(1)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

AD-A241 121

to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and ...ion of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including ...ices, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, ...Project (0704-0188), Washington, DC 20503

2. REPORT DATE
August 1991

3. REPORT TYPE AND DATES COVERED
professional paper

4. TITLE AND SUBTITLE
A GENERAL THEORY FOR THE FUSION OF DATA

5. FUNDING NUMBERS
In-house funding

6. AUTHOR(S)
I. R. Goodman

DTIC
ELECTE
OCT 2 1991
S C D

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Ocean Systems Center
San Diego, CA 92152-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Ocean Systems Center
San Diego, CA 92152-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

91-12103

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

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 sytems 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$_1$ 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 C^3 system. Here such C^3 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."

Published in *DFS-87, 1987 Tri-Service Data Fusion Symposium Technical Proceedings*, Volume 1, June 1987.

14. SUBJECT TERMS

data fusion	probability logic	algebra
C^3	algorithm	

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAME AS REPORT

NSN 7540-01-280-5500

91 10 1 083

Standard form 298

21a. NAME OF RESPONSIBLE INDIVIDUAL	21b. TELEPHONE *(include Area Code)*	21c OFFICE SYMBOL
I. R. Goodman	(619) 553-4014	Code 421

DFS - 87

1987 TRI-SERVICE DATA FUSION SYMPOSIUM

TECHNICAL PROCEEDINGS

VOLUME I

JOHNS HOPKINS UNIVERSITY
APPLIED PHYSICS LABORATORY
LAUREL, MARYLAND

9 - 11 JUNE 1987

SPONSORED BY:

OFFICE OF NAVAL TECHNOLOGY
AND
JOINT DIRECTORS OF LABORATORIES
DATA FUSION SUB-PANEL

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-$\aleph_1$ 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 C^3 system. Here such C^3 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".

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 C^3 systems analysis. On the other hand, relatively little attention will be paid here to detailed 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 operational 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 several 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.

Examples 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" decision 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 C^3 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 C^3 network of nodes. Thus, any ongoing work in the C^3 arena, must effect data fusion efforts. Since 1978, the annual MIT/ONR Workshop on C^3 Systems - with its associated (unclassified) annual Proceedings - has served as one of the primary academic sources for generic C^3 studies. (See [13] for a partial survey of these efforts. See also [14] for a more thorough survey of C^3 work, where many abstracts, analyses, and comparisons and contrasts of C^3 theories and related work are given.) Surprisingly, relatively few comprehensive theories of C^3 systems have been produced, although many valuable papers have been written as a result of the C^3 Workshop on problems of distributive decision-making, hierarchical systems, communications and security, multiple target-tracking and correlation, and various miscellaneous game theoretic and warfare design problems. Among the few theories of C^3 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 intranodally within the context of an appropriately chosen overall C^3 system. That is, fusion occurs typically within decision-making nodes.

(2) All analysis and models of C^3 systems must include subanalysis and models for fusion processes. In particular, this applies to this author's proposed model for C^3 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, C^3 SYSTEMS, AND DATA PROCESSING

Previously, this author proposed a bottoms-up, microscopic, quantitative approach to general C^3 systems [15],[16]. In that approach, a generic C^3 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 comprised 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 is followed by a response or action taken towards other nodes, friendly or hostile. (See Figure 1.) Associated with

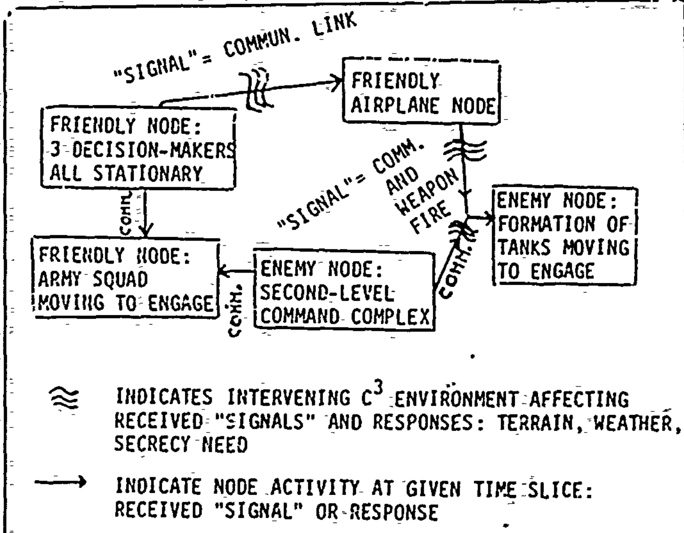

Figure 1. "Signal" and Response Activity in a Portion of Two C^3 Systems.

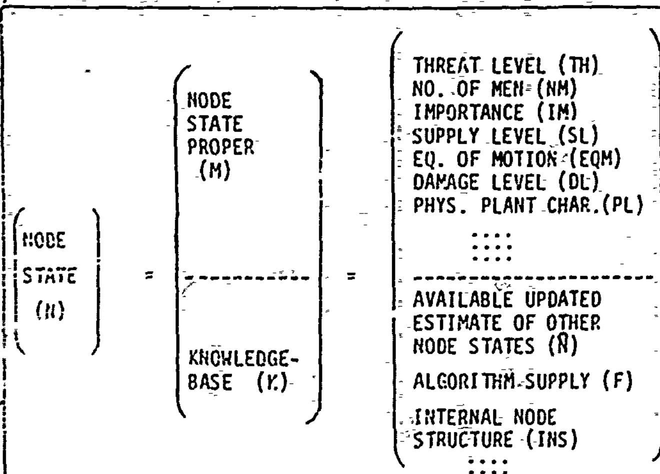

Figure 2. Components of C^2 Node States.

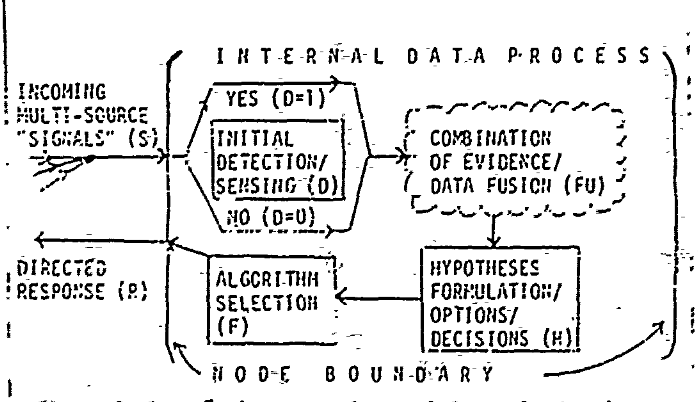

Figure 3. Data Fusion as an Integral Part of a Node's Data Processing Structure.

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 the other nodes (friendly or adversary). Also associated with each node is its internal "signal" processing design, as described in Figure 3. There, 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 infers (see [17], Chapters 1,2 and Figure I, page 14) that a series of underlying processes are 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 observors. 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: $\cdot$, for & ("and" or conjunction); v, for "or" (disjunction); ( )', for "not" (negation); $\ni$, 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 $\ni$ given as $\Rightarrow$, where $\beta \Rightarrow \alpha$ is identified as β' v α, for all wff's α,β.

ALDP 2 = (Modified boolean algebra = pseudo-complemented lattice, Zadeh's (min-max) Fuzzy Sets or Logic). As above, $\ni$ = $\Rightarrow$.

ALDP 3 = (Boolean algebra, Probability Logic); $\ni$ = $\Rightarrow$.

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, with little attention paid to the formal aspects prior to semantic evaluations. (Again, see [4],[5].)

Figure 4 summarizes the above analysis of data fusion.

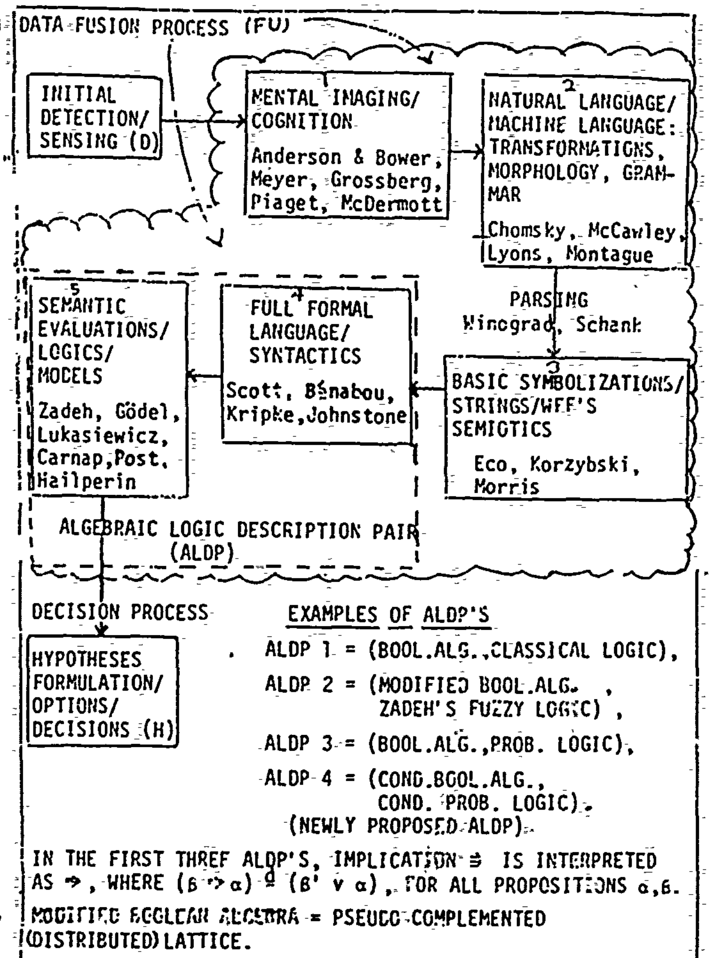

Figure 4. Subprocess Expansion of Data Fusion/Combination of Evidence Process Connecting Initial "Signal" Detections with Hypotheses Formulations.

3. DATA FUSION AS A QUANTITATIVE PART OF AN OVERALL C^3 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 C^3 systems has been established compatible with these qualitative formulations [15], [16]. Inputs to the structure consist basically of ten sorts of known relative primitive relations PRIM among the variables describing a C^3 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 (G) and additive noise (Q). 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 C^3 system a (friendly or hostile) and node number i, while k represents a discrete time index t_k. Specifically, the relation breaks down into 5 intranodal (within nodes) relations, 2 internodal (between nodes) or regression relations, and 3 prior relations for each C^3 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 in (probabilistic) terms of the other variables and node states through PRIM. Thus, we have:

Theorem 1. (See [15], Theorem 1.)

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

$$p(N_{g,k}) = \varphi_{g,k}(PRIM_k) \;, \tag{3.1}$$

where $\varphi_{g,k}$ is a computable functional involving a finite number of integrations and arithmetic operations upon the elements of $PRIM_k$ given in Table 4. ∎

$$PRIM_k \overset{d}{=} (PRIM_k^{(1,1)}, PRIM_k^{(1,2)}; PRIM_k^{(2)}) \;, \tag{3.2}$$

where for C^3 system a, g=(a,i), etc.,

$$PRIM_k^{(1,a)} \overset{d}{=} \Big((1)_{g,k_1}, \ldots, (5)_{g,k_1}, (8)_g, (16)_{g,0}\Big)_{\substack{\text{all } g,h,\\ 0 \le k_1 \le k}} \tag{3.3}$$

and where

$$PRIM_k^{(2)} \overset{d}{=} \Big((6)_{h,g,k_1+1}, (7)_{h,g,k_1+1}, (15)_{h,g,0}\Big)_{\substack{\text{all } g,h,\\ 0 \le k_1 \le k}} \;. \tag{3.4}$$

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

$$(1)_{g,k} = p(H_{g,k} | D_{g,k}, S_{g,k}) \;,$$

$$(2)_{g,k} = p(F_{g,k} | H_{g,k}) \;,$$

$$(3)_{g,k+1} = p(R_{g,k+1} | F_{g,k}, S_{g,k}, N_{g,k}) \;,$$

$$(4)_{g,k+1} = p(N_{g,k+1} | R_{g,k+1}, N_{g,k}) \;,$$

$$(5)_{g,k} = p(D_{g,k} | S_{g,k}, N_{g,k}) \;;$$

Table 1. Relative Primitive Intranodal Relations.

$$(6)_{h,g,k+1} = p(Q_{h,g,k+1}) \text{ with } G_{h,g,k+1} \,,$$

$$(7)_{h,g,k+1} = p(W_{g,k+1} = h | N_0) \;;$$

The basic internodal analysis is developed via additive nonlinear regression relation

$$(S_{g,k+1} | W_{g,k+1} = (h,k)) = G_{h,g,k+1}(R_{h,k}) + Q_{h,g,k+1} \,,$$

where variable $W_{g,k+1}$ indicates original possible possible node source for "signal" at time k, given reception by another node at k+1.

Table 2. Relative Primitive Internodal Relations.

PRIOR/INITIAL TIME

$$(8)_0 = p(N_0) \;,$$

$$(15)_{h,g,0} = p(R_{h,0} | W_{g,1} = h \;, N_0) \;,$$

$$(16)_{g,0} = p(S_{g,0} | N_0) \;.$$

Table 3. Relative Primitive Prior/Initial Relations

$$(9)_{g,k+1} = p(R_{g,k+1} | D_{g,k}, S_{g,k}, N_{g,k}) = \int\!\!\int_{(\text{over all } F_{g,k}, H_{g,k})} (1)_{g,k} \cdot (2)_{g,k} \cdot (3)_{g,k+1} \, dF_{g,k} \, dH_{g,k} \;,$$

$$(10)_{g,k+1} = p(N_{g,k+1} | D_{g,k}, S_{g,k}, N_{g,k}) = \int_{\begin{pmatrix}\text{over all}\\ R_{g,k+1}\end{pmatrix}} (4)_{g,k+1} \cdot (9)_{g,k+1} \, dR_{g,k+1} \;,$$

$$(11)_{g,k+1} = p(R_{g,k+1} | S_{g,k}, N_{g,k}) = \sum_{D_{g,k}=0}^{1} ((9)_{g,k+1} \cdot (5)_{g,k}) \;,$$

$$(12)_{g,k+1} = p(N_{g,k+1}, D_{g,k} | S_{g,k}, N_{g,k}) = (10)_{g,k+1} \cdot (5)_{g,k} \;,$$

$$(13)_{g,k+1} = p(N_{g,k+1} | S_{g,k}, N_{g,k}) = \sum_{D_{g,k}=0}^{1} ((12)_{g,k+1}) \;,$$

$$(14)_{h,g,k+1} = p(S_{g,k+1} | R_{h,k}) = p(Q_{h,g,k+1} = S_{g,k+1} - G_{h,g,k+1}(R_{h,k})) \,,$$

$$(15)_{h,g,k} = p(R_{h,k} | N_0, W_{g,k+1} = h) = \int\!\!\int_{\begin{pmatrix}\text{over all}\\ S_{h,k-1}, N_{h,k-1}\end{pmatrix}} (11)_{h,k} \cdot (16)_{h,k-1} \cdot (18)_{h,k-1} \, dS_{h,k-1} \, dN_{h,k-1} \;,$$

$$(16)_{g,k} = p(S_{g,k} | N_0) = \int\!\!\int_{\begin{pmatrix}\text{over all}\\ h, R_{h,k-1}\end{pmatrix}} (14)_{h,g,k} \cdot (15)_{h,g,k-1} \cdot (7)_{h,g,k-1} \, dh \, dR_{h,k-1} \,,$$

$$(17)_{g,k+1} = p(N_{g,k+1} | N_{g,k}, N_0) = \int_{\begin{pmatrix}\text{over all}\\ S_{g,k}\end{pmatrix}} (13)_{g,k+1} \cdot (16)_{g,k} \, dS_{g,k} \;,$$

$$(18)_{g,k} = p(N_{g,k} | N_0) = \int_{\begin{pmatrix}\text{over all}\\ N_{g,k-1}\end{pmatrix}} (17)_{g,k} \cdot (18)_{g,k-1} \, dN_{g,k-1} \;,$$

$$\boxed{p(N_{g,k}) = \int_{\begin{pmatrix}\text{over all}\\ N_0\end{pmatrix}} (18)_{g,k} \cdot (8)_0 \, dN_0 \;.}$$

Table 4. Structure of $\varphi_{g,k}$ in Theorem 1 Through Sequence of Calculations Involving $PRIM_k$

In turn, a simple two-person zero sum game can be established, called the C^3-decision game. Here, Player I corresponds to entire C^3 system a=1 (say, friendly) and Player II corresponds to entire C^3 system a=2 (say, adversary). In this game, a move by Player j corresponds to a choice (up to given constraints) of $PRIM_k^{(1,j)}$, j=I,II, and the resulting loss or utility due to any such joint move L_k is a function of the marginal updated node state distributions, according to Theorem 1 as

$$L_k(PRIM_k) = MOE_k(\{p(N_{g,k})\} | \text{all } g\})$$

$$= MOE_k(\{\mathcal{Q}_{g,k}(PRIM_k) | \text{all } g\}), \tag{3.5}$$

where MOE_k represents a single figure-of-merit, combining various measures of effectiveness (moe's) or performance (mop's) 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 moe's that could be used include; averaged measure of importance $\overline{IM}_{a,k}$; averaged measure of threat $\overline{TH}_{a,k}$; upper bound total entropy $\overline{ENT}_{a,k}$; and averaged measure of performance $\overline{ACC}_{a,k}$, all computable through $p(H_{g,k})$'s for C^3 system g, by use of Theorem 1. (See also [15], eqs.(59)-(62).) Then one could let

$$MOE_k = MOE_{1,k} - MOE_{2,k} , \tag{3.6}$$

where

$$MOE_{a,k} = \lambda_1 \cdot \overline{IM}_{a,k} + \lambda_2 \cdot \overline{TH}_{a,k} + \lambda_3 \cdot \overline{ENT}_{a,k} + \lambda_4 \cdot \overline{ACC}_{a,k} , \tag{3.7}$$

and the λ_i's are some predetermined weightings.

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

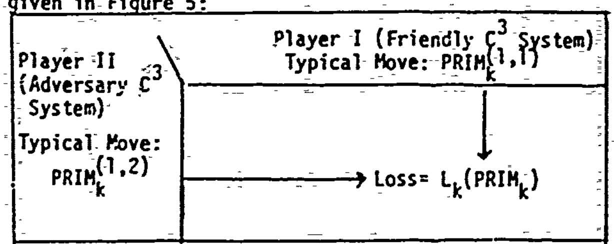

Figure 5. Symbolic Form for C^3 Decision Game.

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 [16] 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 initial sensing and hypotheses 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 :

$$P(FU) = p(H|D,S) , \tag{4.1}$$

where for reasons of convenience from now on we suppress the denotional-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 we are physically in a bunker-a C^3 node- S may be observed loud noise, with D=1 (definitely detected), and H could have possible domain values say $dom(H)=(H_1,\ldots,H_5)$ as given in Table 5.

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

Table 5. Typical Set of Values for dom(H).

Thus, dom(H) could serve as a legimate sample space, if conditional probability p(H|D,S) could be obtained for all possible values of H in 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	= nothing happening
Z_2	= accidental explosion in compartment #1
Z_3	= accidental explosion in compartment #2
Z_4	= enemy shot missile at us and it either hit us or just missed
Z_5	= none of the above situations hold

Table 6. Typical Set of Values for dom(Z).

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:

$$p(H=H_j|D\&S) = \sum_{i=1}^{5} p(H_j\&Z_i|D\&S)$$
$$= \sum_{i=1}^{5} p(Z_i|D\&S) \cdot p(H_j|Z_i\&D\&S) , \tag{4.2}$$

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

$$\text{"If D and S, then } H_j\text{"} \tag{4.3}$$

through the probability values for the expressions

$$\text{"If D and S, then } Z_i\text{" and "If } Z_i \text{ and D and S, then } H_j\text{"} \tag{4.4}$$

Of course, one need not use the above evaluation exactly to obtain useful equivalent values. As it stands, $P(Z_i|D\&S)$ can be interpreted as an error or variability probability for attribute Z_i while $p(H_j|Z_i\&D\&S)$ 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 $p(S|Z_i \& H_j)$ and the joint prior probability $p(Z_i \& H_j)$ are available, assuming here D=1, which by use of Bayes' theorem also yields $p(H{=}H_j|D\&S)$. One standard result is to assume the above probabilities are gaussian, which in the discrete problem here, must serve as very rough approximations- in addition, the sets dom(H) and 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, $p(H_j \& Z_i|S)$ takes on a generalized weighted least squares form. (See, e.g. [18].) The final result, $p(H{=}H_j|S)$, 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, suppose now that H,Z,S are variables such that any of the corresponding "sample spaces" do not truly contain 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 simple corresponding 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. Suppose in the above example in Section 4 (Table 5), the enemy could simultaneously mount the promised offense (H_2), yet also be feeling us out for peace (H_3), or, even additionally, wanting to negotiate (H_4). Thus, in that case, $dom(H)=\{H_1,\ldots,H_5\}$, as it stands, is <u>not</u> a suitable sample space of disjoint elementary events. Indeed, the elementary events H_j are not so elementary, many of them, due to complex causes, being overlapping! 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_j really represent complex phenomena 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 them,

a_3 = past 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;

$a \overset{d}{=} a_1 \times \cdots \times a_7$.

Then ideally, in turn, given enough of these factors, define rigorously the H_j's in terms of combinations of values of the a_k's. One simple approach is to determine the natural domains of values for the a_k's , $dom(a_k)$, k=1,...,7, letting

$$\mathcal{D} \overset{d}{=} dom(a_1)\times\cdots\times dom(a_7) \tag{5.1}$$

and

$$H_j = b_{j,1} \times\cdots\times b_{j,7} \subseteq \mathcal{D} , \tag{5.2}$$

where $b_{j,k} \subseteq a_k$ is determined by H_j, j=1,...,5. Thus, the overlapping of the H_j's in general will not disappear, but rather will be clarified, i.e., in general,

$$H_{j_1} \cap H_{j_2} \neq \emptyset \tag{5.3}$$

Clearly, in this case, if all statistical relations between the newly-introduced factor variables a_k 's and the variables S and Z are known, then the $p(H_j|Z_i\&D\&S)$ can be computed in (4.2). For example, if the a_k given the $Z_i\&D\&S$ are all mutually statistically independent, then

$$p(H_j|Z_i\&D\&S) = \prod_{k=1}^{7} p(a_k \in b_{j,k}|Z_i\&D\&S) , \tag{5.4}$$

and in general

$$\sum_{j=1}^{5} p(H_j|Z_i\&D\&S) > 1 , \tag{5.5}$$

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

One approach to redefining the problem here is to replace the, in general, overlapping H_j's and overlapping Z_i's by suitable partitioning of their domain spaces and then recompute the corresponding conditional probabilities in (4.2) involving the partitioning variables. For example, for convenience, denoting

$$I = \{1,\ldots,5\} , \tag{5.6}$$

for any subset $K \subseteq I$, or equivalently, $K \in P(I)$ (power class of I, the class of all subsets of I), define

$$H_{[K]} \overset{d}{=} \bigcap_{j\in K} H_j - \bigcup_{j\in I-K} H_j \subseteq \mathcal{D} , \tag{5.7}$$

$$H_{(K)} \overset{d}{=} \{H_j | j\in K\} \in P(dom(H)). \tag{5.8}$$

Thus for $K=\emptyset$,

$$H_{[\emptyset]} = H_{(\emptyset)} = \emptyset ; \tag{5.9}$$

for $K=\{j\}$, $j\in I$,

$$H_{(\{j\})} = \{H_j\} , \tag{5.10}$$

and for K=I,

$$H_{(I)} = dom(H) ; \quad H_{[I]} = \bigcap_{j\in I} H_j ; \tag{5.11}$$

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

$$H_{[K]} = H_1 \cap H_2 \cap H_4 - (H_3 \cup H_5) . \tag{5.12}$$

Clearly,

$$\mathcal{H} \overset{d}{=} \{H_{[K]} | K \subseteq I, H_{[K]} \neq \emptyset\} \tag{5.13}$$

is a disjoint exhaustive partitioning of $\mathcal{D}$. In a sense, $\mathcal{H}$ is the tightest disjoint exhaustive partitioning of $\mathcal{D}$ which generates back all H_j's through disjoint unions. Thus, $\mathcal{H}$ can serve as a sample space in place of initial dom(H); the H_j's are in general overlapping compound events of $\mathcal{H}$. Similar comments hold for Z.

Note that the mappings $H_{(\ )}: P(I) \to P(dom(H))$ and $H_{[\ ]}: P(I) \to P(\mathcal{D})$ 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 $H_{[K]} \neq \emptyset$

$$K \leftrightarrow H_{[K]} \leftrightarrow H_{(K)} . \tag{5.14}$$

For any $j\in I$, define the filter class of H_j , or one point coverage class of H_j, as

$$G_{(H_j)} \overset{d}{=} \{H_{(K)} | j\in K\subseteq I\}$$
$$\supseteq F_{(H_j)} \overset{d}{=} \{H_{(K)} | j\in K\subseteq I, H_{[K]} \neq \emptyset\} ; \tag{5.15}$$

define similarly,

$$F_{[H_j]} \overset{d}{=} \{H_{[K]} | j\in K\subseteq I, H_{[K]} \neq \emptyset\}. \tag{5.16}$$

Note also that the mappings $F_{(\ )}:\mathrm{dom}(H) \to PP(\mathrm{dom}(H))$ and $F_{[\ ]}:\mathrm{dom}(H) \to PP(D)$ are injective. Note, further, for any $j \in I$, the bijective relations

$$H_j = \dot{\bigcup}_{\{K|j \in K \subseteq I\}} = \dot{\cup} F_{[H_j]} \leftrightarrow F_{[H_j]} \leftrightarrow F_{(H_j)}. \qquad (5.17)$$

Now let $(\Omega, \mathcal{B}, p)$ be a probability space and $W:\Omega \to D$ be a random variable corresponding to $(z|Z_i \& D \& S)$. In turn, define random subset $S_H^{(1)}$ of $\mathrm{dom}(H)$, $S_H^{(1)}:\Omega \to P(\mathrm{dom}(H))$, where for any $\omega \in \Omega$,

$$S_H^{(1)} \overset{d}{=} \{H_j | j \in I, W(\omega) \in H_j\} . \qquad (5.18)$$

Then it follows that

$$\begin{aligned} W \in H_j \text{ iff } & \underset{\{K|j \in K \subseteq I, H_{[K]} \neq \emptyset\}}{\mathrm{or}} (W \in H_{[K]}) \\ \text{iff } & \underset{\{K|j \in K \subseteq I, H_{[K]} \neq \emptyset\}}{\mathrm{or}} (S_H^{(1)} = H_{(K)}) \\ \text{iff } & S_H^{(1)} \in F_{(H_j)} \\ \text{iff } & H_j \in S_H^{(1)}. \end{aligned} \qquad (5.19)$$

Hence

<u>Theorem</u> 2. (See: [19]; [17], pp.379-381.)

For all $j \in I$,

$$\begin{aligned} \mathrm{poss}(H_j) &\overset{d}{=} p(W \in H_j) = p(H_j \in S_H^{(1)}) \\ &= p(H_j | Z_i \& D \& S). \end{aligned} \qquad (5.20)$$

∎

The significance of this theorem will be more apparent below. Note also that unless dom(H) is a disjoint partitioning itself of D, (5.5) holds; but it always follows that

$$\sum_{K \subseteq I} p(W \in H_{[K]}) = 1 \qquad (5.21)$$

Again, similar results hold for dom(Z) replaced by a suitable space resulting from appropriately chosen factors.

On the other hand, often we do <u>not</u> know all the relevant factors or subvariables 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 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 $p(H_j|Z_i \& D \& S)$ and $p(Z_i|D \& S)$ 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.

<u>Theorem</u> 3. ([17], Chapter 5)

If $\mathrm{poss}_H:\mathrm{dom}(H) \to [0,1]$ is <u>any</u> function, perhaps representing the expert opinions of a panel, as human integrators of information, taking into account the complex and possible overlapping natures of the events in dom(H), then by letting U be any uniformly distributed random variable over [0,1] and defining the nested random subset $S_H^{(2)}$ of dom(H) by

$$\begin{aligned} S_H^{(2)} &\overset{d}{=} \mathrm{poss}_H^{-1}([U,1]) \\ &= \{H_j | j \in I,\ \mathrm{poss}_H(H_j) \geq U\}, \end{aligned} \qquad (5.22)$$

it follows that for all $j \in I$,

$$H_j \in S_H^{(2)} \quad \text{iff} \quad \mathrm{poss}_H(H_j) \geq U , \qquad (5.23)$$

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

$$\begin{aligned} \mathrm{poss}_H(H_j) &= p(H_j \in S_H^{(2)}) = p(S_H^{(2)} \in G_{(H_j)}) \\ &= \sum_{j \in K \subseteq I} p(S_H^{(2)} = H_{(K)}). \end{aligned} \qquad (5.24)$$

∎

<u>Remarks.</u>

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 - [17], Chapter 5) one point coverage equivalent to the given arbitrary possibility function over dom(H). (For comparisons of choices among such candidate random sets, see [20], where entropy is used as one criterion.) Each domain value H_j is naturally identifiable with the filter class $G_{(H_j)}$ containing all possible sets of H_i's having also H_j in them, i.e., all possible sets of interactions $H_{(K)}$, j in K. Thus it is not unreasonable that the given possibility value assigned to H_j can also be expressed rigorously as a probability involving the next higher order interaction domain P(dom(H)) above dom(H). Again, as before, all results hold for Z.

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 P(dom(H)), replacing 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 coverage 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. (See, e.g. [17], 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. (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 intranodal relation (1) in Table 1, sans the probability evaluation, and in natural language form:

$$Q \stackrel{d}{=} \text{"If D \& S, then H"}. \tag{6.1}$$

In symbolic form, where $\cdot$ represents &, v represents "or", $(\ )'$ represents "not", $\ni$ represents implication,

$$Q = (D \cdot S \ni H). \tag{6.2}$$

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

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

$$\alpha \ \& \ T_0 = \alpha, \tag{6.3}$$

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 existance of an absolute falsehood F_0 and let it play the role of an additive zero w.r.t. "or".

(b) "&" and "or" are commutative and associative with "&" 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 Figure 4 all satisfy these conditions.

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

$$\underset{Z_i \in \mathrm{dom}(Z)}{\text{or}} (Z_i) = T_0. \tag{6.4}$$

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

$$\underset{Z_i \in \mathrm{dom}(Z)}{\text{v}} (Z_i) = 1. \tag{6.5}$$

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

$$Q = \text{"If D \& S, then H or } \psi\text{"}, \tag{6.6}$$

where

$$\psi \stackrel{d}{=} \underset{Z_i \in \mathrm{dom}(Z)}{\text{or}} (Z_i \ \& \text{ not } Z_i). \tag{6.7}$$

In many formal languages, the Law of Excluded Middle holds so that for all propositions α,

$$\alpha \ \& \text{ not}(\alpha) = F_0. \tag{6.8}$$

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,

$$Q = (D \cdot S \ni (H \vee \psi)), \tag{6.9}$$

where

$$\psi = \underset{Z_i \in \mathrm{dom}(Z)}{\vee} (Z_i \ \& \ Z_i'). \tag{6.10}$$

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

$$Q = (D \cdot S \ni \underset{Z_i \in \mathrm{dom}(Z)}{\vee} (H \cdot Z_i \vee Z_i \cdot Z_i')). \tag{6.11}$$

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 $\alpha_1, \ldots, \alpha_m$, β,

$$(\beta \ni (\underset{i=1}{\overset{m}{\vee}} \alpha_i)) = \underset{i=1}{\overset{m}{\vee}} (\beta \ni \alpha_i). \tag{6.12}$$

(d) Implication chains relative to "&". Thus for any propositions α, β, γ,

$$(\gamma \ni (\alpha \cdot \beta \vee \beta \cdot \beta')) = (\gamma \ni \beta) \cdot (\gamma \cdot \beta \ni \alpha). \tag{6.13}$$

Again, it can be shown quite readily the first 3 ALDP examples in Figure 4 are such that their formal language components satisfy as well (c) and (d), when implication is interpreted as

$$\ni \ = \ \Rightarrow, \tag{6.14}$$

where for all α, β

$$(\beta \Rightarrow \alpha) \stackrel{d}{=} (\beta' \vee \alpha). \tag{6.15}$$

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

<u>Theorem 4</u>.

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

$$Q = \underset{Z_i \in \mathrm{dom}(Z)}{\vee} \mathcal{f}(Z_i;D,S;H), \tag{6.16}$$

where for all Z_i in dom(Z),

$$\begin{aligned} \mathcal{f}(Z_i;D,S;H) &\stackrel{d}{=} (D \cdot S \ni Z_i \cdot H) \\ &= g(Z_i;D,S) \cdot h(H;Z_i;D,S), \end{aligned} \tag{6.17}$$

where

$$g(Z_i;D,S) = (D \cdot S \ni Z_i) \tag{6.18}$$

can be interpreted as an <u>attribute variability</u> or error form and

$$h(H;Z_i;D,S) = (Z_i \cdot D \cdot S \ni H) \tag{6.19}$$

can be interpreted as an <u>inference rule</u> connecting Z_i and H.

∎

Thus, from the remarks preceeding Theorem 4, the formal language for Classical Logic and Probability Logic, boolean algebra, with implication given in (6.14),(6.15), 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 decomposition 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 <u>copula</u> ϕ_g as a mapping $\phi_g:[0,1]^n \to [0,1]$ which is the same as a cumulative probability distribution function over $[0,1]^n$ such that each marginal distribution, of one dimension corresponds to a random variable U_j uniformly distributed over [0,1], $i=1,\ldots,n$. (Copulas can be used to solve elegantly the important problem of determining all possible joint distributions given specified marginals. See [22].)

For purpose of simplicity here, define a co-copula ϕ_{or} as a mapping $\phi_{or}:[0,1]^n \to [0,1]$ which coincides with the disjunction probabilities corresponding to the conjunction ones for some given copula. Thus if U_i is any r.v. uniformly distributed over $[0,1]$, for $i=1,\ldots,n$, and $(U_1,\ldots,U_n)$ has some legitimate joint distribution, then $\phi_\&$ defined as follows will be a copula and ϕ_{or} defined below will be the co-copula corresponding to $\phi_\&$:

For any $c_i \in [0,1]$, $i \in I_o \overset{d}{=} \{1,\ldots,n\}$,

$$\phi_\&(c_1,\ldots,c_n) = p(\underset{i=1}{\overset{n}{\&}} (U_i \le c_i)), \qquad (6.20)$$

$$\phi_{or}(c_1,\ldots,c_n) = p(\bigvee_{i=1}^{n} (U_i \le c_i))$$

$$= \sum_{\emptyset \ne K \subseteq I_o} (-1)^{\mathrm{card}(K)+1} \cdot \phi_\&(c_{(K)}) , \qquad (6.21)$$

where analogous to previous notation

$$c_{(K)} \overset{d}{=} \{c_i | i \in K\}, \qquad (6.22)$$

by use of the modularity or Poincaré expansion property of probabilities. (For further properties of copulas and related functions, see e.g. [17], section 2.3.6.) Consider also the following related concepts:

Define a t-norm - also denoted as $\phi_\&$ - as a mapping $\phi_\&:[0,1]^n \to [0,1]$ which is associative, commutative, non-decreasing, continuous, and possessing boundary conditions

$$\phi_\&(1,x) = x \quad ; \quad \phi_\&(0,x) = 0 \quad , \qquad (6.23)$$

for all $0 \le x \le 1$, and such that

$$\phi_\& \ge \min. \qquad (6.24)$$

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

$$\phi_{or}(x_1,\ldots,x_n) = 1 - \phi_\&(1-x_1,\ldots,1-x_n), \qquad (6.25)$$

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

$$\phi_\&(x,x) < x \quad ; \qquad (6.26)$$

dually, define a t-conorm to be archimedean iff

$$\phi_{or}(x,x) > x \; , \qquad (6.27)$$

for all $0<x<1$.

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 (prod,probsum) are the only such functions which are both(copula,co-copula)and(t-norm,t-conorm)pairs simultaneously; further, the latter pair is also archimedean, where "prod" denotes ordinary arithmetic product, while "probsum" denotes formal probability "sum" (displaying modularity of probability) as the demorgan transform of prod. (See [23], Section 4.)

(prod,sum) is a non-demorgan 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 $(\phi_\&,\phi_{or})$, there always exists a corresponding continuous non-increasing function $h:[0,1] \to \mathbb{R}^+$, with $h(1) = 0$ and $\mathbb{R}^+$ denoting the extended real line including $+\infty$, such that, assuming the above pair is also demorgan,

$$\phi_\&(x_1,\ldots,x_n) = h^{-1}(\min(h(o),\sum_{i=1}^{n} h(x_i))) \; ; \qquad (6.28)$$

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

$$\alpha \overset{d}{=} (D \cdot S \ni Z) \;\; ; \; \alpha_i \overset{d}{=} (D \cdot S \ni Z_i) \; ; \qquad (6.29)$$

$$\beta \overset{d}{=} (Z \cdot D \cdot S \ni H) \;\; ; \; \beta_{ij} \overset{d}{=} (Z_i \cdot D \cdot S \ni H_j). \qquad (6.30)$$

Then make the following semantic evaluation of Q preserving the formal structure in Theorem 4:

$$\mathrm{poss}(Q = Q_j) = \mathrm{poss}(Q = (D \cdot S \ni H_j))$$

$$= \underset{i \in I}{\phi_{or}} (\phi_\&(\mathrm{poss}_\alpha(\alpha_i), \mathrm{poss}_\beta(\beta_{ij}))). \qquad (6.31)$$

In particular, the evaluation of Q using Zadeh's original fuzzy set theory or Fuzzy Logic is easily seen to be a special case of (6.31), when

$$\phi_\& = \min \quad , \; \phi_{or} = \max. \qquad (6.32)$$

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 $\phi_\&$ and ϕ_{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(#1 and #2):

$$Z = \begin{pmatrix} \text{geolocation parameters for \#1, for \#2} \\ \text{sensor system parameters for \#1for \#2} \\ \text{hull lengths for \#1, for \#2} \\ \text{classifications for \#1, for \#2} \\ \text{flag colors for \#1, for \#2} \end{pmatrix} . \qquad (6.33)$$

Also, for this example, H (denoted in [12] by C) represents correlation level between #1 and #2 (between 0 and 1 when evaluated), while D=1 is assumed and S represents observed (in error) counterpart of Z. Then the inference rules $\mathrm{poss}_\beta(\beta_{ij})$ correspond to some expert-derived (or derived by analytic or physical considerations) relation between some combination of degrees of matching attributes in general with possible correlation levels H ; the terms $\mathrm{poss}_\alpha(\alpha_i)$ represent error distributions between true and observed auxiliary attributes Z. 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 $\mathrm{poss}(Q=Q_j)$.

On the other hand, if we choose

$$\phi_\& = \mathrm{prod} \quad , \; \phi_{or} = \mathrm{sum} \quad . \qquad (6.34)$$

then (6.31) reduces to the classical probabilistic data fusion evaluation given in (4.2).

Next, consider the evaluation of data fusion as given in (6.31) when $\phi_{\&}$ is any copula and ϕ_{or} is the co-copula determined by $\phi_{\&}$ 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 (fixing D and S)

$$\mathrm{dom}(\alpha) = \{\alpha_i \mid i \in I\} \text{ '=' } \mathrm{dom}(Z) = \{Z_i \mid i \in I\}, \quad (6.35)$$

$$\mathrm{dom}(\beta) = \{\beta_{ij} \mid i \in I, j \in J\} \text{ '=' } \mathrm{dom}(Z) \times \mathrm{dom}(H)$$
$$= \{(Z_i, H_j) \mid i \in I, j \in J\}, \quad (6.36)$$

where I and J are suitably chosen index sets.

Let

$$\underset{\sim}{U} \overset{d}{=} (U_i, U_{ij})_{\substack{i \in I,\\ j \in J}} \quad (6.37)$$

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

$$\begin{aligned} \alpha_i \in S_\alpha \quad &\text{iff } U_i \le \mathrm{poss}_\alpha(\alpha_i) \\ \alpha_i \notin S_\alpha \quad &\text{iff } U_i > \mathrm{poss}_\alpha(\alpha_i) \end{aligned} \quad (6.38)$$

and

$$\begin{aligned} \beta_{ij} \in S_\beta \quad &\text{iff } U_{ij} \le \mathrm{poss}_\beta(\beta_{ij}) \\ \beta_{ij} \notin S_\beta \quad &\text{iff } U_{ij} > \mathrm{poss}_\beta(\beta_{ij}) \,. \end{aligned} \quad (6.39)$$

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

$$S_\alpha = S_\alpha^{(2)} \quad , \quad S_\beta = S_\beta^{(2)} \quad (6.39)$$

as given in Theorem 3. Determine $\phi_{\&}, \phi_{or}$ through $\underset{\sim}{U}$.

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

$$\mathrm{poss}(Q = Q_j) = \sum_{\emptyset \ne K \subseteq I} (-1)^{\mathrm{card}(K)+1} \cdot M_{K,j} \quad (6.40)$$

where for all subsets K

$$M_{K,j} \overset{d}{=} \underset{i \in K}{\phi_{\&}}(\phi_{\&}(p(U_i \le \mathrm{poss}_\alpha(\alpha_i)), p(U_{ij} \le \mathrm{poss}_\beta(\beta_{ij}))))$$

$$= p(\underset{i \in K}{\phi_{\&}}(U_i \le \mathrm{poss}_\alpha(\alpha_i),\ U_{ij} \le \mathrm{poss}_\beta(\beta_{ij})))$$

$$= p(\ \underset{i \in K}{\&}((\alpha_i \in S_\alpha)\ \&\ (\beta_{ij} \in S_\beta))\ ). \quad (6.41)$$

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

$$\mathrm{poss}(Q = Q_j) = p(\ \underset{i \in I}{\mathrm{or}}\ ((\alpha_i \in S_\alpha) \& (\beta_{ij} \in S_\beta))\ )$$

$$= p(\ A_j \cap (S_\alpha \times S_\beta) \ne \emptyset), \quad (6.42)$$

where

$$A_j \overset{d}{=} \{(\alpha_i, \beta_{ij}) \mid i \in I\} = \{(Z_i, Z_i, H_j) \mid i \in I\}. \quad (6.43)$$

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 (6.43), we obtain the following result:

<u>Theorem 5.</u>

Given variables C,S,H and auxiliary variable Z as before, then under the assumptions leading to eq. (6.31) and assuming the constructions in (6.35)-(6.39), it follows that for all $j \in J$,

$$\mathrm{poss}(Q = Q_j) = \mathrm{poss}(\ A_j \cap (S_\alpha \times S_\beta) \ne \emptyset\ )$$
$$= p(\ A_j \cap (S_\alpha \times S_\beta) \ne \emptyset\ )$$
$$= \mathrm{plaus}_{S_\alpha \times S_\beta}(A_j)\ , \quad (6.44)$$

where $\mathrm{plaus}_{S_\alpha \times S_\beta}$ denotes the <u>plausibility</u> or upper probability measure with respect to random subset $S_\alpha \times S_\beta$ of $\mathrm{dom}(\alpha) \times \mathrm{dom}(\beta)$.

∎

<u>Remarks.</u>

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 <u>full</u> specification of the associated random (sub)sets. 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 Theorem 5, where an equivalent plausibility description is given, it is only specified over the A_j'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 superset coverages of a random set; any doubt measure is determined by the subset coverge of a random set.

In any case, Theorem 5 shows that a homomorphic 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), $\underset{\sim}{U}$, 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_α and S_β are not only statistically independent, but are the maximal entropy one point equivalent representatives for poss_α and poss_β, 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.l.o.g. Z' is the vector of probabilistic attributes and Z" is the vector of non-probabilistic ones. Note that by the canonical representation theorem mentioned in Section 6 (see eq.(6.28)), if an archimedean t-norm, t-conorm pair is chosen for the evaluation in (6.31), then poss(Q) becomes a monotone transform τ_h, say, for generator function h of $\phi_{\&}$, of a sum of terms over $i \in I$, where

$$\tau_h(x) \overset{d}{=} 1 - h^{-1}(\min(h(0), x)), \quad (6.46)$$

for all $x \in R^+$, and the i_1^{th} term, $i = (i_1, i_2)$, is

$$h(1-\phi_{\&}(poss_{\alpha'}(Z'_{i_1}, G(Z'_{i_1}, H_j)))), \qquad (6.47)$$

where α is partitioned as Z into (α', α'') and

$$G(Z_{i_1}, H_j) \stackrel{d}{=} \phi_{or}_{(Z''_{i_2} \subset dom(\alpha''))}(\phi_{\&}(poss_{\alpha''}(Z''_{i_2}), poss_{\beta}(Z'_{i_1}, Z''_{i_2}, H_j)))). \qquad (6.48)$$

Note that dom(Z') is finite, as well as all other domains of relevant variables, in order for finite argument functions $\phi_{\&}$ and ϕ_{or} 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 $\lim_{dom(Z') \to R^m} (poss(Q))$ to mean that the limit of poss(Q) will be taken, if it exists, as dom(Z') and $poss_{\alpha'}$ are refined so that all cell sizes approach point limits and thus $poss_{\alpha'}$ approaches a joint p.d.f. form corresponding to random variable (Z'|D&S). Then we can show the following:

<u>Theorem</u> 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-conorm pair $\phi_{\&}, \phi_{or}$ chosen for the data fusion evaluation (6.41).

Then

$$\lim_{dom(Z') \to R^m} (poss(Q=Q_j)) = \tau_h(\nu_h \cdot E_{Z'}(\kappa(G(Z', H_j)))), \qquad (6.49)$$

where

$$\nu_h \stackrel{d}{=} (-d\, h(x)/dx)_{x=1}, \qquad (6.50)$$

and all $0 \le \kappa \le 1$,

$$\kappa(x) \stackrel{d}{=} (\partial\phi_{\&}(x,y)/\partial y)_{y=0}, \qquad (6.51)$$

and where $E_{Z'}$ denotes ordinary statistical expectation w.r.t. r.v. Z', conditioned on D&S throughout, where Z' corresponds to the limiting p.d.f. for $poss_{\alpha'}$. ∎

Thus, up to essentially monotone transforms, the limiting form of the data fusion computations here is an averaged value of the data fusion with (only) fixed domain attributes Z". Further simplification to the classical integral (and continuous) version of (4.2) occurs when the fixed non-probabilisitc attribute components are 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 $k(H;Z_i;D,S)$ linking either observed data, such as D,S 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 $g(Z_i;D,S)$.

The somewhat similar, but more general structure for such systems is given in eq. (7.1).

$$Q_j \stackrel{d}{=} \bigvee_{Z_i \in dom(Z)} \left(\bigwedge_{k=1}^{m} (\underbrace{j_k(Z_i, H_j; D,S)}_{g_{j_{kij}}} \ni \underbrace{k_k(Z_i, H_j; D,S)}_{g_{k_{kij}}}) \right) \qquad (7.1)$$

representing $(D \cdot S \ni H)$, where for all k, j_k and k_k are, possibly expert-derived, boolean functions, i.e., combinations of operations $\cdot$, $\vee$, ( )'.

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.

<u>Example</u> 1. ALDP 1.

ALDP 1 = (boolean algebra Ω with (6.14) valid for $\ni$, Classical (two-valued) Logic)

The calculus of relations for implications for the formal language part here, Ω boolean with (6.14):

For all $\alpha_i, \beta_i, \alpha_o, \beta_o \in \Omega$, $i=1,\ldots,m$, $m=1,2,\ldots$,

$$\bigvee_{i=1}^{m} (\beta_i \ni \alpha_i) = ((\bigwedge_{i=1}^{m} \beta_i) \ni (\bigvee_{i=1}^{m} \alpha_i)), \qquad (7.2)$$

$$\bigwedge_{i=1}^{m} (\beta_i \ni \alpha_i) = ((\bigvee_{i=1}^{m} \alpha_i' \cdot \beta_i \vee \bigwedge_{i=1}^{m} \beta_i) \ni (\bigwedge_{i=1}^{m} \alpha_i)). \qquad (7.3)$$

Thus, if $\beta_1 = \cdots = \beta_m = \beta_o$, then (7.2) and (7.3) become homomorphic relations for fixed antecedents:

$$\bigvee_{i=1}^{m} (\beta_o \ni \alpha_i) = (\beta_o \ni (\bigvee_{i=1}^{m} \alpha_i)), \qquad (7.4)$$

$$\bigwedge_{i=1}^{m} (\beta_o \ni \alpha_i) = (\beta_o \ni (\bigwedge_{i=1}^{m} \alpha_i)). \qquad (7.5)$$

But negation is in general <u>not</u> a homomorphic relation:

$$(\beta_o \ni \alpha_o)' = \alpha_o' \cdot \beta_o \ne (\beta_o \ni (\alpha_o' \cdot \beta_o)). \qquad (7.6)$$

Also, for all $\alpha_o, \beta_o, \gamma_o \in \Omega$, (7.7)

$$(1 \ni \alpha_o) = \alpha_o; (\beta_o \ni \alpha_o) = (\beta_o \ni \alpha_o \cdot \beta_o); (\gamma_o \ni (\alpha_o \ni \beta_o)) = (\gamma_o \ni \beta_o)(\beta_o \gamma_o \ni \alpha_o).$$

Consider now the semantic evaluation part. Denoting the evaluation of any proposition variable α, having domain of possible (or not) values in $\Omega(dom(\alpha)) \subseteq \Omega$) as function $poss_{\alpha}: dom(\alpha) \to [0,1]$, for any $\alpha_i \in dom(\alpha)$

$$poss_{\alpha}(\alpha_i) = 0, \text{ i.e., } \alpha_i \notin \alpha$$

or (7.8)

$$poss_{\alpha}(\alpha_i) = 1, \text{ i.e., } \alpha_i \in \alpha,$$

and variable α can be identified with a subset of Ω:

$$\alpha = \{\alpha_i | \alpha_i \in dom(\alpha_i) \,\&\, poss_{\alpha}(\alpha_i) = 1\}, \qquad (7.9)$$

with $poss_{\alpha}$ playing the role of an ordinary set membership function. Then, Classical Logic, as a truth-functional logic (see, e.g. [27] for further elaboration) has the following homomorphic forms, for all proposition variables (and similarly for all propositions) α, β:

$$poss_{\alpha \vee \beta} = \max(poss_{\alpha}, poss_{\beta}), \qquad (7.10)$$

$$poss_{\alpha \cdot \beta} = \min(poss_{\alpha}, poss_{\beta}), \qquad (7.11)$$

$$poss_{\alpha'} = 1 - poss_{\alpha}, \qquad (7.12)$$

$$\text{poss}_0 = 0 \ , \ \text{poss}_1 = 1 \ , \tag{7.13}$$

and hence

$$\text{poss}_{\beta \ni \alpha} = \max(1-\text{poss}_\beta \ , \ \text{poss}_\alpha), \tag{7.14}$$

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 ALDP 2.

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

$$\text{poss}(Q=Q_j) = \text{poss}_{D \cdot S \ni H}(H_j)$$

$$= \max_{Z_i \in \text{dom}(Z)} \Big(\min_{(k=1,\ldots,m)} (\max(1-\hat{j}_{kij} \ , \ \hat{k}_{kij})) \Big), \tag{7.15}$$

where for all k,i,j

$$\hat{j}_{kij} \overset{d}{=} \text{poss}_{j_k}(Z_i, H_j; D, S) \ , \tag{7.16}$$

$$\hat{k}_{kij} \overset{d}{=} \text{poss}_{k_k}(Z_i, H_j; D, S), \tag{7.17}$$

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.(7.15) reduces to:

$\text{poss}(Q=Q_j) = 1$ <u>iff</u> there is some attribute value Z_i such that for each k, k=1,...,m, $(j_k \ni k_k)$ when evaluated at Z_i, H_j, D, S, is true, i.e., $\text{poss}_{(j_k \ni k_k)}(Z_i, H_j, D, S) = 1$, or equivalently, Z_i, H_j, D, S all fire inference rule $(j_k \ni k_k)$: either j_k is false at this evaluation (vacuous antecedent being satisfied) or more non-trivially, k_k is true for this evaluation ; (7.18)

$\text{poss}(Q=Q_j) = 0$ <u>iff</u> no such attribute value Z_i as above exists. (7.19)

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

$$\text{poss}(Q=Q_j) = \text{poss}(q(H_j;D,S) \ni \lambda(H_j;D,S))$$

$$= \max(1-\text{poss}(q(H_j;D,S)), \text{poss}(\lambda(H_j;D,S))) \tag{7.20}$$

where

$$q(H_j;D,S) \overset{d}{=} \mathop{\cdot}_{Z_i \in \text{dom}(Z)} \Big(\bigvee_{k=1}^{m} (k'_{kij} \cdot j_{kij}) \vee \mathop{\cdot}_{k=1}^{m} j_{kij} \Big) \tag{7.21}$$

and

$$\lambda(H_j;D,S) \overset{d}{=} \bigvee_{Z_i \in \text{dom}(Z)} \Big(\mathop{\cdot}_{k=1}^{m} k_{kij} \Big), \tag{7.22}$$

where, in turn, (7.10)-(7.14) could be used to evaluate further poss(q) and poss(λ), which of course should lead back to (7.15) and thus (7.18), (7.19), as a check.

The philosophy of approach in this example is that for the modeling of data fusion, in the context of 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.

<u>Example</u> 2. ALDP 2.

ALDP 2 = (modified boolean algebra Ω_0 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 implications for the formal language part here, Ω_0, is the same formally as that for Ω as in Example 1, except for the following slight modifications given in the two statements (I), (II) below:

(I) The middle equation in (7.7) will be valid, provided that $\alpha_0 \leq \beta_0$, i.e., $\alpha_0 = \alpha_0 \cdot \beta_0$, otherwise in general it is not true.

(II) Adjoin the term $\vee \ \beta_0 \cdot \beta'_0$ to the consequent of $\ni$ on the left hand side of the equality for the far right chaining equation in (7.7).

Then the semantic evaluations proceed in formally the same way as for ALDP 1, 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 (7.8). 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 ALDP 2. On the other hand, eqs. (7.20)-(7.22) hold here, with appropriate modifications following those in (I), (II) above.

<u>Example</u> 3. ALDP 3.

ALDP 3 = (boolean algebra Ω with (6.14), Probability Logic).

Since Ω 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], especially Chapter 2, Sections 26 and 27 for background), we have the usual basic (finitely additive) probability properties, for a given probability measure $p:\Omega \to [0,1]$, playing the role of the semantic evaluation poss in the two previous examples. (In order to use the more standard notation, p is used in place of poss) Only for purposes of comparisons the following well-known properties are given:

For all propositions $\alpha_0, \beta_0 \in \Omega$,

$$p(\alpha_0 \vee \beta_0) = p(\alpha_0) + p(\beta_0) - p(\alpha_0 \cdot \beta_0), \tag{7.23}$$

the modularity property, extending to the Poincaré expansion, used previously in this paper. Here for all $\alpha_1, \ldots, \alpha_n \in \Omega$, letting $I_n = \{1,\ldots,n\}$, n=1,2,... ,

$$p\Big(\bigvee_{i=1}^{n} \alpha_i\Big) = \sum_{\emptyset \neq K \subseteq I_n} (-1)^{\text{card}(K)+1} \cdot p\Big(\mathop{\&}_{i \in K} \alpha_i\Big) \ , \tag{7.24}$$

$$p(\alpha'_0) = 1 - p(\alpha_0) \ , \tag{7.25}$$

$$p(0) = 0 \ , \ \ p(1) = 1 \ , \tag{7.26}$$

resulting in the following evaluations for implication (by (6.14), for $\ni$) and some less-known inequalities involving conditional probabilities:

$$p(\beta_0 \ni \alpha_0) = p(\beta_0' \vee \alpha_0) = 1-p((\beta_0' \vee \alpha_0)') = 1-p(\beta_0 \cdot \alpha_0')$$
$$= p(\alpha_0|\beta_0) + p(\alpha_0'|\beta_0) - p(\beta_0 \cdot \alpha_0')$$
$$= p(\alpha_0|\beta_0) + p(\alpha_0'|\beta_0) - p(\alpha_0'|\beta_0) \cdot p(\beta_0)$$
$$= p(\alpha_0|\beta_0) + p(\alpha_0'|\beta_0) \cdot p(\beta_0')$$
$$\geq p(\alpha_0|\beta_0) \qquad (7.27)$$
$$\geq p(\alpha_0 \cdot \beta_0) \,, \qquad (7.28)$$

where the conditional probability is defined as usual as, e.g.,

$$p(\alpha_0|\beta_0) \overset{d}{=} p(\alpha_0 \cdot \beta_0)/p(\beta_0), \qquad (7.29)$$

provided $p(\beta_0) > 0$.

The above inequalities are strict, in general, and show that, basically, we cannot identify implication, as defined in the formal language (Ω) via eq.(6.14), with a "conditional object" such as $(\alpha_0|\beta_0)$, otherwise this would, following evaluations by p and making the natural identification

$$p((\alpha_0|\beta_0)) = p(\alpha_0|\beta_0) \,, \qquad (7.30)$$

contradict the inequality in (7.27). Hence the behavior of conditional probabilities, while roughly resembling that of the probability of implications is not the same - indeed, one can, by choosing judiciously β_0 close to o in some natural sense, make $p(\beta_0 \ni \alpha_0)$ approach unity, while for the same choice of α_0, β_0, $p(\alpha_0|\beta_0)$ approaches zero. The significance of these results will be explored further in the next section, where we develop an ALDP (4) where formal implications $\alpha_0 \ni \beta_0$ can be identified with "conditional objects" $(\alpha_0|\beta_0)$, whose semantic evaluations as in (7.30) are conditional probabilities; but in light of the above remarks, necessarily these entities lie outside of the original space of propositions Ω.

Returning to the data fusion form in (7.1), the semantic evaluation for Probability Logic becomes, using first (7.24) and then (7.5),

$$p(Q=Q_j) = p(D \cdot S \ni H_j)$$
$$= \sum_{\emptyset \neq K \subseteq dom(Z)} (-1)^{card(K)+1} \cdot p(q_j^{(K)} \ni \lambda_j^{(K)}), \qquad (7.31)$$

which can be further evaluated through use of (7.27) (equality part) in conjunction with (7.23)-(7.26), where similar to (7.21), (7.22), but differing in the operations involving Z_i,

$$q_j^{(K)} \overset{d}{=} \bigvee_{\substack{Z_i \in K \\ k \in I_m}} k_{kij}' \cdot j_{kij} \;\; \bigvee_{\substack{Z_i \in K \\ k \in I_m}} \cdot j_{kij} \qquad (7.32)$$

and

$$\lambda_j^{(K)} \overset{d}{=} \underset{\substack{Z_i \in K \\ k \in I_m}}{\cdot} k_{kij} \,. \qquad (7.33)$$

Alternatively, by using both (7.4) and (7.5) from the calculus of inference relations, and then applying p, one obtains the same as (7.20), with "poss" replaced by "p". Thus,

$$p(Q=Q_j) = p(q(H_j;D,S) \ni \lambda(H_j;D,S)), \qquad (7.34)$$

which can be evaluated through the equality part of (7.27) or through the expansion

$$p(\beta_0 \ni \alpha_0) = p(\beta_0') + p(\alpha_0) - p(\beta_0' \cdot \alpha_0)$$
$$= p(\beta_0') + p(\alpha_0 \cdot \beta_0), \qquad (7.35)$$

for all $\alpha_0, \beta_0 \in \Omega$, 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 $(j_{kij} \ni k_{kij})$.

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 $\ni$ was interpreted in the formal language components as $\rightarrow$ 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 $(j_{kij} \ni k_{kij})$ is $p(k_{kij}|j_{kij})$ for some choice of probability measure p over Ω, 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 Thesis [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 $\rightarrow$ 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 v, &, but not ( )', conditional probabilities are compatible with homomorphic relations holding for all three operations, for any fixed antecedent, i.e., obviously, for all $\alpha_0, \beta_0, \gamma_0 \in \Omega$,

$$p((\alpha_0|\gamma_0)') = 1 - p(\alpha_0|\gamma_0) = p(\alpha_0'|\gamma_0), \quad (8.1)$$

$$p((\alpha_0|\gamma_0) \vee (\beta_0|\gamma_0)) = p(\alpha_0 \vee \beta_0|\gamma_0), \quad (8.2)$$

$$p((\alpha_0|\gamma_0)\cdot(\beta_0|\gamma_0)) = p(\alpha_0\cdot\beta_0|\gamma_0). \quad (8.3)$$

Recall also the operation + over Ω, which in terms of v, •, ( )' is, for any $\alpha_0, \beta_0 \in \Omega$

$$\alpha_0 + \beta_0 = \alpha_0\cdot\beta_0' \vee \alpha_0'\cdot\beta_0, \quad (8.4)$$

and conversely,

$$\alpha_0 \vee \beta_0 = \alpha_0 + \beta_0 + \alpha_0\cdot\beta_0 \quad (8.5)$$

$$\alpha_0' = \alpha_0 + 1. \quad (8.6)$$

Thus there is a bijective relationship between $(\Omega, \vee, \cdot, (\ )')$, a boolean algebra and $(\Omega, +, \cdot)$, 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 actual subsets of a fixed universal set say X where the correspondences hold:

$$\left\{\begin{array}{l} 1 \leftrightarrow X \ ;\ + \leftrightarrow \Delta \text{ (symmetric set difference)};\\ 0 \leftrightarrow \emptyset \ ;\ \vee \leftrightarrow \cup \text{ (set union)};\\ \cdot \leftrightarrow \cap \text{ (set intersection)};\\ (\ )' \leftrightarrow C \text{ or } X-(\ ) \text{ (set complemenet)};\\ \leq \text{ (partial order over } \Omega) \leftrightarrow \subseteq \text{ (subset relation)} \end{array}\right. \quad (8.7)$$

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

Noting that also, for any $\alpha_0, \beta_0 \in \Omega$,

$$p(\alpha_0|\beta_0) = p(\alpha_0\cdot\beta_0|\beta_0), \quad (8.8)$$

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

Theorem 7. Characterization of conditional objects [34]

Given boolean ring Ω, there is a unique space $\tilde{\Omega}$ of smallest possible classes - according to subset partial ordering-denoted as the conditional objects $(\alpha_0|\gamma_0)$, $(\beta_0|\gamma_0)$, $(\beta_0|\zeta_0)$, ..., for all $\alpha_0, \beta_0, \gamma_0, \zeta_0$, .. $\in \Omega$, such that the measure-free counterparts of (8.1)-(8.3) and (8.8) hold. That is,

$$(\alpha_0|\gamma_0)' = (\alpha_0'|\gamma_0), \quad (8.9)$$

$$(\alpha_0|\gamma_0) \vee (\beta_0|\gamma_0) = (\alpha_0 \vee \beta_0|\gamma_0), \quad (8.10)$$

$$(\alpha_0|\gamma_0)\cdot(\beta_0|\gamma_0) = (\alpha_0\cdot\beta_0|\gamma_0), \quad (8.11)$$

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

$$(\alpha_0|\gamma_0) + (\beta_0|\gamma_0) = (\alpha_0 + \beta_0|\gamma_0) \quad (8.12)$$

to hold; and

$$(\alpha_0|\gamma_0) = (\alpha_0\cdot\gamma_0|\gamma_0). \quad (8.13)$$

Specifically, such conditional objects constitute all possible principal ideal cosets of ring Ω, where for any $\alpha_0, \gamma_0 \in \Omega$,

$$\begin{aligned}(\alpha_0|\gamma_0) &= \Omega\cdot\gamma_0' + \alpha_0\\ &= \Omega\cdot\gamma_0' + \alpha_0\cdot\gamma_0 = \Omega\cdot\gamma_0' \vee \alpha_0\cdot\gamma_0\\ &= \{x\cdot\gamma_0' + \alpha_0\cdot\gamma_0 | x \in \Omega\} \subseteq \Omega, \end{aligned} \quad (8.14)$$

the principal ideal coset generated by γ_0' with residue α_0.

Proof: Use first the basic homomorphism theorem for quotient rings and the equivalence class property of cosets applied to (8.13). Again, see [34].

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 ?

Previous work in this direction includes: Hailperin [37], who extended some of Boole's original ideas and developed essentially the same entities as produced 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 "qualitiative 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 $\alpha_0, \beta_0, \gamma_0, \delta_0 \in \Omega$ arbitrary. The natural class extension of • applied now to $(\alpha_0|\beta_0) \cdot (\gamma_0|\delta_0)$, noting each conditional object is in reality via (8.14) a subset of Ω, yields:

$$\begin{aligned}(\alpha_0|\beta_0)\cdot(\gamma_0|\delta_0) &= \{q\cdot r | q \in (\alpha_0|\beta_0), r \in (\gamma_0|\delta_0)\}\\ &= \{(x\cdot\beta_0' + \alpha_0)\cdot(y\cdot\delta_0' + \gamma_0) | x,y \in \Omega\}\\ &\subseteq \Omega. \end{aligned} \quad (8.15)$$

The basic structure of the conditional object extension $\tilde{\Omega}$ of Ω is summarized next.

Theorem 8. Basic structure of $\tilde{\Omega}$ [34],[35],[36].

(i) In terms of quotient rings,

$$\tilde{\Omega} = \bigcup_{\gamma_0 \in \Omega} (\Omega/\Omega\cdot\gamma_0') = \bigcup_{\gamma_0 \in \Omega} (\Omega/\Omega\cdot\gamma_0). \quad (8.16)$$

(ii) Conditioning as defined here can be identified essentially as the functional inverse of one-sided conjunction, i.e., conditional objects $(\alpha_0|\gamma_0)$ all satisfy the basic relation analogous to (7.29) for conditional probabilities and a related condition:

$$(\alpha_0|\gamma_0)\cdot\gamma_0 = \alpha_0\cdot\gamma_0 \quad (8.17)$$

and

$$(\alpha_0|\gamma_0) = \{x|x \in \Omega,\ x\cdot\gamma_0 = \alpha_0\cdot\gamma_0\}. \quad (8.18)$$

(iii) The natural class extensions of all boolean operations from Ω to $\tilde{\Omega}$ are well-defined/closed with ring-like properties, i.e., in the same previous sense, $\tilde{\Omega}$ is a modified boolean algebra.

(iv) $\Omega \subseteq \tilde{\Omega}$,

since for all $\alpha_0 \in \Omega$, (8.14) shows immediately that

$$(\alpha_0|1) = \{\alpha_0\} . \quad (8.19)$$

(v) Also, partial order $\leq$ defined over Ω, characterized by , for any α_0 , $\beta_0 \in \Omega$,

$$\alpha_0 \leq \beta_0 \text{ iff } \alpha_0 = \alpha_0\cdot\beta_0 \text{ iff } \beta_0 = \beta_0 \vee \alpha_0 \ ,(8.20)$$

can be extended directly to $\tilde{\Omega}$ with the same characterizations as in (8.20) where (unconditional) objects in Ω are replaced by conditional ones in $\tilde{\Omega}$. Then, combining this with (iii) and (iv) establishes $(\tilde{\Omega},\vee,\cdot,(\ )',+;\leq)$ as a natural extension of its unconditional counterpart $(\Omega,\vee,\cdot,(\ )',+;\leq)$.

(vi) A basic calulus of operations is , in addition to the properties in (8.9)-(8.13) for any $\alpha_i,\gamma_i \in \Omega$, $i=1,\ldots,m$, $m\geq 1$,

$$\bigvee_{i=1}^{m} (\alpha_i|\gamma_i) = (\bigvee_{i=1}^{m} \alpha_i | \bigvee_{i=1}^{m} \alpha_i\cdot\gamma_i \vee \prod_{i=1}^{m} \gamma_i),(8.21)$$

$$\prod_{i=1}^{m} (\alpha_i|\gamma_i) = (\prod_{i=1}^{m} \alpha_i | \bigvee_{i=1}^{m} \alpha_i'\cdot\gamma_i \vee \prod_{i=1}^{m} \gamma_i),(8.22)$$

$$\sum_{i=1}^{m} (\alpha_i|\gamma_i) = (\sum_{i=1}^{m} \alpha_i | \prod_{i=1}^{m} \gamma_i) \ . \quad (8.23)$$

Noting the reductions of (8.21)-(8.23) when antecedent $\gamma_1=\cdots=\gamma_m=\gamma_0$, as in (8.9)-(8.12), it follows that all boolean operational extensions over Ω coincide with corresponding coset operations when restricted to a fixed quotient ring, here $\Omega/\Omega\cdot\gamma_0'$.

(vii) As a special case of (8.22), the following chaining condition holds for all $\alpha_0,\beta_0,\gamma_0 \in \Omega$:

$$(\alpha_0\cdot\beta_0|\gamma_0) = (\beta_0|\gamma_0)\cdot(\alpha_0|\beta_0\cdot\gamma_0). \quad (8.24)$$

Proof: The most difficult proof is that of (8.22). A sketch of the proof for the case m=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. • over v) 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 Von 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 to force 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 $p:\Omega \to [0,1]$ to $p:\tilde{\Omega} \to [0,1]$ through eq.(7.30) yields for the extension a monotone function. (Again, see [34]-[36], for further details.)

Most importantly here, analogues of calculus of relations for ALDP 1 (eqs.(7.2)-(7.7)) 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 Ω as simply

$$\text{ALDP } 4 = (\tilde{\Omega},p), \quad (8.25)$$

where $p:\tilde{\Omega} \to [0,1]$ is the conditional probability extension of $p:\Omega \to [0,1]$, as mentioned above and where implication is interpreted as conditioning, i.e., for all $\alpha_0,\beta_0 \in \Omega$,

$$(\beta_0 \Rightarrow \alpha_0) = (\alpha_0|\beta_0). \quad (8.26)$$

(Note that implication or conditioning here is restricted to be upon unconditional elements, i.e. elements of Ω, 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 expression Q in (7.1):

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

$$p(Q=Q_j) = p(\bigvee_{Z_i\in \text{dom}(Z)} (\prod_{k=1}^{m} (k_{kij}|j_{kij})))$$

$$= p(\delta(H_j;D,S)|\delta(H_j;D,S)\vee q(H_j;D,S))$$

$$= p(\delta(H_j;D,S))/p(\delta(H_j;D,S)\vee q(H_j;D,S)), \quad (8.27)$$

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

$$\delta(H_j;D,S) \overset{d}{=} \bigvee_{Z_i\in \text{dom}(Z)} (\prod_{k=1}^{m} (k_{kij}\cdot j_{kij})). \quad (8.28)$$

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^3 systems. In Section 2, qualitative relations are established pinpointing the role of data fusion in C^3 systems- especially as perceived by the author in previous work (see Figures 1,2,3), where data fusion is a process within a C^3 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 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 where in all, implication is interpreted as a disjunction of a negation and affirmation. A particular quantitative counterpart of the qualitative model given in the previous section is presented in Section 3. In this model, the collection of all updated marginal node state distributions (in either the classic probability sense or in a multivalued logic sense of broader scope) is shown to depend functionally on essentially ten types of primitive relations (in the probability interpretation, they become conditional probabilities) among the basic variables determining the C^3 system in question. These variables include: S, "signal" nodes N receive; R, response of nodes; D, detection state; H, hypotheses selection; and F, algorithm choice (Theorem 1). In turn, this result is used to establish a zero-sum two person C^3 decision game between adversary and friendly C^3 systems. There, each game move corresponds to a choice of the ten types of primitive relations, up to feasible and compatible conditions, and the resulting loss due to a joint move by both players is some figure-of-merit based upon moe's and mop's, which are in turn evaluated through the node state distributions as a consequence of the primitive relations' forms (Figure 5).

In Section 4, the quantitative expression for data fusion $p(H|D,S)$ (eq.(4.1)) is considered for the classical probability case. An auxiliary variable Z is introduced for the evaluation, representing possible characteristics or attributes which can be used to connect D and S with H through probabilistic conditioning here. This results in the well-known weighted sum of conditional probabilities form (eq.(4.2)). In Section 5, two modifications of the classical probability case are considered. First treated is the situation where variables Z or H in actuality are not random variables due to their "sample spaces of elementary events or domain values not representing truly disjoint (and exhaustive) events, but where the relevant subfactors contributing to these - in actuality, compound - events can be determined at least in a full probabilistic sense. This results, in effect, in random set descriptions replacing the original "distributions" for the variables (Theorem 2). Next, the case where not all subfactors are known is considered. In this situation, if experts are available, possibility functions can be gleaned for the overlapping or vague events, which, in effect, take into account the possible joint occurrences, and thus can yield functions which exceed unity in summation. However, it is shown in Theorem 3, quite similar in form to Theorem 2, that this is always equivalent to the partial specification (through one point coverages) of a random set model, thereby giving rigorous justification for this procedure. The results in Section 5 are further extended in Section 6, where the formal language aspect for data fusion is emphasized (Theorem 4). This result (extending (4.2)) shows data fusion can, under relatively mild assumptions, be expressed as a disjunction of conjunctions of inference rules and variability or error forms connecting D,S, and Z with H. In turn, a general semantic evaluation for data fusion is presented through t-norms, copulas, etc (See (6.31)) This evaluation form generalizes the PACT algorithm which seeks to determine correlation level between track histories through disparate data sources, including possible linguistic-based information [12]. A relation is given in Theorem 5 connecting the above-mentioned general data fusion distribution with random sets and Dempster-Shafer plausibility functions.

In Section 7, the most general formal setting is established and analyzed for describing data fusion. Basically here, data fusion is considered a disjunction of conjunctions of inference rules with antecedents and consequences in general functional forms involving possibly all four relevant variables D,S,Z,H (see eq.(7.1)), essentially the same structure as a general knowledge-based system, such as used in medical diagnosis or parameter estimation. A calculus 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 inference rules through conditional probabilities. For consistency, this requires the full development of a calculus of "conditional objects" (Theorems 7,8). It is shown that this ALDP can be succesfully used to evaluate data fusion probabilities with a level of complexity of 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 NI (Natural Intelligence).

Alternative structures for data fusion may also be investigated - as opposed e.g., to that given here in (6.16) or (7.?) 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 iterated 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 (IR) and the Joint Directors of Laboratories, Technical Panel for C^3, Basic Research Group (JDL, TPC^3, BRG).

11. REFERENCES

1. M. Kaku and J. Trainer, *Beyond Einstein*, Bantom Books, New York, March, 1987.

2. F. White, *Data Fusion Lexicon*, prepared for Joint Directors of Laboratories, Technical Panel for C^3 (JDL, TPC^3), Data Fusion Sub-Panel (DFSP), Draft Version, May, 1987, Naval Ocean Systems Center, San Diego, CA.

3. F. White, *Data Fusion Research Survey*, for JDL, TPC^3, DFSP, Draft Version, May, 1987, Naval Ocean Systems Center, San Diego, CA.

4. H.L. Wiener, W.W. Willman, I.R. Goodman, and J.H. Kullback, *Naval Ocean-Surveillance Correlation Handbook, 1978*, NRL Rept. 8340, Oct.31, 1979, Naval Research Lab., Wash., D.C.

5. I.R. Goodman, H.L. Wiener, and W.W. Willman, *Naval Ocean-Surveillance Correlation Handbook, 1978*, NRL Rpt. 8402, Sept. 17, 1980, Naval Research Lab., Wash., D.C.

6. I.R. Goodman, *A General Model for the Contact Correlation Problem*, NRL Rpt. 8417, July 27, 1983, Naval Research Lab., Wash., D.C. (See shortened version in 18th IEEE Conf. Decis. & Cntrl., Dec., 1979, 383-388.)

7. M.J. Shensa and V.P. Broman, *A Basic Theory for a Bayesian Recursion Approach to Data Association*, NOSC Tech. Rpt. 1034, Dec., 1984, Naval Ocean Systems Center, San Diego, CA.

8. C.-Y. Chong and S. Mori, "Fusion algorithm for hierarchical multitarget tracking", *Proc. 7th MIT/ONR Wrkshp. C^3 Sys.*, Dec., 1984, 179-184.

9. C.L. Bowman and C.L. Morefield, "Multisensor fusion of target attributes and kinematics", *Proc. 3rd MIT/ONR Wrkshp. C^3 Sys., Vol. IV*, Sept., 1980, 1-42.

10. R. Dillard, *New Methodologies for Automated Data Fusion Processing*, NOSC Tech. Rpt. TR 364, Sept., 1978. (See also, *Research Needs for Artificial Intelligence Applications in Support of C^3*, NOSC Tech. Rpt. TR 1009, Dec., 1984.)

11. E.H. Shortliffe, *Computer-Based Medical Consultations: MYCIN*, Elsevier Co., New York, 1976.

12. I.R. Goodman, *PACT: An Approach to Combining Linguistic-Based and Probabilistic Information for Correlation and Tracking*, NOSC Tech. Doc. 878, March, 1986, Naval Ocean Systems Center, San Diego, CA.

13. I.R. Goodman, "Combination of evidence in C^3 systems", *Proc. 8th MIT/ONR Wrkshp. C^3 Sys.*, Dec., 1985, 161-166.

14. Joint Directors of Laboratories, Technical Panel for C^3, Basic Research Group (JDL, TPC3, BRG), *C^3 Handbook*, Draft Version, Jan., 1987, Naval Ocean Systems Center and Naval Personnel Research and Development Center, San Diego, CA.

15. I.R. Goodman, "A probabilistic / possibilistic approach to modeling C^3 systems", *Proc. 9th MIT/ONR Wrkshp. C^3 Sys.*, Dec., 1986, 53-58.

16. I.R. Goodman, "A probabilistic / possibilistic approach to modeling C^3 systems: Part II", *Proc. First Symp. C^2 Research*, 1987, to be published.

17. I.R. Goodman and H.T. Nguyen, *Uncertainty Models for Knowledge-Based Systems*, North-Holland Co., Amsterdam, 1985.

18. C.R. Rao, *Linear Statistical Inference and Its Applications, 2nd ed.*, John Wiley Co., New York, 1973.

19. I.R. Goodman, "A unified approach to modeling and combining evidence through random set theory", *Proc. 6th MIT/ONR Wrkshp. C^3 Sys.*, Dec., 1983, 42-47.

20. I.R. Goodman, "Some new results concerning random sets and fuzzy sets", *Inf. Sci., 34*, 1984, 93-113.

21. E.M. Oblow, *Extension of O-Theory to Problems of Logical Inferencing*, Rpt. No. ORNL/TM-10107 (CESAR 87105), Oak Ridge National Lab., March, 1987.

22. A. Sklar, "Random variables, joint distributions, and copulas", *Kybernetika* (Czechoslavakia), *6*, 1973, 449-453.

23. I.R. Goodman, "Some fuzzy set operations which induce homomorphic random set operations", *Proc. 26th Conf. Gen. Sys. Research. (AAAI)*, 1982, 417-426.

24. G. Shafer, *A Mathematical Theory of Evidence*, Princeton Univ. Press, 1976.

25. H.T. Nguyen, "On random sets and belief functions", *J. Math. Analysis & Applic., 65*, 1978, 531-542.

26. I.R. Goodman, "Some asymptotic results for the combination of evidence problem", *Math. Modelling, 8*, 1987, 216-221.

27. N. Rescher, *Many-valued Logic*, McGraw-Hill Co., New York, 1969.

28. D. Dubois and H. Prade, *Fuzzy Sets and Systems*, Academic Press, New York, 1980.

29. *Proceedings of (First) Workshop on Uncertainty and Probability in Artificial Intelligence*, based on conference held at UCLA, Aug 14-16, 1985, AAAI & RCA

30. R.C. Stalnaker, "A theory of conditionals", *Amer. Phil. Quarterly, Monograph Series No. 2*, N. Rescher ed., Oxford, 1968, 98-112.

31. D. Lewis, "Probabilities of conditionals and conditional probabilities", *The Phil. Rev., LXXXV, No. 3*, July, 1976, 297-315.

32. P. Calabrese, "The probability that p implies q", Cal. State College Rpt. and abstract in Amer. Math. Soc. Notices, 22, No. 3, April, 1975, A 430-31; see also, "An algebraic synthesis of the foundations of logic and probability", to appear in *Info. Sci.*, 1987-88.

33. E. Mendelson, *Boolean Algebra and Switching Circuits*, Schaum Outline Series, McGraw-Hill, N.Y., 1970.

34. I.R. Goodman and H.T. Nguyen, *An Algebraic Theory of Conditioning with Applications to Uncertainty Modeling*, monograph to be submitted, 1987-88.

35. I.R. Goodman, A measure-free approach to conditioning, *Proc. 3rd AAAI Workshop Uncert. A.I.*, 1987, based on conference held at Seattle, July 10-12, 1987.

36. I.R. Goodman and H.T. Nguyen, "Conditional objects and the modeling of uncertainties", to appear in *Fuzzy Computing*, M.M. Gupta et al., eds., North-Holland Co., New York, 1987-88.

37. T. Hailperin, *Boole's Logic and Probability*, North-Holland Co., New York, 1976.

38. Z. Domotor, *Probabilistic Relational Structures and Their Applications*, Tech. Rpt. 144, May 14, 1969, Psychol. Series, Inst. for Math. Stud. in Soc. Sci., Stanford University.

39. E.W. Adams, *The Logic of Conditionals*, D. Reidel Co., Dordrecht, Holland, 1975.

40. S. Rudeanu, *Boolean Functions and Equations*, North-Holland Co., Amsterdam, 1974.

41. I. Rubin and I. Mayk, "Markovian modeling of canonical C^3 systems components", *Proc. 8th MIT/ONR Wrkshp. C^3 Sys.*, Dec., 1985, 15-23.

42. A. Levis, "Information processing and decisionmaking organizations: a mathematical description", *Proc. 6th MIT/ONR Wrkshp. C^3 Sys.*, Dec., 1983, 30-38. (See the related paper by M. Tomovic and A. Levis, "On the design of organizational structures for command and control", *Proc. 7th MIT/ONR Wrkshp. C^3 Sys.*, Dec., 1984, 131-138.)