The problem of data fusion is in a real sense the problem of how to model the real world with all of its great complexities. A miniaturized version of this is the multiple target tracking and data association problem. There, a number of pieces of information arrive, typically from disparate sources—such as from various sensing systems and from human sources in the form of narrative descriptions in natural language. A procedure has already been established for dealing with this type of situation, called succinctly the PACT algorithm (PACT = Possibilistic Approach to Correlation and Tracking). The technique is based upon the premise that all arriving information can be adequately treated through some appropriate choice of classical or multivalued logic such as Probability Logic, Fuzzy Logic, Lukasiewicz-K, Logic, or some (t-norm, t-conorm, negation function) general logic as discussed in a recent text of Goodman and Nguyen, Uncertainty Models for Knowledge-Based Systems. Moreover, it can be demonstrated that for a large class of logics chosen, a version of a partially specified Probability Logic may be used instead. Indeed, other approaches to uncertainty, such as the Dempster-Shafer approach, can also be strongly related to Probability Logic through the vehicle of random set modeling. In any case, the structure of the PACT algorithm is based upon a generalized chaining and disjunction relation, which in a classical probability setting reduces to the usual posterior probability description as a weighted sum of intermediate probabilities, an alternative form of Bayes' formulation. In the PACT algorithm, joint inference rules are represented which connect various combinations of matches of the intermediate attributes relevant to correlation (such as geolocation, radar parameters, visual narratives, etc.) to the consequential correlation levels between track histories. In addition, error relations involving these attributes are also represented.

In the present paper, the PACT technique is extended to the full combination of evidence problem, viewed as being identical to the general data fusion problem. In addition, data fusion is also intimately linked with internodal activity within a larger C3 system. Here such C3 systems are identified as networks of interacting decision-maker node complexes. Some general examples of data fusion in this context are presented, including a new approach to the use of marginal conditional probabilities measuring validity of inference rules via "conditional objects."

<table>
<thead>
<tr>
<th>21a. NAME OF RESPONSIBLE INDIVIDUAL</th>
<th>21b. TELEPHONE (include Area Code)</th>
<th>21c. OFFICE SYMBOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>L. R. Goodman</td>
<td>(619) 553-4014</td>
<td>Code 421</td>
</tr>
</tbody>
</table>

UNCLASSIFIED
A GENERAL THEORY FOR THE FUSION OF DATA

I.R. Goodman
Command & Control Department
Code 421
Naval Ocean Systems Center
San Diego, California 92152

Abstract

The problem of data fusion is in a real sense the problem of how to model the real world with all of its great complexities. A miniaturized version of this is the multiple target tracking and data association problem. There, a number of pieces of information arrive, typically from disparate sources - such as from various sensing systems and from human sources in the form of narrative descriptions in natural language. A procedure has already been established for dealing with this type of situation, called succinctly the PACT algorithm. (PACT = Possibilistic Approach to Correlation and Tracking.) The technique is based upon the premise that all arriving information can be adequately treated through some appropriate choice of classical or multivalued logic such as Probability Logic, Fuzzy Logic, Lukasiewicz-H Logic, or some (t-norm, t-conorm, negation function) general logic as discussed in a recent text of Goodman and Nguyen. Uncertainty Models for Knowledge-Based Systems. Moreover, it can be demonstrated that for a large class of logics chosen, a version of a partially specified Probability Logic may be used instead. Indeed, other approaches to uncertainty, such as the Dempster-Shafer approach, can also be strongly related to Probability Logic through the vehicle of random set modeling. In any case, the structure of the PACT algorithm is based upon a generalized chaining and disjunction relation, which in a classical probability setting reduces to the usual posterior probability description as a weighted sum of intermediate probabilities, an alternative form of Bayes' formulation. In the PACT algorithm, joint inference rules are represented which connect various combinations of matches of the intermediate attributes relevant to correlation (such as geolocation, radar parameters, visual narratives, etc.) to the consequential correlation levels between track histories. In addition, error relations involving these attributes are also represented.

In the present paper, the PACT technique is extended to the full combination of evidence problem, viewed as being identical to the general data fusion problem. In addition, data fusion is also intimately linked with internal activity within a larger C3 system. Here such C3 systems are identified as networks of interacting Corison-maker node complexes. Some general examples of data fusion in this context are presented, including a new approach to the use of marginal conditional probabilities measuring validity of inference rules via "conditional objects".

1. INTRODUCTION

For the past several years, throughout many fields of science and technology, researchers have been seeking unification and extension of past results in order to explain reality better and to be able to predict future developments. Recent events in theoretical physics involving "superstring" theory, an attempt at developing a Grand Unified Theory of the Universe, underscore this quest [1].

In a more modest way, this paper seeks to establish a theory unifying, coordinating, and extending the somewhat appearing distinct concepts of data fusion, combination of evidence, and C3 systems analysis. On the other hand, relatively little attention will be paid here to data-based computational techniques which are particular to certain types of common data: fusion problems such as regression procedures for combining stochastic sensor information, or maximum likelihood or Bayesian procedures for putting together geolocation data arriving from different sources relative to a given target of interest. All of the above-mentioned techniques are essentially special cases of a much more general combination of evidence approach on which this paper will concentrate.

In the past there has been much dispute as to what constitutes data fusion. A reasonable three-fold definition has been proposed in [2], which, except for a minor modification (as shown below), will be the basis for the work here. In a related vein, mention should be made of the recent (unclassified) survey of data fusion techniques [3]. The basic definition for data fusion, for completeness, is given below:

(i) "The integration of information from multiple sources to produce the most comprehensive and specific unified data about an entity."

(ii) "The analysis of intelligence information from multiple sources covering a number of different events to produce a comprehensive report of activity that assesses its significance. The analysis is often supported by the inclusion of operatinal data."

(iii) "Intelligence usage, the logical blending of related information / intelligence from multiple sources." [After fusion, the sources of the inputs and single pieces of information must not be evident to the user. This we believe to be too restricted, [IRG].]

One of the most common examples of fusion of data occurs in the multiple target-tracking problem. Here, information arrives in disparate form. Typically, this includes sensor information emanating from possibly different types of sources, such as radar, acoustic, non-acoustic, infra-red, and various others. In addition, non-mechanical/ human sensor sources may be present in the form of natural language narratives or descriptions, possibly in a parsed form, suitable for symbolizations. Much of the arriving information can be related to the targets' observed or predicted positions, velocities, or related equations of motion. On the other hand, some of the data may refer to other characteristics or attributes of the targets.
certainly of the latter include: hull lengths, vessel shapes, observed flag colors, names, classifications, and other non-geolocational sensor parameter estimates.

Nevertheless, as recently as a few years ago, the great majority of approaches to target data fusion were concerned only with target positions and other geolocation data and ignored, at least in a formal way, most of the other potentially useful stochastic and non-stochastic (such as linguistic) information. For a solid justification of this conclusion, see [4] and [5], where a comprehensive survey of multiple target-tracking techniques was carried out. For comprehensive mathematical treatments of such "classical" data association and correlation, see [6], e.g. For an exception to the above statement concerning the restriction of fusion to geolocation-only information, see, e.g. [7],[8],[9].

However, with the advent of AI in the form of expert and knowledge-based systems, it is apparent that this additional information could be utilized. (See, e.g. [10].) Following the lead of medical diagnostic systems such as MYCIN [11], many such systems (not necessarily military oriented) utilize only two-valued logic in conjunction with some use of probabilities to represent confidences. On the other hand, some approaches take a "softer" viewpoint as to the nature of descriptions and employ throughout some form of multivalued logic (such as the PACT algorithm [12]).

Moreover, data fusion is intimately related to the functioning of C³ systems. Indeed, in many cases, data fusion may be perceived as an interacting decision process occurring within each decision-maker node relative to the entire C³ network of nodes. Thus, any ongoing work in the C³ arena must affect data fusion efforts. Since 1978, the annual MIT/ONR Workshop on C³ Systems - with its associated (unclassified) annual Proceedings has served as one of the primary academic sources for generic C³ studies. (See [13] for a partial survey of these efforts. See also [14] for a more thorough survey of C³ work, where many abstracts, analyses, and comparisons and contrasts of C³ theses and related work are given.) Surprisingly, relatively few comprehensive theories of C³ systems have been produced, although many valuable papers have been written as a result of this effort. Workshop on problems of distributive decision-making, communications systems, and security, and target-tracking and correlation, and various miscellaneous data theoretical and operational design problems. Among the few theories of C³ that should be mentioned [41] and [42], the latter taking a related view of fusion. Based upon the above remarks, it is the author's conclusion that:

1. Data fusion, as commonly applied, is a process occurring intradurally within the context of a given node's process within decision-making nodes.
2. All analysis and models of C³ systems must include subanalysis and models for fusion-processes. In particular, this applies to this author's proposed model for C³ systems [15],[16].
3. Data fusion in its most generic sense can be equated with the combination of evidence problem, a well-known problem arising in the modeling of uncertainties for knowledge-based systems. (For further elaboration and background, see [17].)

2. DATA FUSION, C³ SYSTEMS, AND DATA PROCESSING

Previously, this author proposed a bottom-up, microscale quantitative approach to generic C³ systems [15],[16]. In that approach, a generic C³ system is identified as a network of node complexes of decision-makers, human or automated, interfacing with each other in general. Each node receives "signals" which may be ordinary communication signals, either from friendly or hostile sources (possibly unaware), or which may be received weapon fire. In general, these "signals" are stacked vectors colored of incoming data from several different nodes. In turn, each node, which may consist of a single decision-maker or some coalition of decision-makers and which may include passive type decision-makers, such as "followers", then processes the data, this followed by a response or action taken towards other nodes, friendly or hostile. (See Figure 1.) Associated with
each node is the node state (see Figure 2.) describing the current state-of-affairs given in terms of a number of functions such as threat level, equations of motion, and supply level. In addition, there is an associated knowledge base reflecting the node’s local knowledge of other nodes (other friendly or adversary). Also associated with each node is its internal "signal" processing design, as described in Figure 3. (5) Data fusion plays a central role in transmitting detected "signals" to hypotheses formulations, which in turn through algorithm selection leads to an output response to other nodes (again, these may be friendly or adversary).

Next, since we identify data fusion with the combining of evidence, all of the knowledge-based system techniques associated with the latter are available. In particular, this refers (see [17], Chapters 1, 2 and Figure 1, page 14) to a series of underlying processes involved in data fusion. Basically, there are five such processes (including natural language in its broadest context) given below in sequence of information processing:

(1) Cognition: Human and/or machine in recognizing the pattern of received "signals", recalling that "signals" refer to either ordinary signals or any other received input, including weapons fired.

(2) Natural Language Formulation: This is relevant to all narratives produced by human observers. Machine language could also be put in this area, if used in the same context. Parsing leads to the next process:

(3) Primitive symbolic formulation of data, including strings of well-formed formulas according to basic syntax, without further or refined constraints on structures. Formulations include use of basic quantifiers and connectors: for "and" or conjunction; v for "or" (disjunction); ) for "not" (negation); , for "if then" (implication).

(4) Full formal language formulation of data: Use of rules of syntax, constraints on wff’s, such as commutativity, associativity, idempotence, distributivity, etc.

(5) Full compatible (homomorphic-like) semantic evaluations or logic chosen (or model selected).

Any consistent or compatible choice of a full formal language (4) and a semantic evaluation or logic (5) we will call an algebraic logic description pair (ALDP).

Three common choices for ALDP are:

ALDP 1 = (Boolean algebra (or ring), Classical two-valued Logic) with implication given as →, where F → F is identified as B v a, for all wff’s c.c.

ALDP 2 = (Modified boolean algebra = pseudo-complemented lattice, Zadeh’s (min-max) Fuzzy Sets or Logic). As above, B → A

ALDP 3 = (Boolean algebra, Probability Logic); B → B

A fourth useful (Conditional Probability Logic): ALDP will be introduced later. In the past, often only ALDP 1 or ALDP 3 were chosen, in effect, to the exclusion of multivalued logical choices. That is, either Classical Logic or Probability Logic, or some combination, would be chosen for the basic model to combine information or fuse data, without attention paid to the formal aspects prior to semantic evaluations. (Again, see [4],[5].)

Figure 4 summarizes the above analysis of data fusion.

3. DATA FUSION AS A QUANTITATIVE PART OF AN OVERALL C3 SYSTEM AND DECISION GAME

So far, in this development toward a general theory for the fusion of data, only general qualitative descriptions have been given for the processes involved. However, as mentioned before, a quantitative model for generic C3 systems has been established. Interactions have been formulated [15],[16]. Inputs to the structure consist basically of ten sorts of known relevant primitive relations PRIM among the variables describing a C3 system. These variables are: node (n) hypotheses selection (h); detection (d) of incoming "signals" (s); algorithm selections (f); initial node responses (r), prior to environmental distortion (e) and additive noise (n). To each variable is affixed subscripts (g, k) (or (h, g, k)) where g(a, i) denotes the identification of a particular node in question in terms of the C3 system a (friendly or hostile) and node number 1, while k represents a discrete time index t. Specifically, the relation breaks down into 5 intranodal (within nodes) relations, 2 intermodal (between nodes); or regression relations, and 3 prior relations for each C3 system. These relations are expressed in terms of conditional or unconditional probabilities, as they stand, but the results can be extended, with appropriate replacements, to a multivalued logic setting. (Again, see [15].) Then by making certain reasonable sufficiency assumptions among the variables and utilizing basic properties of conditional probabilities, it can be shown that each updated node state can be obtained explicitly (probabilistic) terms of the other variables and node states through PRIM. Thus, we have:
Theorem 1. (See [15], Theorem 1.)

Suppose \( \text{PRIM}_k \) and \( N_{g,k} \) are as described above with \( \text{PRIM}_k \) given in further details in eqs. (3.2)-(3.4) and Tables 1-3. Then under the above-mentioned sufficiency conditions,

\[
p(N_{g,k}) = \phi_{g,k}(\text{PRIM}_k),
\]

(3.1)

where \( \phi_{g,k} \) is a computable functional involving a finite number of integrations and arithmetic operations upon the elements of \( \text{PRIM}_k \) given in further detail (as in Table 4).

\[
\text{PRIM}_k \neq \{ \text{PRIM}_{(1,1)}, \text{PRIM}_{(1,2)} \} \cup \{ \text{PRIM}_{(2,1)} \}
\]

(3.2)

where for \( c^2 \)-system \( a, g(a, 1), \) etc.,

\[
\text{PRIK} \left( \begin{array}{c}
1, \ldots, 5 \\
\text{g}, \text{k}
\end{array} \right) \begin{array}{c}
(1, g), (5, 9), (6, 16), (9, 5), (10, 1), (13, 5), (15, 3)
\end{array}
\]

(3.3)

and where

\[
\text{PRIK}_k \left( \begin{array}{c}
1, \ldots, 5 \\
\text{g}, \text{k}
\end{array} \right) \begin{array}{c}
(1, g), (5, 9), (6, 16), (9, 5), (10, 1), (13, 5), (15, 3)
\end{array}
\]

(3.4)

The numerical symbols \( 5_{g,k} \), etc. are shortened forms for the primitive relations given in Tables 1-3:

\[
\begin{align*}
(1) \ g, k &= p(N_{g,k} | g, k, S_{g,k}) \\
(2) \ g, k &= p(H_{g,k} | g, k) \\
(3) \ g, k &= p(R_{g,k+1} | g, k, S_{g,k}) \\
(4) \ g, k &= p(N_{g,k+1} | g, k+1, S_{g,k}) \\
(5) \ g, k &= p(0_{g,k} | S_{g,k}) \\
\end{align*}
\]

Table 1. Relative Primitive Intralodal Relations.

\[
\begin{align*}
(6) \ h, g, k+1 &= p(0_{g,k+1} | h, g, k+1) \text{ with } h, g, k+1' \\
(7) \ h, g, k+1 &= p(0_{g,k+1} | h, g, k+1' \text{ for } h, g, k+1'' \\
\end{align*}
\]

The basic intralodal analysis is developed via an additive nonlinear regression relation

\[
\begin{align*}
\text{(S, g,k)+1} | \text{g,k+1} = (h, g,k+1) \\
\text{where variable } \text{h, g,k+1} \text{ indicates original possible node source for "signal" at time k, given reception by another node at k+1.

Table 2. Relative Primitive Intralodal Relations.

\[
\begin{align*}
\text{PRIOR/INITIAL TIME} \\
(2) \ h, g, k &= p(N_{h, g, k}) \\
(15) \ h, g, k &= p(0_{h, g, k} | h, g, k) \\
(16) \ h, g, k &= p(S_{h, g, k} | h, g, k) \\
\end{align*}
\]

Table 3. Relative Primitive Prior/Initial Relations.

\[
\begin{align*}
(5) \ g, k+1 &= p(N_{g,k+1} | 0_{g,k+1}, S_{g,k+1}) \\
\end{align*}
\]

In turn, a simple two-person zero sum game can be established, called the C-linear decision game. Here, Player I corresponds to entire \( c^2 \)-system \( a \) (say, friendly) and Player II corresponds to entire \( c^2 \)-system \( a \) (say, adversary). In this game, a move by Player I corresponds to a choice (up to given constraints) of \( \text{PRIK}_k \) as \( \text{PRIK}_1 \), \( \text{PRIK}_2 \), and the resulting loss or utility due to any such joint move is a function of the marginal updated node state distributions according to Theorem 1 as

\[
L_k(\text{PRIK}_k) = \text{MOC}_k(p(0_{h, g, k} | h, g, k)) \\
\]

Table 4. Structure of \( 5_{g,k} \) in Theorem 1 Through Sequence of Calculations Involving: PRIK_k
where MOE$_k$ represents a single figure-of-merit, combining various measures of effectiveness (MOEs) or performance (MOPs) for the two C$^3$ systems. (Note, that although ideally the entire joint node state distribution of the two C$^3$ systems should be sought, in practice this is difficult to do, because of the great combinatoric computations involved.) Typical MOEs that could be used include: averaged measure of importance $\phi$, averaged measure of threshold $\tilde{\phi}$, upper bound total entropy $\Phi$, and averaged measure of performance $\tilde{\Phi}$. (See also [15], eqs. (39)-(52).) Then one could let

$$\text{MOE}_k = \text{MOE}_{1,k} - \text{MOE}_{2,k},$$

where

$$\text{MOE}_{a,k} = \lambda_1 \Phi_{a,k} + \lambda_2 \phi_{a,k} + \lambda_3 \tilde{\phi}_{a,k} + \lambda_4 \tilde{\Phi}_{a,k},$$

and the $\lambda_i$'s are some predetermined weightings.

Symbolically, the C$^3$ decision game appears as given in Figure 5:

![Figure 5. Symbolic Form for C$^3$ Decision Game.](image)

Finally, one can then apply all the usual game-theoretic methods to this C$^3$ game, such as seeking Bayes decision functions for moves, least favorable strategies (all subject to practical constraints), minimax strategies, the game value, and various sensitivity measures. It is the long-range hope that such a game will be a useful decision-aid in planning command strategy. At present, a relatively simple implementation scheme is being carried out for testing the feasibility of such an approach to C$^3$ systems. (See [15] for further details.)

### 4. Structure for Data Fusion: The Classical Probability Case

With the general C$^3$ system context for data fusion established in the previous sections, let us now return to the task of developing a general quantitative structure for data fusion. In light of the previous remarks (again, see Figure 3), fusion is a process intermediate with internal sensing and hypothesis formulations, within a C$^3$ node complex of decision-makers. In addition, the fusion process decomposes into natural subprocesses (see Figure 4). Thus, in essence, we wish to expand the first relative primitive intranodal relation appearing in Table 1:

$$P(FU) = p(H(D,S)),$$

where for reasons of convenience from now on we suppress the denotational-time indices, unless necessary. As stated before, $p$ need not necessarily refer to ordinary probability evaluation, but may represent other evaluations such as possibilities for Zadeh's Fuzzy Logic or for more general multivalued truth systems.

In determining the above evaluation, another variable $Z$ is often present. $Z$ represents the vector of auxiliary or "nuisance" characteristics or attributes which can be useful in connecting $H$, the variable representing possible hypotheses or decisions as to what unknown parameter value or situation or diagnosis is occurring, with input data $S$ and detection state $D$. Thus for example, if one is physically in a bunker - a C$^3$ node - $S$ may be observed loud noise, with $D=1$ (definedly detected), and $H$ could have possible domain values say $\text{dom}(H) = (H_1, H_2, H_3)$ as given in Table 5:

$$H_1 = \text{no change in previous situation}$$
$$H_2 = \text{enemy is about to mount the promised big offense}$$
$$H_3 = \text{enemy is just feeling us out}$$
$$H_4 = \text{enemy wants to negotiate}$$
$$H_5 = \text{none of the above situations hold}$$

Thus, dom$(H)$ could serve as a legitimate sample space, if conditional probability $p(H(D,S))$ could be obtained for all possible values of $H$ in $\text{dom}(H)$, i.e. (H(D,S)) could be interpreted as a random variable over dom$(H)$. In this case, suppose also that $Z$ is an auxiliary variable representing any of a likewise collection of disjoint exhaustive situations locally going on at the bunker. Here, let dom$(Z)$ be given as in Table 6 below:

$$Z_1 = \text{nothing happening}$$
$$Z_2 = \text{accidental explosion in compartment 1}$$
$$Z_3 = \text{accidental explosion in compartment 2}$$
$$Z_4 = \text{enemy's not at us and it either hit us or just missed}$$
$$Z_5 = \text{none of the above situations hold}$$

Thus, again by disjointness and exhaustion, it is reasonable to conclude that dom$(Z)$ could serve as a legitimate sample space and $Z$ can be interpreted as a random variable. All of this leads to the evaluation of the conditional probabilities $p(Z|D,S)$, which together with the values for $p(H|D,S)$ can be used to obtain the standard "integrated-out" form for the posterior distribution of $H$ as given below:

$$p(H|D,S) = p(H,Z|D,S) _{i=1}^5 = p(Z_i|D,S) p(H_j|Z_i,D,S) ,$$

using the standard chaining property of conditional probabilities and replacing the antecedent comma notation by conjunctions. One could reasonably interpret the evaluation in (4.2) as the probability value for the expression

$$\text{"If } D \text{ and } S, \text{ then } H_j \text{"}$$

through the probability values for the expressions

$$\text{"If } D \text{ and } S, \text{ then } Z_i \text{"} \text{ and } \text{"If } Z_i \text{ and } D \text{ and } S, \text{ then } H_j \text{"}$$

Of course, one need not use the above evaluation exactly to obtain useful equivalent values. As it stands, $p(Z|D,S)$ can be interpreted as an error or variability probability for attribute $Z$, while $p(H|Z,D,S)$ can be understood to mean the inference rule probability connecting $Z$ and $D$ and $S$ with $H$. On the other hand, often
the conditional data or regression probability
$p(S|Z_{1}, H_{j})$ and the joint prior probability $p(Z_{1}|H_{j})$
are available, assuming here $D=1$, which by use of
Bayes' theorem also yields $p(H_{j}|S,D)$. One standard
result is to assume the above probabilities are
gaussian, which in the discrete problem here, must
result in very rough approximations. In addition, the
sets $\text{dom}(\alpha)$ and $\text{dom}(Z)$ are not easily ordered
compatible with a real domain for gaussian random
variables. Then, if the mean of the conditional data
distribution is linear in the data $S$, $p(H_{j}|S,D)$
takes on a generalized weighted least squares form.
(See, e.g., [18]). The final result, $p(H_{j}|S,D)$, as in
(4.2), is then a mixture of the probabilities of
such least squares estimators.

5. STRUCTURE FOR DATA FUSION: THE CLASSICAL
PROBABILITY CASE MODIFIED

Retaining the same terminology as before, sup-
pose now that $H, Z, S$ are variables such that any of
the corresponding "sample spaces" do not truly con-
tain disjoint exhaustive events; in particular, the
disjointness condition may be violated more often
than exhaustiveness--which we will assume here is
always satisfied. Then it follows that the simple corre-
ponding probability measures as in Section 4 cannot
be immediately assigned. Nor should "brute-
force" normalization procedures be employed, unless
absolutely necessary. For example, consider $H$. Sup-
pose in the above example in Section 4 (Table 5), the
enemy could simultaneously mount the promised offense
($H_{1}$), or, even additionally wanting to negotiate ($H_{2}$).
Thus, in that case, $\text{dom}(H)=\{H_{1}, \ldots, H_{n}\}$, as it stands, is not a
suitable sample space of disjoint elementary events.
Indeed, the elementary events $H_{i}$ are not so element-
ary, many of them, due to complex causes, being over-
lapping! Equivalently, $H$ in its current form may not be a legitimate random variable. What to do?

Note first that it is reasonable to assume that
the simple labels $H_{i}$ really represent complex phe-
enomena and may be better described through factors contributing to them. For example, some factors for $H$ in Table 5 are:

$a_{1}$ = Importance of node,
$a_{2}$ = Relative strengths of us and then,
$a_{3}$ = Post and present incoming salvo rate,
$a_{4}$ = Duration of war to this point,
$a_{5}$ = What the enemy knows about us: location,
$a_{6}$ = Present weather conditions,
$a_{7}$ = Safety level-coordination level to
prevent accidents;
a \notin a_{1}, \ldots, a_{7}.

Then ideally, in turn, given enough of these
factors, define rigorously the $H_{i}$'s in terms of com-
binations of values of the $a_{i}$'s. One simple approach is to determine the natural domains of values for the
$a_{i}$'s, $\text{dom}(a_{i})$, $i=1, \ldots, n$, letting

$$\Omega \equiv \text{dom}(a_{1}) \times \cdots \times \text{dom}(a_{n})$$

and

$$H_{j} = b_{j_{1}} \times \cdots \times b_{j_{7}} \in \Omega.$$  

where $b_{j_{1}}, \ldots, b_{j_{7}}$ is determined by $H_{j}$, $j=1, \ldots, n$. Thus,
the overlapping of the $H_{i}$'s in general will not ap-
appear, but rather will be clarified, i.e., in general,

$$H_{j_{1}} \cap H_{j_{2}} \neq \emptyset.$$  

Clearly, in this case, if all statistical relations
between the newly-introduced factor variables $a_{i}$'s
and the variables $S$ and $Z$ are known, then the
$p(H_{j}|Z_{1}, D)$'s can be computed in (4.2). For example,
if the $a_{i}$'s given the $Z_{1}$'s are all mutually statisti-
cally independent, then

$$p(H_{j}|Z_{1}, D) = \Pi p(a_{k} \in b_{j_{k}}, k|Z_{1}, D) \quad (5.4)$$

and in general

$$p(Z_{j}|H_{j}, D) = \Pi p(H_{j}|Z_{1}, D) \quad (5.5)$$

and the computation in (4.2) involving summing over
the domain of $Z$ is no longer valid if $Z$ also repre-
sents, as $H$, possibly complex overlapping events.

One approach to redifining the problem here is to replace the, in general, overlapping $H_{i}$'s and over-
lapping $Z_{i}$'s by suitable partitioning of their domain
spaces and then recompute the corresponding condition-
alg probabilities in (4.2) involving the partitioning
variables. For example, for convenience, denoting

$$I = \{1, \ldots, 5\}$$

for any subset $K \subseteq I$, or equivalently, $K_{p}(I)$ (power
class of I, the class of all subsets of I), define

$$H[K] \equiv \{ H_{j} \in I \mid j \in K \}$$

and for $K=\emptyset$

$$H[\emptyset] = \emptyset.$$  

for example, for $K=\{1,2,4\}$,

$$H[K] = H_{1} \cup H_{2} \cup H_{4}.$$  

Clearly,

$$H[K] \subseteq H[K] \subseteq H[I] \subseteq \emptyset.$$  

is a disjoint exhaustive-partitioning of $\Omega$. In a sense,
$H[K]$ is the strongest disjoint exhaustive partitioning
of $\Omega$ which generates back all $H_{i}$'s through disjoint
unions. Thus, $H$ can serve as a sample space in place of
initial $\text{dom}(H)$; the $H_{i}$'s are in general overlapping
compound events of $H$. Similar comments hold for $Z$.

Note that the mappings $H_{j} : P(I) \rightarrow \text{dom}(H)$
and $H[I] : P(I) \rightarrow \text{dom}(H)$ are injective (1-to-1 into),
for all $K \subseteq I$ such that $H[K] \neq \emptyset$. Hence we have the
bijective relation for all $K$ such $H[K] \neq \emptyset$

$$K \rightarrow H[K] \rightarrow H[K].$$  

(5.14)

For any $j \in I$, define the filter class of $H_{j}$, or
the point coverage class of $H_{j}$, as

$$F[H_{j}] \equiv \{ H[K] | j \in K \}.$$  

(5.15)

define similarly,

$$F[H_{j}] \equiv \{ H[K] | j \in K [K] \}.$$  

(5.16)
Note also that the mappings \( f_j : \text{dom}(H) \rightarrow \text{PP}(\text{dom}(H)) \) and \( f_j : \text{dom}(H) \rightarrow \text{PP}(D) \) are injective. Note, further, for any \( j \), the bijective relations

\[
H_j = G(H_j) = f_j(F(H_j)) = f_j(F(H_j)). \tag{5.17}
\]

Now let \((0,8,p)\) be a probability space and \( \mathcal{H} \subseteq \mathcal{D} \) be a random variable corresponding to \((\mathcal{U}, \mathcal{F}, \P)\). In turn, define random subset \( S_H^{(1)} \) of \( \text{dom}(H) \), \( S_H^{(1)} : n = \text{PP}(\text{dom}(H)) \), where for any \( u \in U \),

\[
S_H^{(1)} \in [H_j \cup \{H_j\} \cup \{H_j\}] \quad \text{iff} \quad H \in \{S_H^{(1)} \},
\]

\[
\text{iff} \quad H \in \{S_H^{(1)} \},
\]

\[
\text{iff} \quad H \in \{S_H^{(1)} \},
\]

\[
\text{iff} \quad H \in \{S_H^{(1)} \}.
\]

Hence

\[
V \in H_j \text{ iff } \{V \in \text{PP}(\text{dom}(H)) \} \text{ is a random variable corresponding to } \{H_j \cup \{H_j\} \cup \{H_j\} \text{ for any } u \in U \}.
\]

Theorem 2. \((5.20, 5.21, 5.22, 5.23)\)

For all \( j \),

\[
\text{poss}(H_j) = p(W \in H_j) = p(H_j \in S_H^{(1)}) = p(H_j \in \{S_H^{(1)} \},
\]

\[
\text{iff} \quad H \in \{S_H^{(1)} \}.
\]

The significance of this theorem will be more apparent below. Note also that unless \( \text{dom}(H) \) is a disjoint partitioning itself of \( \mathcal{D} \), \( \text{PP}(\text{dom}(H)) \) holds, but it always follows that

\[
\sum_{K} p(W \in H_j) = 1
\]

\( \text{Kg} \)

Again, similar results hold for \( \text{PP}(\text{dom}(2)) \) replaced by a suitable space resulting from appropriately chosen factors.

On the other hand, often we do not know all the relevant factors or variables contributing to given compound events and even if these variables can be pinpointed, often we do not know their natural domains or perhaps do not know the distributional relationships involved, etc. Thus the technique of constructing directly a product space, such as \( \mathcal{D} \) for \( H \), as above, may not be appropriate.

However, we can still make the basic identifications in (5.14) and (5.17), where we omit all the square-bracket expressions. Suppose now that probabilistic evaluations are available such as \( p(H_j \in \{S_H^{(1)} \} \) and \( p(H_j \in \{S_H^{(1)} \} \) for all \( i \) and \( j \), but that the possible overlapping nature of the compound events is taken into account. For example, these calculations could be obtained from experts by soliciting the individual marginal possibilities occurring without regard to the joint or overall occurrences of the remaining events.

Can these individual probabilities or possibilities be made compatible in a rigorous manner with the previous random set construction? The answer is yes.

Theorem 3. \((5.17, 5.20, 5.21)\)

If \( \text{poss}(\text{dom}(H)) = [0,1] \) is any function, perhaps representing the expert opinions of a panel, as human integrators of information, taking into account the complex and possible overlapping nature of the events in \( \text{dom}(H) \) and defining the nested random subset of \( \text{dom}(H) \) by

\[
S_H^{(2)} \subseteq \text{poss}^{-1}(\{U, 1\}) = \{H_j \cup \{H_j\} \cup \{H_j\} \text{ iff } \text{poss}(H_j) \geq U \},
\]

it follows that for all \( j \),

\[
H_j \in S_H^{(2)} \text{ iff } \text{poss}(H_j) \geq U,
\]

whence there exists a legitimate probability measure \( p: \text{PP}(\text{dom}(H)) \rightarrow [0,1] \) such that

\[
\text{poss}(H_j) = p(H_j \in S_H^{(2)}) = p(S_H^{(2)} \in \text{dom}(H_j))
\]

\[
= \sum_{i \in \text{Kg}} p(S_H^{(2)} \in \text{dom}(H_j)).
\]

Remarks:

Note first that the two definitions for \( S_H \) will differ in general in structure, but are both among many other possible definitions for such random sets. \( \text{dom}(3) \) (among many other possible definitions for such random sets) one point coverage equivalent to the given arbitrary possibility function over \( \text{dom}(H) \). (For comparisons of choices among such candidate random sets, see \( [20] \), where entropy is used as one criterion.) Each domain value \( H_i \) is naturally identifiable with the filter class \( G(H_j) \) containing all possible sets of \( H_i \) having also \( H_j \) in them, i.e., all possible sets of \( \text{S_H}^{(2)} \) \( j \) in \( K \). Thus it is not unreasonable that the given possibility value assigned to \( H_i \) could be estimated rigorously as a probability involving the next higher order interaction domain \( p(\text{dom}(H)) \) above \( \text{dom}(H)_j \). Again, as before, all results hold for \( 2 \).

In a word, the possibilistic or general fuzzy set approach is seen to be essentially a weakened form of the full random set approach, where any one of the one point coverage equivalent random sets \( S \) is fixed for the modeling over \( \text{PP}(\text{dom}(H)) \), replacing \( \text{dom}(H) \). This can be thought of as being somewhat analogous to the situation where a probability distribution describing a problem is only partially specified, such as up to the mean and variance.

Finally, homomorphic-like relations (involving the one point-covetage relations) can be established between a number of operations established among possibility functions or fuzzy sets, representing generalized unions, intersections, and other set-like operations, and corresponding ordinary set counterparts applied to the one point coverage equivalent random sets. \((\text{See, e.g.} \text{[17]}, \text{Chapter 6})\) Some of these relations will be used in Section 6 for representing data fusion in terms of the general combination of evidence problem. \((\text{In a related vein, see [21]} \) for some recent work using random sets in modeling problems.)

6. STRUCTURE FOR DATA FUSION: THE GENERAL FIXED ANTECEDENT CASE

The results of the previous section point up some of the difficulties involved in evaluating probabilities for apparently disjoint elementary events which are in reality compound overlapping and difficult to define precisely.

Following the philosophy of approach outlined in Figure 4, we will establish a general procedure for treating the combination of evidence problem, which reduces to the probability or possibility cases when appropriate. Ideally, this procedure should reflect
cognition (box 1 in Figure 4), the first stage following initial "signal" detection, but for purposes of simplicity this will be omitted in the present paper.

In particular, consider the crucial expression \( Q \) for data fusion appearing as primitive intramodal relation (1) in Table 1, sans the probability evaluation, and in natural language form:

\[
Q \equiv \text{ "If } D \cup S, \text{ then } H."
\]

(6.11)

In symbolic form, where \( \equiv \) represents \( \exists \), \( \wedge \) represents "or", \( \not\equiv \) represents "not", \( \equiv \) represents implication.

\[
Q = (D \cup S \equiv H).
\]

(6.2)

Suppose next, the following two basic properties hold for the natural language used:

(a) Letting \( T_0 \) represent absolute truth, for any proposition \( a \),

\[
a \equiv T_0 = a .
\]

(6.3)

i.e., \( T_0 \) plays the role of a multiplicative unity w.r.t. "and", and can be denoted w.l.o.g. as 1. Dually, we can assume the existence of an absolute falsehood \( F_0 \) and let it play the role of an additive zero w.r.t. "or".

(b) "and" and "or" are commutative and associative with "not" being distributive over "or".

These properties are quite mild and will serve in no way here to restrict our choice of ALDP (algebraic logic description pair). The four examples in Table 4 all satisfy these conditions.

(i) Suppose also that auxiliary attribute variable \( Z \), used to connect \( D \) and \( S \) with \( H \), is such that

\[
\text{or } (Z_i) = T_0 .
\]

(6.4)

Equivalently, this means that the possible "values" of \( Z \) are exhaustive, even if they overlap. Symbolically,

\[
Z_i \in \text{dom}(Z).
\]

(6.5)

(ii) Suppose, further, that \( Z \) relative to \( D, S \), and \( H \), is such that

\[
Q \equiv \text{ "If } D \cup S, \text{ then } H \text{ or } \phi ."
\]

(6.6)

where \( \phi \equiv \text{ or } (Z_i \not\equiv T_i). \)

(6.7)

\[
Z_i \in \text{dom}(Z).
\]

(6.8)

In many formal languages, the Law of Excluded Middle holds so that for all propositions \( a \),

\[
a \equiv \text{ not}(a) = F_0 .
\]

(6.9)

But in many multiple-valued logics, such as Zadeh's Fuzzy Sets, (6.8) does not hold, and an alternate condition must be sought to obtain the desired results we seek. (See also Example 2, Section 7.)

Symbolically,

\[
Q = (D \cup S \equiv (H \vee \phi)) ,
\]

(6.10)

where

\[
\phi = \vee (Z_i \equiv Z_i') .
\]

(6.11)

Then if we apply (a),(b),(i),(ii) to (6.1), we obtain in symbolic form

\[
Q = (D \cup S \equiv (H \vee Z_i' \vee Z_i) ) .
\]

(6.12)

Next, two more restrictive assumptions are made:

(c) The antecedent of implication is distributive over "or"; equivalently, a homomorphism exists relative to "or" for a fixed implication antecedent. Thus for any propositions \( a_1, a_2, \ldots, a_n \),

\[
\text{or } (a \equiv (a_1 \equiv a_2 \equiv \ldots \equiv a_n)) \equiv (a \equiv a_1 \equiv a_2 \equiv \ldots \equiv a_n) .
\]

(6.13)

(d) Implication chains relative to "\&". Thus for any propositions \( a, b, \gamma \),

\[
\gamma (a \equiv (b \equiv b')) = (\gamma \equiv a) (\equiv b \equiv b') .
\]

(6.14)

(See Examples 1-3, Section 7, where ALDP 1-3 are presented in some detail. For ALDP 4, see Section 8.)

Theorem 4.

Suppose a formal language of propositions satisfies constraints (a),(b),(c),(d). Suppose also that variables \( D, S, H, Z \) are to be interpreted as before in the general sense and are such that (i) and (ii) are satisfied, then

\[
Q = \vee (Z_i: D; S; H) ,
\]

(6.15)

where for all \( Z_i \in \text{dom}(Z) \),

\[
\phi = g(Z_i: D; S; H) .
\]

(6.16)

where

\[
\phi = g(Z_i: D; S) = (D \equiv Z_i) .
\]

(6.17)

can be interpreted as an attribute variability or error form and

\[
A(H: Z_i: D; S) = (Z_i \equiv D \equiv H) .
\]

(6.18)

Thus from the remarks preceding Theorem 4, the formal language for Classical Logic and Possibility Logic, boolean algebras, with implication given in (6.16),(6.17), satisfies (6.16)-(6.19). Similarly, the modified boolean algebra representing the formal language of Zadeh's Fuzzy Logic (min-max type) also satisfies the above formal relations for the decompostion of the key expression for data fusion \( Q \).

In turn, we seek the full semantic evaluation of the data fusion expression through probability or possibility or other means, compatible with the results of Theorem 4.

In order to accomplish the above goal, we first review some concepts which may not be too familiar to many. Define a copula \( \phi_k \) as a mapping \( \phi_k: [0,1]^m \rightarrow [0,1] \) which is the same as a cumulative probability distribution function over \([0,1]^m\) such that each marginal distribution, of one dimension corresponds to a random variable \( U_k \) uniformly distributed over \([0,1]\), \( k=1,\ldots,m \). Copulas can be used to solve elegantly the important problems of determining all possible joint distributions given specified marginals. See [22].
For simplicity here, define a co-copula \( c \) or as a mapping \( \epsilon_{\text{co}}: [0,1]^{n} \to [0,1] \) which coincides with the disjunction probabilities corresponding to the conjuction ones for some copula. Thus if \( U_i \) is any r.v. uniformly distributed over \([0,1] \), for \( i=1, \ldots, n \), and \((Y_1, \ldots, Y_n) \) has some legitimate joint distribution, the \( \epsilon_{\text{co}} \) defined as follows will be a copula and \( \epsilon_{\text{co}} \) defined below will be the co-copula corresponding to \( \epsilon_{\text{co}} \):

\[
\epsilon_{\text{co}}(x_1, \ldots, x_n) = \min \{ x_i \} , \quad \forall x_i \in [0,1].
\]

Similarly, define a t-conorm as the decorman transform of some t-norm

\[
\epsilon_{\text{or}}(x_1, \ldots, x_n) = 1 - \epsilon_{\text{and}}(1-x_1, \ldots, 1-x_n), \quad \forall x_i \in [0,1].
\]

for all \( x_1, \ldots, x_n \in [0,1] \). Also, define an archimedean t-norm as a t-norm which for all \( 0\lt x < 1 \),

\[
\epsilon_{\text{and}}(x,y) < y ;
\]

dually, define a t-conorm to be archimedean iff

\[
\epsilon_{\text{or}}(x,y) > x ,
\]

for all \( 0\lt x < 1 \).

Consider some examples of conjunction and disjunction function pairs being copulas or t-norms with co-copulas or t-conorms.

First, it should be noted that \((\min, \max)\) and \((\text{prod}, \text{probsum})\) are the only such functions which are both copulas, co-copulas, and t-norm, t-conorm pairs simultaneously; the latter pair is also archimedean, where "prod" denotes ordinary arithmetic product, while "probsum" denotes formal probability "sum" (displaying modularity of probability) as the copular transform of product. (See [23], Section 4.)

\((\text{prod}, \text{sum})\) is a non-deorgan archimedean pair, where "sum" is to be interpreted as ordinary arithmetic sum, but bounded by unity; the latter is a t-conorm but not a co-copula.

Finally, to complete this brief preliminary discussion, the important canonical representation theorem for archimedean pairs of t-norms, t-conorms, states that for any such pair \((\epsilon_{\text{and}}, \epsilon_{\text{or}})\), there always exists a corresponding continuous non-increasing function \( h: [0,1] \to [0,1] \), with \( h(1) = 0 \) and \( R^n \) denoting the extended real line including \( -\infty \), such that, assuming the above pair is also demorgan,

\[
\epsilon_{\text{and}}(x_1, \ldots, x_n) = h^{-1}(\min(h(y_1), \ldots, h(y_n))) ; \quad \text{for all } y_i \in [0,1].
\]

Conversely, any such \( h \) as above generates a legitimate archimedean pair, where the t-norm part is given in (6.28).

Next, for convenience define for all \( i,j \)

\[
\epsilon_{\text{or}}(0,1) = (D-S \oplus Z) ; \quad \epsilon_{\text{or}}(0,1) = (D-S \oplus Z_i) ;
\]

\[
0 \leq Z \leq Z \leq H ; \quad \epsilon_{\text{or}}(0,1) = (Z_i \oplus D-S \oplus H_j). \]

Then make the following semantic evaluation of \( \epsilon_{\text{or}} \) preserving the formal structure in theorem 4:

\[
\text{pos}(q, q) = \text{pos}(q, D-S \oplus H_j) = \phi_{\text{or}}(\phi_{\text{or}}(\phi_{\text{or}}(c(0), c(1)), c(1)), c(1)), c(1)), c(1)), c(1)).
\]

In particular, the evaluation of \( \epsilon_{\text{or}} \) using Tadeh's original fuzzy set theory or fuzzy logic is easily seen to be a special case of (6.31), when

\[
\epsilon_{\text{or}}(\epsilon_{\text{or}}(c(0), c(1)), c(1)), c(1)), c(1)).
\]

More generally, the PACT algorithm [12], briefly mentioned previously, can also be shown to be essentially a special case of the data fusion evaluation given in (6.31), where now \( \phi_{\text{or}} \) and \( \phi_{\text{or}} \) are in certain parameterized families of conjunction and disjunction functions. In the PACT algorithm, data association or "correlation" is to be determined to hold or not for a feasible pair of developing track histories, where in addition to geolocation information, present may be other attribute forms. A typical example is where \( Z \) represents the following potential matching attributes for the two tracks \( \hat{q} \#2 \):

\[
Z = \begin{cases} \text{geolocation parameters for #1, for #2 sensor system parameters for #1, for #2 null lengths for #1, for #2, classifications for #1, for #2} \end{cases}
\]

Also, for this example, \( h \) (denoted in [12] by \( \epsilon \)) represents correlation level between \#1 and \#2 (between \( \epsilon \) and 1 when evaluated), while \( \hat{q} \#1 \) is assumed and \( \hat{q} \#2 \) represents observed (in error) counterpart of \( \hat{q} \#2 \). Then the inference rules of (6.29) correspond to some expert-derived (or derived by other analytic or physical considerations) relation between some combination of degrees of matching attributes in general with possible correlation levels \( h \); the terms \( \text{pos}(\hat{q} \#1) \) represent error distributions between true and observed auxiliary attributes. PACT can operate upon a mix of probabilistic information and attributes and linguistic-based information and attributes, as shown in (6.33), where typically the first, second, and possibly the third entries are in stochastic form, while the remaining entries are narrative-based and given in natural language. The basic PACT output, before further integration into an overall tracking-correlator design, is the posterior description of correlation-based upon observed or reported data involving the track history pair in question, as is represented in (6.31) by \( \text{pos}(\hat{q} \#1, \hat{q} \#2) \).

On the other hand, if we choose

\[
\epsilon_{\text{or}} = \text{prod} \quad \text{or} \quad \epsilon_{\text{or}} = \text{sum} ,
\]

262
Next, consider the evaluation of data fusion as given in (6.31) when $\mathcal{A}$ is any copula and $\mathcal{Q}$ or is the co-copula determined by $\mathcal{A}$ as in (6.21), compatible with the data fusion problem as modeled here. Thus, similar to the specific example given in Section 5, but with generality in mind, using (6.29)-(6.30), let $(\text{fixing} D$ and $S)$

$$\text{dom}(A) = (\{U_i\}_{i \in I})^* \text{ dom}(Z) = \{Z_i\}_{i \in I}, \quad (6.35)$$

$$\text{dom}(B) = (\{U_j\}_{j \in J})^* \text{ dom}(H) = \{J_{ij}\}_{i \in I, j \in J}, \quad (6.36)$$

where $I$ and $J$ are suitably chosen index sets. Let

$$U \subset (U_i, U_j)_{i \in I, j \in J} \quad (6.37)$$

be any stochastic process where each marginal $U_i$ and $U_j$ is some random variable uniformly distributed over $[0,1]$. Then define random subsets $S_{\alpha}$ and $S_{\beta}$ of dom($\alpha$) and dom($\beta$) by, for all $i \in I, j \in J$,

$$\alpha_i \in S_{\alpha} \iff U_i \leq \text{pos}_{\alpha}(\alpha_i) \quad (6.38)$$

and

$$\beta_{ij} \in S_{\beta} \iff U_i \leq \text{pos}_{\beta}(\beta_{ij}) \quad (6.39)$$

Note that if the $U_i$ are all identical and, separately, the $U_{ij}$ are all identical, then

$$S_{\alpha} = S_{\beta} (2), \quad S_{\beta} \in \mathcal{R} \quad (6.39)$$

as given in Theorem 3. Determine $\mathcal{A}, \mathcal{B}$ or through $U_i$.

Then it follows that the evaluation of data fusion in (6.31) becomes, using (6.21), (6.35)-(6.39),

$$\text{pos}(Q_{\alpha} Q_{\beta}) = \sum (-1)^{\text{card}(K)+1} \text{ pos}_{\alpha}(\{U_i \text{ pos}_{\alpha}(\alpha_i), U_{ij} \text{ pos}_{\beta}(\beta_{ij})\}) \quad (6.40)$$

where for all subsets $K$

$$N_{K,j} = \{U_i \text{ pos}_{\alpha}(\alpha_i), U_{ij} \text{ pos}_{\beta}(\beta_{ij})\} \quad (6.41)$$

But, using the Poincaré expansion of probabilities, (6.40) and (6.41) yield

$$\text{pos}(Q_{\alpha} Q_{\beta}) = p(\alpha \in S_{\alpha} \wedge \beta_{ij} \in S_{\beta} \wedge \text{ dom}(\alpha) \wedge \text{ dom}(\beta)) \quad (6.42)$$

where

$$\text{dom}(\alpha) = \{U_i \text{ pos}_{\alpha}(\alpha_i), U_{ij} \text{ pos}_{\beta}(\beta_{ij})\} \quad (6.43)$$

Noting that the expression in the right side of eq. (6.31) can be written in a natural way in terms of possibilities analogous to that in (6.42), we obtain the following result:

Theorem 5.

Given variables $D, S, H$ and auxiliary variable $Z$ as before, then under the assumptions leading to eq.

$$(5.31) \text{ and assuming the constructions in } (6.39)-(6.39), \text{ it follows that for all } \mathcal{A},$$

$$\text{pos}(Q_{\alpha} Q_{\beta}) = p(\alpha \in S_{\alpha} \wedge \beta_{ij} \in S_{\beta} \wedge \text{ dom}(\alpha) \wedge \text{ dom}(\beta)) \quad (6.44)$$

where $\text{plaus}_{\alpha} S_{\beta} \text{ denotes the plausibility or upper probability measure with respect to random subset } S_{\alpha} \wedge S_{\beta} \text{ of dom(} \alpha \text{ dom(} \beta \text{).}$

Remarks.

For related results and general background, see [17], Chapters 3 and 4. Shafer [24] independently has developed use of plausibility measures and other bijectively related functions, such as "belief" and "doubt" measures in modeling combination of evidence problems. However, Nguyen [25] has emphasized, via Choquet's Capacity Theorem, which characterizes such functions in terms of both their random set connections and their generalized Poincaré expansion forms, that such "measures" require full specification of the associated random (subsets). Contrast such modeling with that employing possibility functions in a general multiple logic context, as given above, using some pair of conjunction and disjunction functions. As shown in the previous section and here, the latter approach only in effect requires knowledge of the one point coverage functions of the relevant random sets involved. Even in Theory 5, where an equivalent plausibility description is given, it is only specified over the $\mathcal{A}$'s. In short, any plausibility measure is determined by the incidence function of some appropriate random set with all ordinary subsets of the space--any belief measure is determined by the supersets--by the random set--any doubt measure is determined by the subset--by the random set.

In any case, Theorem 5 shows that a homomorphism relation exists between the possibilistic incidence form of data fusion evaluation as given originally in (6.31) and the corresponding equivalent probability form in (6.44). If in (6.37), $U_i$ instead of being chosen identical for all $U_i$ and all $U_{ij}$ separately, is such that all $U_i$ are statistically independent of each other and of all $U_{ij}$ which are also all independent, then the resulting $S_{\alpha}$ and $S_{\beta}$ are not only statistically independent, but are the maximal entropy one point equivalent representatives for pos_{\alpha} and pos_{\beta}, respectively. (See [20].)

In another direction, the following important asymptotic result holds for the data fusion expression in (6.31): Noting that variable $Z$ can represent a complex of attributes, some probabilistic in nature, others linguistic-based in nature, so that their descriptions can be possibilistic but not probabilistic, partition $Z$ accordingly into

$$Z = (Z^*, Z^*) \quad (6.45)$$

where $w_1, w_2, \ldots, Z$ is the vector of $p_{\alpha i}$ attributes and $Z^*$ is the vector of $\alpha_1, \ldots, \alpha_n$ linguistic ones. Note that by the canonical representation theorem, mentioned in Section 6 (see eq. (6.28)), any archimedean $\alpha$-measure, $\alpha$-norm pair is chosen for the evaluation in (6.31), then $\text{pos}(Z)$ becomes a monotone transform $\text{r}_{\alpha}$, say, for generator function $h$ of $\alpha$, of a sum of terms over $i$, where

$$h(x) = h^{-1}(\min(h(0), x)) \quad (6.46)$$
for all \( z \in \mathbb{R}^n \), and the \( i^{th} \) term, \( i = (i_1, i_2) \), is
\[
h\left(1-z(p_{a_1}(z'_{1} z'_{2})_{E_{2}}(z', H_2))\right),
\]
where \( a \) is partitioned as \( z \) into \( (a', a'') \) and
\[
5(z'_{1}, H_{2}) \subseteq \{1-z(p_{a_1}(z'_{1} z'_{2})_{E_{2}}(z', H_2))\}.
\]

Note that \( \text{dom}(z') \) is finite, as well as all other \( a \) of relevant variables. In order for finite argument functions \( z'_{1} \) and \( z'_{2} \), to be well-defined. In some cases, these finite domains are the result of discretizations and truncations of initial natural domains which are infinite and/or continuous, especially those corresponding to continuous probability density functions. In this context, suppose all probabilistic attributes, making up \( z' \) are such that they correspond to actual probability density functions which have all been so discretized as above. Denote the symbol
\[
\lim \text{dom}(z') \text{ to mean that the limit of } \text{dom}(z') \text{ will be taken, if it exists, as } \text{dom}(z') \text{ and } \text{dom}, \text{ are refined so that all cell sizes approach point limits and thus } \text{dom}, \text{ approaches a joint p.d.f. form corresponding to random variable } (z' | \text{DBS}). Then we can show the following:

Theorem 6. Asymptotic limiting form for data fusion. (See [26].)

Suppose that all of the above assumptions hold together with some mild analytic conditions for the Archimedean t-norm, t-norm pair \( t_{a} \), for chosen for the data fusion evaluation (6.41).

Then
\[
\lim \tau \left( p_{a_{1} a_{2}}(z') \right) = \tau_{a_{1} a_{2}} \left( \xi_{a_1 a_2}(z' | H_{1}) \right),
\]
where
\[
\xi_{a_1 a_2}(z' | H_{1}) \triangleq \frac{1}{\text{dom}(z')^{a}}
\]
and all \( a \) and \( z \) for chosen for the data fusion evaluation (6.41).

Thus, up to essentially monotone transforms, the limiting form of the data fusion computations here is an averaged value of the data fusion within (early) fixed domain attributes \( z' \). For simplification to the classical integral and (continuous) version of (4.2) occurs when the fixed non-probabilistic attributes component is missing. These results can be used for data checks when modeling via (6.31). (See, e.g., [12].)

For other controversies involving probability vs. possibility vs. Dempster-Shafer belief, doubt, etc., see [17], (especially, Chapter 10).

7. STRUCTURE FOR DATA FUSION: THE GENERAL COMBINATION OF EVIDENCE CASE

Let us return to the formal language aspect of data fusion as given in Theorem 4. In general knowledge-based systems such as medical diagnosis ones consist of a collection of inference rules corresponding to \((H, 2, 0, 1)\) linking either observed data, such as \( 0, 0, 5 \) or portions of intermediate variable \( Z \) with other portions of \( Z \) or with diagnoses directly, played by the role of variable \( H \). Similar comments hold for the attribute variability term \( q(Z, 0, 5) \).

The somewhat similar, but more general structure for such systems is given in eq. (7.3).
\[
\{1-z(p_{a_1}(z'_{1} z'_{2})_{E_{2}}(z', H_2))\} \subseteq \{1-z(p_{a_1}(z'_{1} z'_{2})_{E_{2}}(z', H_2))\}
\]

Next, to complete the general data fusion theory again, referring to figure 4, we must choose an ALDP, i.e., a pair consisting of a compatible choice of formal language followed by a semantic evaluation or logic.

Consider then as reasonable candidates for the evaluation of (7.1). ALDP 1,2,3 as in Figure 4.

Example 1. ALDP 1.

ALDP 1 = (boolean algebra \( \mathcal{B} \) with (6.14) valid for \( \mathcal{B} \), Classical (two-valued) Logic)

The calculus of relations for implications for the formal language part here, \( \mathcal{B} \) boolean with (6.14):
\[
\begin{align*}
\text{For all } a_1, a_2, a_3, a_4, \text{ and } a_5, b_5, c_1, & \\
\text{m} & \\
\text{m} & \\
\text{m} & \\
\text{m} & \\
\text{m} & \\
\text{m} & \\
\text{m} & \\
\text{m} & \\
\end{align*}
\]

Thus, if \( \beta_1 \equiv \neg \beta_2 \equiv \beta_3 \equiv \beta_4 \equiv \beta_5 \equiv \beta_0 \), and (7.2) and (7.3) become homomorphic relations for fixed antecedents:
\[
\begin{align*}
\text{m} & \\
\text{m} & \\
\text{m} & \\
\end{align*}
\]

But negation is in general not a homomorphic relation:
\[
\text{(6.10) a\not\equiv a' \triangleright (6.11) a \not\equiv a')}
\]

Also, for all \( a_0, a_1, a_2, \) and \( a_3 \),
\[
\begin{align*}
\text{(1.0) a_0 a_1 a_2 a_3} & \equiv (a_0 a_1 a_2 a_3) \equiv (a_0 a_1 a_2 a_3) \equiv (a_0 a_1 a_2 a_3) \\
\end{align*}
\]

Consider then the semantic evaluation part. Denoting the evaluation of any proposition variable \( a \), having domain of possible (or not) values in \( \text{dom}(a) \) as function \( p_{a} : \text{dom}(a) = \{0, 1\} \), for any \( a \) and \( a' \),
\[
\begin{align*}
\text{poss}(a) & = 0, i.e., a_1 \not\equiv a_2 \\
\text{poss}(a) & = 1, i.e., a_1 \equiv a_2 \\
\text{and variable } a & \text{ can be identified with a subset of } \Omega:
\end{align*}
\]

with \( \text{poss} = \text{playing the role of an ordinary set membership function}. Then, Classical Logic, as a truth-functional logic (see, e.g., [27] for further elaboration)

The following homomorphic forms, for all proposition variables (and similarly for all propositions) \( a, b, c \),
\[
\begin{align*}
\text{poss} & = \max(\text{poss}_a, \text{poss}_b), \\
\text{poss} & = \min(\text{poss}_a, \text{poss}_b), \\
\text{poss} & = 1 - \text{poss}_a.
\end{align*}
\]
where

\[ \text{poss}_{\mathcal{B}}(a) = \max(1 - \text{poss}_{\mathcal{B}}, \text{poss}_{a}) \]

(7.14)

and hence

\[ \text{poss}_{\mathcal{B} \& a} = \max(1 - \text{poss}_{\mathcal{B}}, \text{poss}_{a}) \]

(7.15)

where in all of the above equations, all functions are understood to be evaluated at arbitrary common domain points component-wise.

The usual presentation - which is equivalent - is through truth tables, but the above display allows for natural generalizations to Zadeh's (min-max) fuzzy logic in ALD2.

It also follows that the semantic evaluation of the data fusion form in (7.1) becomes here:

\[ \text{poss}(Q|D) = \text{poss}_{D,S} \& H_{(j)} \]

(7.16)

\[ = \max \left( \min \left( \text{poss}(D|Z_{k1} Z_{k2} \cdots Z_{km}) \right) \right) \]

(7.17)

where for all \( k, l, j \)

\[ \int_{k_{l}} \text{poss}_{j} \left( Z_{k_{l}1} H_{j}; D, S \right) \]

(7.18)

\[ \int_{k_{l}} \text{poss}_{j} \left( Z_{k_{l}1} H_{j}; D, S \right) \]

(7.19)

and where the expressions in (7.16) and (7.17), if necessary, can be evaluated further using (7.10)-(7.14).

But since we have here a simple two-valued logic, eq.(7.10) reduces to:

\[ \text{poss}(Q|D) = 1 \quad \text{iff there is some attribute value} \]

\[ Z_{i} \quad \text{such that for each} \]

\[ k, l, j \]

\[ \text{that when evaluated at} \]

\[ Z_{k_{l}} H_{j}; D, S \]

\[ \text{is true, i.e.,} \]

\[ \text{poss}_{j} \left( Z_{k_{l}} H_{j}; D, S \right) = 1 \]

(7.19)

or equivalently, \( Z_{k_{l}} H_{j}; D, S \) all fire inference rule \( (j \& k) \text{ is false at this evaluation} \)

(7.19)

(7.19)

Otherwise, one can evaluate (7.1), by first directly applying the calculus of relations for inferences in (7.2)-(7.3)) and then evaluate the result semantically. Thus

\[ \text{poss}(Q|D) = \text{poss}(Q|D) \& H_{(j)} \]

(7.20)

\[ = \max(1 - \text{poss}(Q|D), \text{poss}(Q|D)) \]

(7.21)

where

\[ \text{poss}(Q|D) \]

(7.22)

and

\[ \text{poss}(Q|D) \]

(7.23)

where, in turn, (7.10)-(7.14) could be used to evaluate further poss(q) and poss(a), which of course should lead back to (7.10) and thus (7.10)-(7.19), as a check.

The philosophy of approach in this example is that for the modeling of data fusion, in the context or medical diagnosis, for example, although truth can only be 0 or 1, by introducing sufficiently many inference rules in the knowledge-based system, multiple-valued truth logics can be avoided.

Example 2. ALD2.

ALD2 = (modified boolean algebra \( \Omega \) with (6.14), Zadeh's (min-max) Fuzzy Logic)

As mentioned earlier (again, see Figure 4 and associated remarks in Section 2), "modified" boolean means a pseudo-complemented (distributive) lattice, or roughly a boolean-like system without the law of excluded-middle and all its consequences holding.

(See [28], pp. 14-16 for a related discussion. [28] as a whole also serves as a good introduction to Zadeh's Fuzzy Logic.)

The calculus of relations for implication for the formal language part here, \( \alpha \), is the same formally as that for \( \alpha \) as in Example 1, except for the following slight modifications given in the two statements (1), (11) below:

(1) The middle equation in (7.5) will be valid, provided that \( \alpha_{0} \leq \alpha_{0} \), i.e., \( \alpha = \alpha_{0} \alpha_{0} \), otherwise in general it is not true.

(II) Adjoin the term \( \alpha_{0} \alpha_{0} \) to the consequent of \( \alpha \) on the left-hand side of the equality for the far-right chaining equation in (7.5).

Then the same semantic evaluations proceed in formally the same way as for ALD1, but here the range of values of each possibility function is in the unit interval \([0,1]\), instead of being restricted to the set \([0,1]\), replacing (7.3). Thus eqs. (7.9)-(7.17) all remain valid here. Eq.(7.18) and eq.(7.19) are no longer valid in the context of ALD2. On the other hand, eqs. (7.20)-(7.22) hold here, with appropriate modifications following those in (1), (11) above.

Example 3. ALD3.

ALD3 = (boolean algebra \( \Omega \) with (6.14), Probability Logic)

Since \( \Omega \) is the same as in Example 1, all of the relations in eqs. (7.2)-(7.7) hold here also. On the other hand, the semantic evaluation aspect - Probability Logic - differs considerably from the two previous examples. In this non-truth-functional logic (see again [27], specifically Chapter 2, Sections 25 and 27 for background), we have the usual basic (finitely additive) probability properties, for a given probability measure \( \rho = [0,1] \), playing the role of the semantic evaluation poss in the two previous examples. (In order to use the more standard notation, \( \rho \) is used in place of \( \text{poss} \)). Only for purposes of comparisons the following well-known properties are given:

For all propositions \( \alpha_{0}, \alpha \in \Omega \),

\[ p(\alpha \cup \beta) = p(\alpha) + p(\beta) - p(\alpha \cap \beta) \]

(7.23)

the modularity property, extending to the\( \cup \)-expansion, used previously in this paper, are for all \( \alpha_{0} \leq \alpha \leq \beta \leq \gamma \leq \delta \leq \epsilon \leq \zeta \leq \eta \leq \theta \leq \iota \leq \kappa \leq \ldots \)

\[ p(\alpha \cup \beta) = p(\alpha) + p(\beta) \]

(7.24)

for \( \alpha \cup \beta \).

\[ p(\alpha \cup \beta) = 1 - p(\alpha \cap \beta) \]

(7.25)

\[ p(\alpha) = 0 \quad \text{p}(1) = 1 \]

(7.26)

resulting in the following evaluations for implication (by (6.14), for \( \rightarrow \)) and some less-known inequalities

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.27)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.28)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.29)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.30)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.31)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.32)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.33)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.34)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.35)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.36)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.37)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.38)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.39)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.40)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.41)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.42)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.43)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.44)

\[ p(\alpha \rightarrow \beta) = 1 - p(\alpha \cap \beta) \]

(7.45)
Involving conditional probabilities:
\[
p(\theta_0 \oplus \theta_0) = p(\theta_0 \vee \theta_0) = 1 - p(\theta_0 \wedge \theta_0) = p(\theta_0 | \theta_0) + p(\theta_0 | \theta_0) - p(\theta_0 \wedge \theta_0) = p(\theta_0 | \theta_0) + p(\theta_0 | \theta_0) + p(\theta_0)
\]
\[
\geq p(\theta_0) + p(\theta_0) - p(\theta_0)
\]
(7.20)

where the conditional probability is defined as usual as, e.g.,
\[
p(\theta_0 | \theta_0) = \frac{p(\theta_0 \wedge \theta_0)}{p(\theta_0)}\]
(7.29)

The above inequalities are strict, in general, and show that, basically, we cannot identify implication, as defined in the formal language (a) via eq. (6.14), with a "conditional object" as (7.21), (7.22), but differing in the used-to which can be further evaluated through use of (6.15), (6.16). For the evaluation of implication via (6.14) with conditional propositions, as pointed out in the discussion in the paper, we seek a formal implication via (6.14) with conditional objects, whose semantic evaluations as (7.30) are conditional probabilities; but in light of the above remarks, necessarily these entities lie outside of the original space of propositions \( \mathcal{P} \).

Returning to the data fusion form in (7.1), the semantic evaluation for Probability Logic becomes, using first (7.24) and then (7.5),
\[
p(\mathcal{Q} \mathcal{O}_j) = p(\mathcal{O} \vee \mathcal{O}_j) = \sum (-1)^{\text{card}(\mathcal{K})+1} \cdot p(\mathcal{Q} \mathcal{O}_j \mathcal{O}_j), \quad (7.31)
\]

which can be further evaluated through use of (7.20) (equality part) in conjunction with (7.23)-(7.26), where similar to (7.21), (7.22), but differing in the operations involving \( \mathcal{K}_j \)
\[
(\mathcal{K}_j \mathcal{K}_j \mathcal{K}_j \mathcal{K}_j) \cup (\mathcal{K}_j \mathcal{K}_j \mathcal{K}_j \mathcal{K}_j) \cup (\mathcal{K}_j \mathcal{K}_j \mathcal{K}_j \mathcal{K}_j)
\]
(7.32)

and
\[
(\mathcal{K}_j \mathcal{K}_j \mathcal{K}_j \mathcal{K}_j) \cup (\mathcal{K}_j \mathcal{K}_j \mathcal{K}_j \mathcal{K}_j) \cup (\mathcal{K}_j \mathcal{K}_j \mathcal{K}_j \mathcal{K}_j)
\]
(7.33)

Alternatively, by using both (7.4) and (7.5) from the calculus of inference relations, and then applying this, one obtains the same as (7.20), with "pos" replaced by "\( \mathcal{P} \)." Thus,
\[
p(\mathcal{Q} \mathcal{O}_j) = p(q(\mathcal{Q} \mathcal{O}_j), \mathcal{O}_j | \mathcal{O}_j) = p(q(\mathcal{Q} \mathcal{O}_j) + p(\mathcal{Q} \mathcal{O}_j) - p(\mathcal{Q} \mathcal{O}_j)
\]
(7.30)

which can be evaluated through the equality part of (7.27) or through the expansion
\[
p(\theta_0 \oplus \theta_0) = p(\theta_0) + p(\theta_0) - p(\theta_0) - p(\theta_0)
\]
(7.35)

for all \( \mathcal{Q} \theta_0 \in \mathcal{P} \), followed by use again of the basic properties of probability function \( p \) in (7.23)-(7.26).

Obviously, in the above schemes, the number of computations involving probabilities of the conjunctions of relevant events or propositions can be quite large and, as well, it may be difficult to evaluate each such conjunction, unless some simplified dependency or other relations are assumed for certain of the events. As a consequence, several techniques have been established for evaluating combination of evidence in a knowledge-based system, when marginally one has available estimates of probabilities, or related certainties or likelihoods or confidences, etc., for each of the inference rule forms \( f_j(k_j, k_j) \).

Some of these procedures are ad hoc in nature, others are more analytically based. For a compendium, see [29].

8. DATA FUSION AND CONDITIONAL OBJECTS

In Section 7, we have seen how a general inference rule structure for data fusion can be evaluated through three different approaches ALDP 1-3. In all of these, the key connector for inference \( \mathcal{Q} \) was interpreted in the formal language-components as \( \mathcal{Q} \) as given in eq. (6.14). On the other hand a natural - and commonly used - semantic evaluation for inference rules is through conditional probabilities. That is, the evaluation of a typical form \( f_j(k_j, k_j) \) is
\[
p(k_j | k_j) \forall j \)
where \( q_j \) is some choice of probability measure \( p \) over \( \mathcal{Q} \), the set of all events or propositions, which for purposes of simplicity, from now on is assumed to be a boolean algebra. With this choice of evaluation, apropos to the spirit of this paper, we seek a formal language which will be compatible with these evaluations, i.e., will form an ALDP.

However, as pointed out in the discussion in the previous section, centered around (7.27), one cannot identify implication via (6.14) with conditioning as evaluated in (7.30). The apparently commonly-held belief that such an identification can be made with no serious consequences, often called in the literature of logic as Stalnaker's Theorem [30], was attacked by Lewis [31] and independently by Calabrese [32]. The latter indeed showed, by use of a simple canonical expansion, that not only \( \mathcal{Q} \) in (6.14) would not work, but any boolean function of two variables could not be used to play the role of conditioning, compatible with conditional-probability evaluations.

Moreover, it would be particularly desirable, to replace this rather flawed situation, with an ALDP which would yield feasible computations for data fusion or at least be on the same order of complexity as ALDP 1, 2, 3. Note of course, if truly all inference rule antecedents are identical, as is the case essentially in Sections 4, 5, 6, then there is no real need to work with conditional objects, since all conditioned events can be simply considered as unconditional ones relative to their intersections with the fixed common antecedent, or one can stick with the interpretation of implication as in (6.14). Compatible with this result, note the homomorphic relations for implication \( \Rightarrow \) w.r.t. disjunction and conjunction but not negation - as given in eqs. (7.4), (7.5).

But, for the modeling of data fusion through inference rules with varying antecedents, no such-direct simplification occurs and the development of such conditional objects would address the problem. Although we have stated above that implication operator \( \Rightarrow \) for a fixed antecedent yields homomorphic relations for
theorem, but not (.), conditional probabilities are compatible with homomorphic relations holding for all three operations, for an fixed antecedent, i.e., obviously, for all \( a, b, c \in \mathcal{N} \),

\[
p(a \mid y_0) = 1 - p(a \mid y_0^c) = p(a \mid y_0^c), \quad (8.1)
\]

\[
p(a \mid y_0) \cup (b \mid y_0) = p(a \cup y_0^c), \quad (8.2)
\]

\[
p(a \mid y_0) \cap (b \mid y_0) = p(a \cap y_0^c), \quad (8.3)
\]

Recall also the operation \( + \), which in terms of \( \vee, \wedge \), \( (.) \) is,

\[
a \vee b = a 
\]

\[
\text{and conversely,}
\]

\[
a \vee b = a + b - a b. \quad (8.4)
\]

Thus there is a bijective relationship between \((a, \vee, +, \cdot, \cdot, (.)\)\), a boolean algebra and \((a, \vee, +, 0, 1)\), a boolean ring. (For further discussion and properties, see [33])

Furthermore, recall the Stone Representation Theorem ([33], Chapter 5) which establishes an order-preserving isomorphism between any given boolean ring and a corresponding boolean ring of all subsets of a fixed universal set say \( X \) where the correspondences hold:

\[
\begin{align*}
1 & \rightarrow x : \rightarrow \partial X \text{ (symmetric set difference)}; \\
0 & \rightarrow 0: \rightarrow \text{u (set-union)}; \\
\ast & \rightarrow n \text{ (set intersection)}; \\
( ) & \rightarrow c \text{ (set complement)}; \\
\langle \rangle & \rightarrow \text{a (subset relation)}
\end{align*}
\]

All following results can be interpreted in terms of ordinary subsets and the alternative boolean algebra or boolean ring structures.

Noting also, for any \( a, b, c \in \mathcal{N} \),

\[
p(a \mid b) = p(a \mid b^c), \quad (8.5)
\]

the next result shows that under quite rigid and simple conditions, conditional objects are essentially characterized:

Theorem 7. Characterization of conditional objects. [34]

Given boolean ring \( \mathcal{N} \), there is a unique space \( \mathcal{N} \) of smallest possible classes according to subset partial ordering of \( \mathcal{N} \) as the conditional objects \( a \mid y_0 \), \( (a \mid y_0) \), \( (a \mid y_0) \ldots \), for all \( a, b, c \in \mathcal{N} \), such that the measure-free counterparts of (8.1)-(8.3) and (8.8) hold. That is,

\[
(a \mid y_0) = (a \mid y_0), \quad (8.9)
\]

\[
(a \mid y_0) \cup (b \mid y_0) = (a \cup b) \mid y_0, \quad (8.10)
\]

\[
(a \mid y_0) \cap (b \mid y_0) = (a \cap b) \mid y_0, \quad (8.11)
\]

and equivalent to (8.9)-(8.11), one can require ens. (8.11) and

\[
(a \mid y_0) + (b \mid y_0) = (a \cup b) \mid y_0, \quad (8.12)
\]

(8.11) and

\[
(a \mid y_0) \mid y_0 = (a \mid y_0) \mid y_0. \quad (8.12)
\]

Specifically, such conditional objects constitute all possible principal ideal cosets of ring \( \mathcal{N} \), where for any \( a, y_0 \in \mathcal{N} \),

\[
(a \mid y_0) = (a \mid y_0). \quad (8.12)
\]

Thus, for a fixed antecedent, even though, as stated earlier the resulting conditional objects could be identified as subsets or subevents of the antecedent (noting Stone's Representation Theorem), nevertheless the actual algebraic structures of these entities will be of non-trivial use: Suppose we wish to perform boolean operations on conditional objects with differing antecedents; how can this be accomplished, compatible with the results in Theorem 7?

Previously work in this direction includes: Hall in [37], who extended some of Boole's original ideas and developed essentially the same entities as reduced here, but from a different and more complicated perspective, with relatively little attention paid to developing operators among conditional objects with different antecedents, using the technique of universal algebras and "partially defined" operators; Domotor [38], who following the direction of "qualitative probability structures", as used in preference orderings and subjective probability, developed rather complicated expressions for combining conditional objects; not realizing the rich structure inherent in such entities; Adams [39], among others in the literature, who considered "conditional logics" which appear to be somewhat related to the concept produced here, but differ considerably in structure, and Calabrese [32] who was apparently the first to attempt to develop directly conditional objects from a logical-consequence viewpoint, which can be shown to be equivalent to that given here([36, Section 2]), but Calabrese proposed ad hoc definitions for boolean operations on conditional objects with varying antecedents.

In the approach taken here, developing all results from first principles considerations, the required operations upon conditional objects are defined simply as the natural class or component-wise extensions of the original operations. Thus, for example, let \( a, b, c \in \mathcal{N} \) arbitrary. The natural class extension of \( * \) applied now to \( (a \mid y_0) \ast (y_1 \mid y_0) \), noting each conditional object is in reality via (8.11) a subset of \( \mathcal{N} \), yields:

\[
(a \mid y_0) \ast (y_1 \mid y_0) = (a \ast r)(c \mid y_0) \ast (y_1 \mid y_0), \quad (8.14)
\]

The basic structure of the conditional object extension \( \mathcal{N} \) of \( \mathcal{N} \) is summarized next.

Theorem 8. Basic Structure of \( \mathcal{N} \) is summarized next.

\[
\tilde{\mathcal{N}} = (0)(\tilde{\mathcal{N}}), \quad (8.15)
\]

(1) In terms of quotient rings,

\[
\tilde{\mathcal{N}} = (0)(\tilde{\mathcal{N}}), \quad (8.16)
\]

(11) Conditioning as defined here can be identified essentially as the functional inverse of one-sided conjunction, i.e., conditional objects \( a \mid y_0 \) all sat-
is by the basic relation analogous to (7.29) for conditional probabilities and a related condition:

\[(a_0 | y_0) \cdot \gamma_0 = a_0 \cdot \gamma_0 \quad (8.17)\]

\[\{a_0 | y_0\} = \{x | x \in \mathcal{N}, x \cdot \gamma_0 = a_0 \cdot \gamma_0\}. \quad (8.18)\]

(iii) The natural class extension of all boolean operations from \(\mathcal{N} \times \mathcal{N}\) are well-defined/closed with ring-like properties, i.e., in the same previous sense, \(\mathcal{N}\) is a modified boolean algebra.

(iv) Since for all \(a_0 \in \mathcal{N}\), (8.14) shows immediately that

\[(a_0 | 1) = (a_0 \cdot 1). \quad (8.19)\]

(v) Also, partial order \(\leq\) defined over \(\mathcal{N}\), characterized by, for any \(a_0, b_0 \in \mathcal{N}\),

\[a_0 \leq b_0 \iff a_0 = b_0 \iff b_0 = b_0 \land a_0 \land b_0 \quad (8.20)\]

can be extended directly to \(\mathcal{N}\) with the same characters in them. (8.20) where (unconditional) objects in \(\mathcal{N}\) are replaced by conditional ones in \(\mathcal{N}\). Then, combining this with (iii) and (iv) establishes \((\mathcal{N}, \lor, \land, \cdot, \lnot)\) as a natural extension of its unconditional counterpart \((\mathcal{N}, \lor, \land, \cdot, \lnot)\).

(vi) A basic calculus of operations is, in addition to the properties in (8.9)-(8.13) for any \(a_1, \gamma_1 \in \mathcal{N}\),

\[\begin{align*}
\forall \{a_1 | \gamma_1\} &= \{x \land \gamma_1 \cdot a_1 | \gamma_1 \cdot x \ \land \gamma_1\}, \quad (8.21) \\
\forall \{a_1 | \gamma_1\} &= \{a_1 | \gamma_1 \cdot \gamma_2 | \gamma_1 | \gamma_2\}, \quad (8.22) \\
\forall \{a_1 | \gamma_1\} &= \{a_1 | \gamma_1 \cdot \gamma_2 | \gamma_1 | \gamma_2\}, \quad (8.23)
\end{align*}\]

Noting the reductions of (8.21)-(8.23) as in (8.9)-(8.12), it follows that all boolean operational extensions over \(\mathcal{N}\) coincide with corresponding coset operations when restricted to a fixed quotient ring, here \(\mathcal{N} \cdot \gamma_0\).

(vii) As a special case of (8.22), the following chaining holds for all \(a_0, \beta, \gamma_0 \in \mathcal{N}:

\[(a_0 \cdot \beta) | \gamma_0 = (a_0 | \beta) \cdot (a_0 | \beta) | \gamma_0. \quad (8.24)\]

Proof: The most difficult proof is that of (8.22). A sketch of the proof for the case \(n = 2\) is given in [35], Theorem 3.1; a full proof is presented in [34] where all other proofs are also given.

Remarks

Apropos to Theorem 8(i), it follows that all results in the theory and application of linear (w.r.t. \(\cdot\) over \(\beta\)) boolean equations, such as presented in [40], can be reinterpreted in terms of conditional objects. Extensions of the concept of conditioning to more general structures than boolean, such as modified boolean, or for Neumann regular, or to a category theory setting, have been considered [34].

Many other mathematical properties have been derived for conditional objects, including: characterizations for iterated conditional objects, i.e., conditional objects whose antecedent and consequence are also conditional objects; extensions of Stone's Representation Theorem to conditional objects; development of an outer approximation technique for closure for non-boolean functions, including arithmetic operations over conditional objects; relations established between ordinary conditional random variables and a randomized version of conditional objects; and establishment of various probabilistic connections, such as measure-free independence; measure-free bayesian and sequential learning forms; and the proof that the extension of any probability measure \(p : [0,1] \to [0,1]\) through eq. (7.30) yields the conditional extension of a monotone function. (Again, see [36]-[38], for further details.)

Most importantly here, analogues of calculus of relations for ALDP 1 (eqs. (7.2)-(7.5)) hold for conditional objects, as Theorem 8 shows. Moreover, the hypotheses for Theorem 4 all hold here. At this point let us define ALDP 4, for a given boolean algebra \(\mathcal{N}\) as simply

\[\text{ALDP } 4 = \left(\overline{\mathcal{N}}, p\right)\quad (8.25)\]

where \(\overline{p} : [0,1] \to [0,1]\) is the conditional probability extension of \(p : [0,1] \to [0,1]\), as mentioned above and where implication is interpreted as conditioning, i.e., for all \(a_0, b_0 \in \mathcal{N}\),

\[\left(\overline{b_0} \implies a_0\right) = \left(a_0 | \overline{b_0}\right) \quad (8.26)\]

(Note that implication or conditioning here is restricted to be upon unconditional elements, i.e., elements of \(\mathcal{N}\), not upon other properly-conditional objects. Some results indicate a possible identification of iterated conditional forms with simple conditional objects ([36], Section 4); so that in a sense this restriction may be unnecessary.)

Finally, consider use of ALDP 4 in evaluating data fusion expressions \(q\) in (7.1):

Direct use of (8.21) and (8.22) show that

\[p(\overline{q}_0 \cdot q_3) = p\left(\overline{q}_0 \cdot \sum_{k=1}^{m} \left(k_i \cdot k_{ij} \cdot k_{ij}\right)\right) \quad (8.27)\]

\[= p(\overline{q}_0 \cdot q_3) = p(\overline{q}_0 \cdot \sum_{k=1}^{m} \left(k_i \cdot k_{ij} \cdot k_{ij}\right)) \quad (8.28)\]

etc., where \(q\) is given in eq. (7.21) and

\[\overline{q}_0 \cdot \sum_{k=1}^{m} \left(k_i \cdot k_{ij} \cdot k_{ij}\right) = \overline{q}_0 \cdot \sum_{k=1}^{m} \left(k_i \cdot k_{ij} \cdot k_{ij}\right) \quad (8.29)\]

Thus, due to the calculus of operations given in Theorem 8, computations for data fusion using ALDP 4, with implication interpreted as a conditioning compatible with conditional probabilities, appears no more complex than that for the other choices of ALDP's.

9. CONCLUDING DISCUSSION

Summary

This paper presents a number of results contributing toward a cohesive top-down theory of data fusion.

In Section 1, a general overview of the data fusion problem is presented, with the conclusion that data fusion is identifiable as the combination of evidence occurring within decision nodes of \(C\) systems. In Section 2, qualitative relations are established pinpointing the role of data fusion in \(C\) systems specially as perceived by the author in previous work (see Figures 1, 2, 3), where data fusion is a process within a \(C\) decision-maker node intermediate with incoming "signal" detection and hypotheses selection.
Also, the concept of an ALDP (algebraic logic description-pair) is introduced as part of the total evaluation procedure involving data fusion (Figure 4). Three important examples of ALDP's are given, corresponding to classical logic, fuzzy logic, and probability logic. In all, implication is interpreted as a disjunction-of a negation and affirmation. A particular quantification of operations involving implications is reviewed for each ALDP and then applied to the evaluation of data fusion (Examples 1,2,3). Finally, a fourth ALDP is determined in Section 8, based on interpreting the probability rules through conditional probabilities. For consistency, this requires the full development of a calculus of "conditional objects" (Theorem 7.8). It is shown that this ALDP can be successfully used to evaluate data fusion probabilities with a level of complexity and calculations not exceeding that of the alternative methods, but here allowing rigorously for conditional probability interpretations of implications.

Future Work and Open Problems

In this paper the cognitive process phase has been used only implicitly in the evaluation of data fusion distributions. Future work will be directed toward more-direct use of mental imaging and related thought processes. This is because in addition to the "formalistics" involved in translating detected signals (or "signals", using the more general sense) as shown in the sequence of processes in Figure 4, heuristic processes may also be used, possibly shortening the process path or providing alternative means as for example in TI (Natural Intelligence).

Alternative structures for data fusion may also be investigated — as opposed-e.g., to that given here in (6.16) or (7.1) — in formal language form. Recursive computations for general data fusion may also be possible, analogous to the well-known Kalman filter or related maximum likelihood forms. In a similar vein, progressive change for hypotheses distributions based upon newly arriving data may also be monitored through entropy measurements. Details of this have yet to be established for the general case we seek here.

Finally, conditional object theory must certainly be developed further, if only to be able to better treat iterative conditioning and required approximations or truncations of computations for data fusion evaluation — when made through conditional probability evaluation of inference forms, i.e., through ALDP 4.

10. ACKNOWLEDGMENTS

This work was supported jointly by the Naval Ocean Systems Center Program for Independent Research (AIR) and the Directors of Laboratories, Technical Panel of 'Basic Research Group (JULTP-3880).

11. REFERENCES


