rich mathematical structure to model and understand decision making structures within systems engineering and project management. The models produced by information theory provide guidance in the structuring of boards, both their membership and their relationship to other board structures. Since boards are established to control the system or project, control theory provides a model of the individual decision making process. This provides a starting point for the future addition of cognitive processes in the decision making model. Set theory understanding of information theory is also helpful in understanding the relationships between board members and the information that they share among the board.

References

[i] Tishby, N., Polani, D., "Information Theory of Decisions and Actions", accessed 16 Feb 2018 at http://www.cs.huji.ac.il/labs/learning/Papers/IT-PAC.pdf .

[ii] Dorpinghaus, M., Roldant, E., Nerit, I., Mery, H., Julicher, F., "An information theoretic analysis of sequential decision-making", IEEE International Symposium on Information Theory (ISIT), Aachen, Germany, 2017.

[iii] Reza, Fazlollah M., “An Introduction to Information Theory”, Dover, NY, 1994.

[iv] Merton, Robert K., “The Unanticipated Consequences of Social Action (1936)”, in Social Structure and Science, ed. by Piotr Sztompka, University of Chicago Press, 1996, pp. 173-182.

[v] Fan, J., “An information theory account of cognitive control”, Front. Hum. Neurosci., 02 September 2014

[vi] Thornton, C, “A New Way of Linking Information Theory with Cognitive Science”, accessed 16 Feb 1018 at https://pdfs.semanticscholar.org/c944/20f28388bc053fa9e1170e547e51aa4fc6c2.pdf.