

# Categorical Analytics Based On Consensus Theory

Harry Katzan, Jr., Savannah State University, USA

## ABSTRACT

*Categorical analytics is an admixture of computational methods with the express purpose of facilitating the multifaceted process of unstructured decision making. The complex subject is based on consensus theory and includes structured analytics, categories, entropy, and the combination of evidence. The methodology is applicable to a wide range of business, economic, social, political, and strategic decisions. The paper includes an election application to demonstrate the concepts.*

**Keywords:** Consensus theory, uncertainty, information theory, category, entropy, evidence, combination of evidence, structured analytics, Dempster's rule.

## INTRODUCTION

A category is a means of structuring a problem domain so that relevant information may be assessed and unstructured decisions can be made. The objective is to engage in a predictive modality in which one or more future events may be identified and analyzed.

Let  $C_i$  be one of the categories used to stratify the problem domain such that the collection

$$C = \{C_1, C_2, \dots, C_n\}$$

represents a complete conceptualization of the dynamics under investigation and  $n$  is the number of categories.

Associated with each category is a set of probabilities representing an assessment of a future outcome based on its underlying categorical imperative. Thus, a category is a mechanism for isolating a single view of the problem under consideration. The ontological definition of a *category*, as a conceptual entity with no attributes in common with other categories, is adopted in this paper. The mutually exclusive set of possibilities under investigation is known as the frame of discernment.

Consensus theory is covered first followed by a presentation of an approach to the structural analysis of categories. Finally, an election application is used as a demonstrative example that gives some insight into how the methods can be applied to other problems.

## BASIC CONCEPTS

A *frame of discernment* (hereafter referred to as simply a *frame*) is a set of mutually exclusive and collectively exhaustive possibilities for the value of an element of a system within a particular problem domain (referred to as the *domain*). Three obvious examples of frames are:

$$\begin{aligned} E &= \{\text{McCain, Obama}\} \\ M &= \{\text{Up, Unchanged, Down}\} \\ J &= \{\text{Guilty, Deadlock, Innocent}\} \end{aligned}$$

In general, a frame is a means of representing the possibilities under consideration.

Clearly, the elements in a frame are in fact propositions that can be interpreted in the usual fashion as events or states. Thus, if component  $s_i$  of system  $S$  over domain  $V$  were associated with the symbol “McCain,” then that state is equivalent to the proposition, “The true value of  $V$  for component  $s_i$  is *McCain*,” or in ordinary language, “ $s_i$  prefers or has voted for *McCain*.”

The set  $S$  of propositions  $S_i$ ,

$$S = \{S_1, S_2, \dots, S_n\}$$

commonly represents the collection of states of a component of a system under analysis. Clearly, at an agreed upon point in time, one proposition is true and the others are false. If proposition  $S_i$  is true, then we can assert that

$$\begin{aligned} S_i &= T \text{ and} \\ S_j &= F, \text{ where } i \neq j \end{aligned}$$

where  $T$  and  $F$  represent “true” and “false,” respectively.

**Uncertainty**

Prior to an agreed point in time ( $\tau$ ), we obviously do not know the state of the system under analysis or its components with any degree of certainty. The expectation that a part of the system will be in a particular state at time  $\tau$  is denoted by a real number  $p(S_i)$  associated with each of the propositions in the frame  $S = \{S_i\}$ ,  $i=1,2,\dots,n$ , such that

$$\begin{aligned} &0 \leq p(S_i) \leq 1 \\ \text{and} \\ &\sum_{i=1}^n p(S_i) = 1 \end{aligned}$$

It follows that the function  $p$  can be extended to the lattice  $\underline{S}$  of  $S$  such that

$$p(k) = \sum_{i=1}^k p(S_i) = 1$$

where  $k$  is an element of  $\underline{S}$ . This is simply the addition rule for mutually exclusive events.

**Information Theory**

If the likelihood of an event is relatively remote and it happens to turn out to be true at an appointed time, we are surprised. Similarly, if the chance of the same event is very likely and it comes out true, we are not overly surprised. The extent of our surprise represents our “information gain” in the sense that if we knew that information, we wouldn’t be surprised. This is an exceedingly profound principle, although it is simple in concept, and takes the notion of information from the subjective plane to an objective level.

At an appointed time our knowledge of a situation can change dramatically based on “new” information. The difference in information before and after gives us the information content of the message giving us the outcome.

Let  $x$  be the probability that some event  $E$  will occur, where  $0 \leq x \leq 1$ . Assume further that a message is received that  $E$  has occurred. If  $x=0.99$ , you are not surprised and the message has little information content. On the other hand, if the value of  $x$  were 0.01, you are very surprised and the message would indeed have high information content. It would be reasonable to require that the information content  $h(x)$  of a message be a decreasing function of the probability  $x$ . It is generally agreed to take the logarithm of the reciprocal of the probability  $x$  as the decreasing function, such that

$$h(x) = \log(1/x) = -\log x$$

In general, the base of the logarithm is open but base 2 is common in information theory and information can be measured in bits.

Example. Assume there are  $2^N$  possible events and all have the same chance of occurrence  $2^{-N}$ . The amount of information contained in the message that states the particular event that occurred is:

$$h(2^{-N}) = \log(1/2^{-N}) = N \text{ bits}$$

When the probability is not a power of two, then the information content will take on non-integer values.

**Entropy**

Consider a complete system  $S = \{S_1, S_2, \dots, S_n\}$  with corresponding probabilities  $x_1, x_2, \dots, x_n$ , such that

$$\sum_{i=1}^n x_i = 1$$

where  $0 \leq x_i \leq 1$  for  $i=1, 2, \dots, n$ .

When we receive a message that  $S_i$  has occurred, the information content of the message is  $h(x_i) = -\log(1/x_i)$ . Before the message is received, we do not know the event that will occur and the magnitude of the corresponding information content. On the other hand, we can develop the average information content (i.e, the *expected information content*  $H(x)$ ) as:

$$H(x) = \sum_{i=1}^n x_i h(x_i) = \sum_{i=1}^n x_i \log(1/x_i) = -\sum_{i=1}^n x_i \log(x_i)$$

where

$$x_i \log(x_i) = 0 \text{ if } x_i = 0$$

The expected information content of a probability distribution is a measure of its disorder. In a distribution with  $n$  probabilities, the expected information approaches a maximum value as the  $n$  probabilities  $x_i$  approach  $1/n$ . The expected information as a measure of disorder is commonly known as *entropy* (Theil, 1967).

EXAMPLE. Consider the system  $S = \{S_1, S_2\}$  with probabilities

$$p(S) = \{x_1, x_2\} = \{0.6, 0.4\}$$

The entropy of  $S$  is computed as

$$\begin{aligned}
 H(S) &= \sum_{i=1}^2 p(x_i)h(x_i) \\
 &= (0.6)\log(1/0.6)+(0.4)\log(1/0.4) \\
 &= 0.970951
 \end{aligned}$$

The entropy is high because the possibilities approach being equally likely.

**CONSENSUS THEORY**

Consensus theory is a methodology for combining evidence based on Dempster-Shafer theory (Shafer [1976], Katzan [1992, 2006]) and the mathematical combination of evidence (Dempster [1967]). Dempster-Shafer theory has commanded a considerable amount of attention in the scientific and business communities, because it allows a knowledge source to assign a numerical measure to a proposition from a problem space, and provides a means for the measures accorded to independent knowledge sources to be combined. Dempster-Shafer theory is attractive because conflicting, as well as confirmatory, evidence from multiple sources may be combined.

The basis of Dempster-Shafer theory is the frame of discernment ( $\Theta$ ), introduced previously. Accordingly, a knowledge source may assign a numerical measure to a distinct element of  $\Theta$ , which is equivalent to assigning a measure of belief to the corresponding proposition. In most cases, the numerical measure will be a basic probability assignment. A measure of belief may also be assigned to a subset of  $\Theta$  or to  $\Theta$  itself.

**Support Functions**

Consider a frame of discernment  $\Theta$  and its power set denoted by  $2^\Theta$ . For example, given the frame:

$$\Theta = \{a, b, c\}$$

The power set is delineated as follows:

$$\begin{aligned}
 2^{\Theta} = & \{ \{a, b, c\}, \\
 & \{a, b\}, \{a, c\}, \{b, c\}, \\
 & \{a\}, \{b\}, \{c\} \}
 \end{aligned}$$

In Dempster-Shafer theory, a knowledge source apportions a unit of belief to an element of  $2^\Theta$ . This belief can be regarded as a mass committed to a proposition and represents a judgment as to the strength of the evidence supporting that proposition. When viewed in this manner, evidence focuses on the set corresponding to a proposition; this set is called a *focal set*.

The support for a focal set is a function  $m$  that maps an element of  $2^\Theta$ , denoted by  $A$ , onto the interval  $[0,1]$ . Given a frame of discernment  $\Theta$  and function  $m: 2^\Theta \rightarrow [0,1]$ , a support function is defined as follows:

$$\begin{aligned}
 m(\phi) &= 0, \text{ where } \phi \text{ is the null set} \\
 0 &\leq m(A) \leq 1, \text{ and}
 \end{aligned}$$

$$\sum_{A \subset 2^\Theta} m(A) = 1$$

The support function  $m$  is called a *basic probability assignment*, which is assigned by the knowledge engineer or domain specialist.

A support function is called a *simple support function* if it reflects, at most, one focal set not equal to  $\Theta$ . A simple support function assigns a measure of belief to the focal set  $A$ , as follows:

$$\begin{aligned} m(A) &> 0 \\ m(\Theta) &= 1 - m(A) \\ m(B) &= 0, \text{ for all } B \subset \Theta \text{ and } B \neq A \end{aligned}$$

The simple support function for a focal set  $A$  assigns a portion of the total belief exactly to  $A$  and not to its subsets or supersets. The remainder of the belief is assigned to  $\Theta$ . Because certainty function must add up to 1,  $m(\Theta) = 1 - m(A)$ .

It is possible that a body of knowledge or evidence supports more than one proposition, as in the following case. If

$$\begin{aligned} \Theta &= \{a, b, c, d\} \\ A &= \{a, b\} \end{aligned}$$

and

$$B = \{a, c, d\}$$

then the evidence supports two focal sets, which in the example, are  $A$  and  $B$ . If  $m(A) = 0.5$  and  $m(B) = 0.3$ , then  $m(\Theta) = 0.2$ . A support function with more than one focal set is called a *separable support function*. Separable support functions are normally generated when simple support functions are combined.

The notion of combining simple support functions is a practical approach to the assessment of evidence. An analyst obtains information from a knowledge source, and it leads to an immediate conclusion – not with certainty, but with a certain level of belief. This is a normal straightforward means of handling human affairs and is precisely what people do. Then when additional information comes in, the various pieces of evidence are combined to obtain a composite picture of the situation.

**Combination of Evidence**

A method of combining evidence is known as Dempster’s rule of combination (Dempster [1967]). Evidence would normally be combined when it is obtained from two different observations, each over the same frame of discernment. The combination rule computes a new support function reflecting the consensus of the combined evidence.

If  $m_1$  and  $m_2$  denote two support functions, then their combination is denoted by  $m_1 \oplus m_2$  and is called their *orthogonal sum*. The combination  $m_1 \oplus m_2$  is computed from  $m_1$  and  $m_2$  by considering all products of the form  $m_1(X) \bullet m_2(Y)$ , where  $X$  and  $Y$  range over the elements of  $\Theta$ ;  $m_1(X) \bullet m_2(Y)$  is the set intersection of  $X$  and  $Y$  combined with the product of the corresponding probabilities.

For example, consider the frame of discernment

$$\Theta = \{\text{healthy, test, sick}\}$$

and focal sets  $A$  and  $B$ , based on two different observations over the same frame:

$$\begin{array}{ll} A = \{\text{test, sick}\} & B = \{\text{healthy, tests}\} \\ m_1(A) = 0.8 & m_2(B) = 0.6 \\ m_1(\Theta) = 0.2 & m_2(\Theta) = 0.4 \end{array}$$

The combination rule is depicted via the following tableau:

	A	$\Theta$
$m_1$	{test,sick}	{healthy,tests,sick}
$m_2$	0.8	0.2
$B$		
{healthy,tests}	{tests}	{healthy,tests}
0.6	0.48	0.12
$\Theta$		
{healthy,test,sick}	{tests,sick}	{healthy,test,sick}
0.4	0.32	0.08

The entries are then combined as follows:

$$\begin{aligned}
 m_1 \oplus m_2(\{tests\}) &= 0.48 \\
 m_1 \oplus m_2(\{healthy,tests\}) &= 0.12 \\
 m_1 \oplus m_2(\{tests,sick\}) &= 0.32 \\
 m_1 \oplus m_2(\{healthy,tests,sick\}) &= 0.08 \\
 m_1 \oplus m_2 &= 0 \text{ for all other subsets of } \Theta
 \end{aligned}$$

Thus, for  $A_i \cap B_j = A$  and  $m_1 \oplus m_2 = m$ , the combination rule is defined mathematically as:

$$m(A) = \frac{\sum_{A_i \cap B_j = A} m_1(A_i) \bullet m_2(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i) \bullet m_2(B_j)}$$

The denominator reflects a normalization process to insure that the pooled values sum to 1. So, in this instance, the combination rule yields

$$A \oplus B = \{ \{tests\}, 0.48 \}, \{ \{healthy,tests\}, 0.12 \}, \{ \{tests,sick\}, 0.32 \}, \{ \{healthy,tests,sick\}, 0.08 \}$$

This result states that the combined evidence from knowledge sources A and B gives 0.48 as the likelihood of the proposition “tests” and 0.12 as the likelihood of proposition “healthy OR tests,” and so forth.

As another example, consider the same frame of discernment

$$\Theta = \{ healthy, test, sick \}$$

and views C and D, based on two different observation over the frame:

$$\begin{aligned}
 C &= \{ \{healthy\}, 0.6 \}, \{ \{tests\}, 0.3 \}, \{ \{sick\}, 0.1 \} \\
 D &= \{ \{healthy\}, 0.4 \}, \{ \{tests\}, 0.4 \}, \{ \{sick\}, 0.2 \}
 \end{aligned}$$

The combination rule is depicted via the following tableau:

		<i>C</i>		
	$m_1$	{healthy}	{tests}	{sick}
$m_2$		0.6	0.3	0.1

---

		<i>D</i>		
	{healthy}	{healthy}	{∅}	{∅}
	0.4	0.24	0.12	0.04
	{tests}	{∅}	{tests}	{∅}
	0.4	0.24	0.12	0.04
	{sick}	{∅}	{∅}	{sick}
	0.2	0.12	0.06	0.02

The entries are then combined as follows:

$$\begin{aligned}
 m_1 \oplus m_2(\{\text{healthy}\}) &= 0.24 \\
 m_1 \oplus m_2(\{\text{tests}\}) &= 0.12 \\
 m_1 \oplus m_2(\{\text{sick}\}) &= 0.02 \\
 m_1 \oplus m_2(\{\emptyset\}) &= 0.62
 \end{aligned}$$

So, in this instance, the combination and normalization processes yield the result

$$C \oplus D = \{ \{\text{healthy}\}, 0.63 \}, \{ \{\text{tests}\}, 0.32 \}, \{ \{\text{sick}\}, 0.05 \}$$

after normalization by dividing the combined assessment by (1-0.62) or 0.38. Because the problem is well-structured, the representation can be simplified as

$$C \oplus D = \{0.63, 0.32, 0.05\}$$

where the elements exist in a one-to-one correspondence with the frame {healthy, test, sick}.

For views  $A = \{A_1, A_2, \dots, A_n\}$  and  $B = \{B_1, B_2, \dots, B_n\}$ , the combination rule can be simplified as

$$\begin{aligned}
 A \oplus B &= \{A_1 \times B_1 / k, A_2 \times B_2 / k, \dots, A_n \times B_n / k\} & [1] \\
 \text{where } k &= \sum_{i=1}^n A_i \times B_i
 \end{aligned}$$

We will refer to equation [1] as the *simplification rule*.

**STRUCTURAL ANALYTICS**

A problem domain is composed of categories, each of which is defined by a set of alternate structures. In a make-believe universe, for example, the category party could be defined as

$$\text{party} = \{\text{republican, democrat, independent}\}$$

based on a structural assessment, such as demographics. In this instance, the category party is one of many viewpoints of an underlying decision situation, which could be a vote in an election or on an important issue. We are going to argue that in many unstructured decision-making problems, the probabilistic outcome can be based on structural, rather than, preferential elements. What makes an unstructured decision so complex is that there are usually several categories “tugging at the decision maker.” We are going to show how categorical assessments can be combined to form a composite assessment of a decision under consideration. Through the technique of structural

analysis, we are going to assign probabilities to the elements of the frame of discernment from a given category, and then use consensus theory to combine the various probabilities. For example, a vote based on party could go one way and a vote based on gender could go another way. In reality, it would be a combination of the two views.

**Structural Elements**

Each category  $C_i$  is comprised of a set of structural elements  $S_i = \{S_{i1}, S_{i2}, S_{i3}, \dots, S_{imi}\}$ , where  $mi$  is the number of structural elements in category  $C_i$ . Consider a fresh example universe defined as:

$C = \{C_1, C_2\} = \{\text{color, size}\}$ , where  
 color = {white, red, green, blue}, and  
 size = {small, medium, large}

where  $S_1$  and  $S_2$  are defined respectively as

$S_{11} = \text{white}$	$S_{21} = \text{small}$
$S_{12} = \text{red}$	$S_{22} = \text{medium}$
$S_{13} = \text{green}$	$S_{23} = \text{large}$
$S_{14} = \text{blue}$	and $m2 = 3$
and $m1 = 4$	

Each problem domain is represented by a set of categories, each of which is a special lens into the underlying problem. Each category is defined as a set of structural components that define it. The categorical demographics in an election, for example, could be party, gender, age, and so forth. In the immediate example, the categories are COLOR and size.

**Structural Probabilities**

Each object selected at random from a universe of study possesses a probability of occurrence

$$\text{prob}(S_{ij}) \text{ where } \sum_{j=1}^{mi} \text{prob}(S_{ij}) = 1 \text{ for category } i \text{ and structural element } j \text{ in category } i.$$

Accordingly, for category  $C_i$  and its structure  $S_i$ , the probability set would be expressed as :

$$P_i = \{\text{prob}(S_{i1}), \text{prob}(S_{i2}), \dots, \text{prob}(S_{imi})\}$$

For example, the probability set for category #1 (color), could be

$$P_1 = \{0.2, 0.4, 0.3, 0.1\}$$

representing white, red, green, and blue respectively.

Each  $\text{prob}(S_{ij})$  represents the “probability of occurrence” in the universe of study of the structural elements of category  $i$ .

**Analytic Alternatives**

In this form of analysis, each category  $C_i$  has an associated probability set  $P_i$ . Each structural element has a corresponding probability  $\text{prob}(S_{ij})$  in  $P_i$ . That probability represents the likelihood that an object selected at random from category  $C_i$  would be  $S_{ij}$ . Another interpretation is that a value in  $P_i$  gives the proportion of the corresponding structural element in  $C_i$ . The following tableau gives a couple of examples:



Category	Structural Element	Probability
party	republican	0.4
party	democrat	0.5
party	independent	0.1
gender	male	0.4
gender	female	0.6

The structural probabilities, alternately regarded as structural proportions, give a means of describing the environment in which a decision is to take place. In an election, the environment would be the electoral demographics.

**Preference Set**

Each structural element is assigned a *preference set* over the frame of discernment from a knowledge source, such as a poll, survey, or historical document. The probabilities in the preference set are the decision variables. For example, we might know that **republican** prefers candidate M with probability 0.7 and candidate O with probability 0.3. The set {0.7, 0.3} is known as the preference set.

Thus, for each structural element  $S_{ij}$  for all categories, there exists a preference set

$$ps(S_{ij}) = \{p_{ij}(\Theta_1), p_{ij}(\Theta_2), \dots, p_{ij}(\Theta_t)\},$$

and

$$\Theta = \{\Theta_1, \Theta_2, \dots, \Theta_t\}$$

Clearly,  $\sum_{k=1}^t p_{ij}(\Theta_k) = 1$  for all  $i$  and  $j$ .

**Composite Probabilities**

Composite categorical probabilities for each element in the frame of discernment are computed by combining the structural probabilities and corresponding preference set as follows:

$$P(\Theta_{it}) = \sum_{j=1}^n ((\text{prob}(S_{ij}) \cdot ps(S_{ij})) \tag{2}$$

where the index  $i$  runs through the categories and the index  $t$  runs through the alternatives in the frame of discernment.

**Categorical Probabilities**

The composite probabilities represent a summation of the preference for each element of the frame of discernment for each category. The result is a set of independent categorical assessments of the problem domain from different viewpoints represented as probabilities, as follows:

$$C_i = \{P(\Theta_{i1}), P(\Theta_{i2}), \dots, P(\Theta_{it})\}$$

where  $t$  is the cardinality of the frame of discernment, as defined previously.

Using the simplification rule [1], we derive a combined assessment of categories  $C_i$  and  $C_j$  of the form  $C_i \oplus C_j$

So that if

$$C_1 = \{0.54, 0.46\}$$

and

$$C_2 = \{0.58, 0.42\}$$

Then

$$K = C_i \oplus C_j = \{0.62, 0.38\}$$

The evidence is complementary and that fact is demonstrated in the combined assessment.

**Entropy**

In applications with a greater cardinality, entropy can be used as a measure of the “degree of disorganization” among the alternatives, since it is often difficult to make an assessment from a large collection of values. The entropy of  $C_1$  is 0.995378 and the entropy of  $C_2$  is 0.981454, and the entropy of  $K$  is 0.958042, representing a lesser degree of disorganization for the above examples.

**ELEMENTARY APPLICATIONS**

Two applications of the preceding concepts are considered: a voting system and the elicitation of expert opinion.

**Simple Voting System**

Consider a system in which a decision maker is required to vote on a well-structured issue, such as the selection of a candidate. Let the candidates be *Roberts*, *Richards*, and *Williams*. Assume further that the voter is influenced by three groups to cast a vote to their best interests. In a real sense, the voter is being pulled in three directions. Let the influential groups be *Party*, *Cause*, and *Lobby*, delineated as follows:

- Party = {{{Roberts},0.6}, {{Richards},0.3}, {{Williams},0.1}}
- Cause = {{{Roberts},0.4}, {{Richards},0.2}, {{Williams},0.4}}
- Lobby = {{{Roberts},0.4}, {{Richards},0.5}, {{Williams},0.1}}

The information should be interpreted from a probabilistic view. For example, based on party affiliation, the voter will choose *Roberts* with a 0.6 probability, *Richards* with a 0.3 probability, and *Williams* with a 0.1 probability.

Clearly, the frame of discernment is {Roberts, Richards, Williams} and since the problem is well-structured, the basic probability assignments may be summarized as follows:

- Party = {0.6, 0.3, 0.1}
- Cause = {0.4, 0.2, 0.4}
- Lobby = {0.4, 0.5, 0.1}

Table 1 summarizes the application of Dempster’s rule of combination to this problem. First, *Party* is combined with *Cause*, and then the result is combined with *Lobby* to obtain a composite picture of the result when the support functions are fused. In general, the evidence is complementary and this fact is reflected in the entropy that is reduced from the initial party affiliation of 1.29 to 0.965.

Support Function	Probability Assignment	Entropy
Party	{0.6, 0.3, 0.1}	1.29
Cause	{0.4, 0.2, 0.4}	1.52
Lobby	{0.4, 0.5, 0.1}	1.36
Party×Cause (=K)	{0.706, 0.176, 0.118} 1.16	
K×Lobby	{0.739, 0.230, 0.031} 0.965	

**Table 1. Application of the Simplification Rule to the Voting System**

**Elicitation of Expert Opinion**

Typically, experts do not agree, especially when system failure is concerned. Typical examples might be the crash of an expensive fighter aircraft or the collapse of a building. Consider a situation wherein the frame of discernment is {A,B,C} reflecting that the failure could be caused by Component A, Component B, or Component C. Expert #1 believes the failure is due to Component A with probability 0.75, Component B with probability 0.15, or Component C with probability 0.10. Expert #2 believes the failure is due to Component A with probability 0.30, Component B with probability 0.20, or Component C with probability 0.50. The support function are:

$$\text{Expert \#1} = \{ \{A\}, 0.75 \}, \{ \{B\}, 0.15 \}, \{ \{C\}, 0.10 \} \}$$

$$\text{Expert \#2} = \{ \{A\}, 0.30 \}, \{ \{B\}, 0.20 \}, \{ \{C\}, 0.50 \} \}$$

The problem is well-structured, so the probability assignments are summarized as follows:

$$\text{Expert \#1} = \{0.75, 0.15, 0.10\}$$

$$\text{Expert \#2} = \{0.30, 0.20, 0.50\}$$

Table 2 summarizes the application of the simplification rule to this problem. The opinion of the experts is summarized and reflects the differing opinions.

Support Function	Probability Assignment	Entropy
Expert #1 (=X)	{0.75, 0.15, 0.10}	1.05
Expert #2 (=Y)	{0.30, 0.20, 0.50}	1.49
X×Y	{0.634, 0.084, 0.282} 1.23	

**Table 2. Elicitation of Expert Opinion**

The strong opinion of Expert #1 in favor of Component A, reflected in the low entropy, has a major influence on the consensus.

**ELECTION APPLICATION**

One of the most familiar unstructured decision applications is the prior assessment of the outcome of an election. The major determinants of how people will vote can be combined into eight well-known categories: party, gender, education, religion, attitude, age, region (of the country), and race. The structural elements for each of the

categories are given in Sheet 1, along with the respective structural probabilities. The columns are titled “Demographics.” For category Party, the structural element Rep has a probability (or proportion), for example, of 0.4. Associated with each structural element is a preference set for that element over the frame of discernment, which is {R,D}. In this case, a person in Party/Rep, would vote for R with probability 0.7 and for D with probability 0.3.

Categorical probabilities are calculated as a set of composite probabilities using equation [2], as shown in Sheet 2, which gives spreadsheet functions that compute the respective probabilistic elements in the category probability set. Sheet 1 gives the computed probabilities for this example in the “Categorical Probabilities” section.

Finally, the consensus probabilities are computed using the simplification rule (equation [1]) in the “Consensus” section of Sheet 2. The results of the actual calculations are given in the “Consensus” section of Sheet 1. The probabilities are combined from top down, starting with the Party category and ending with Race.

The results are more sensitive to demographics than they are to the preferences, as evidenced through experimentation with the spreadsheet recalc facility.

## **SUMMARY**

An admixture of methods has been given to structure a problem domain into categories and compute categorical probabilities from structure elements and preference sets. The categorical probabilities are then combined using Dempster’s rule of combination to obtain a composite assessment of the decision landscape. A demonstrative election application is given.

## **ACKNOWLEDGMENT**

Thanks to M. Lulu, H. Choi, and M. Katzan for reading the paper.

## **REFERENCES**

1. Dempster, A.P. (1967), Upper and Lower Probabilities Induced by a Multivalued Mapping, *The Annals of Statistics* 28:325-339.
2. Katzan, H. (1992), *Managing Uncertainty: A Pragmatic Approach*, New York: Van Nostrand Reinhold Co.
3. Katzan, H. (2006), Consensus, Proceedings of the Decision Science Institute National Conference, San Antonio TX. (November 2006).
4. Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton, NJ: Princeton University Press
5. Theil, H. (1967), *Economics and Information Theory*, New York: American Elsevier Publishing Company, Inc.

	B	C	D	E	F	G	H	I	J	K	L	M	N
1	<b>CONSENSUS (R/D)</b>			<b>Categorical Probabilities (R/D)</b>				<b>Demographics</b>				<b>Preferences (R/D)</b>	
2	0.54	0.46		Party	0.54	0.46		Party	Rep	0.4		0.7	0.3
3									Dem	0.5		0.4	0.6
4									Indep	0.1		0.6	0.4
5	0.618483412	0.381516588		Gender	0.58	0.42		Gender	Male	0.4		0.7	0.3
6									Female	0.6		0.5	0.5
7	0.58	0.42		College	0.46	0.54		College	None	0.4		0.2	0.8
8									Attended	0.2		0.5	0.5
9									Grad	0.4		0.7	0.3
10	0.560386473	0.439613527		Religion	0.48	0.52		Religion	Cath	0.3		0.2	0.8
11									Prot	0.3		0.6	0.4
12									Evan	0.2		0.6	0.4
13									Jewish	0.1		0.7	0.3
14									Other	0.1		0.5	0.5
15	0.618666667	0.381333333		Attitude	0.56	0.44		Attitude	Conserv	0.3		0.8	0.2
16									Liberal	0.3		0.4	0.6
17									Neither	0.4		0.5	0.5
18	0.560386473	0.439613527		Age	0.44	0.56		Age	<=25	0.4		0.2	0.8
19									26-55	0.4		0.5	0.5
20									>=56	0.2		0.8	0.2
21	0.599425837	0.400574163		Region	0.54	0.46		Region	N	0.3		0.4	0.6
22									S	0.2		0.4	0.6
23									M	0.2		0.8	0.2
24									W	0.3		0.6	0.4
25	0.646512413	0.353487587		Race	0.55	0.45		Race	B/C	0.1		0.2	0.8
26									Cauc	0.8		0.6	0.4
27									other	0.1		0.5	0.5

Sheet 1. Spreadsheet for the Election Application showing Demographics, Preferences, Categorical Probabilities, and the Consensus for the Election Application.

	B	C	D	E	F	G
1	CONSENSUS (R/D)				Categorical Probabilities (R/D)	
2	=F2	=G2		Party	=K2*M2+K3*M3+K4*M4	=K2*N2+K3*N3+K4*N4
3						
4						
5	=(B2*F5)/(B2*F5+C2*G5)	=(C2*G5)/(B2*F5+C2*G5)		Gender	=K5*M5+K6*M6	=K5*N5+K6*N6
6						
7	=(B5*F7)/(B5*F7+C5*G7)	=(C5*G7)/(B5*F7+C5*G7)		College	=K7*M7+K8*M8+K9*M9	=K7*N7+K8*N8+K9*N9
8						
9						
10	=(B7*F10)/(B7*F10+C7*G10)	=(C7*G10)/(B7*F10+C7*G10)		Religion	=K10*M10+K11*M11+K12*M12+K13*M13+K14*M14	=K10*N10+K11*N11+K12*N12+K13*N13+K14*N14
11						
12						
13						
14						
15	=(B10*F15)/(B10*F15+C10*G15)	=(C10*G15)/(B10*F15+C10*G15)		Attitude	=K15*M15+K16*M16+K17*M17	=K15*N15+K16*N16+K17*N17
16						
17						
18	=(B15*F18)/(B15*F18+C15*G18)	=(C15*G18)/(B15*F18+C15*G18)		Age	=K18*M18+K19*M19+K20*M20	=K18*N18+K19*N19+K20*N20
19						
20						
21	=(B18*F21)/(B18*F21+C18*G21)	=(C18*G21)/(B18*F21+C18*G21)		Region	=K21*M21+K22*M22+K23*M23+K24*M24	=K21*N21+K22*N22+K23*N23+K24*N24
22						
23						
24						
25	=(B21*F25)/(B21*F25+C21*G25)	=(C21*G25)/(B21*F25+C21*G25)		Race	=K25*M25+K26*M26+K27*M27	=K25*N25+K26*N26+K27*N27
26						
27						

Sheet 2. Spreadsheet for the Election Application Giving Functions for the Calculations in Sheet 1.