

UNCLASSIFIED



Australian Government
Department of Defence
Defence Science and
Technology Organisation

A Critique of Three Decision Support Techniques

Lewis Warren

National Security and ISR Division
Defence Science and Technology Organisation

DSTO-TN-1254

ABSTRACT

This document provides an in-depth critical analysis of three recent non-standard analytical techniques and algorithms. This analysis continues this author's ongoing study of the foundations of extant decision analysis techniques for the purpose of identifying those that are mathematically sound and robust. These techniques have been selected because they have been proposed for application in the military and intelligence arenas where decisions can have great strategic importance. Our analysis indicates that all rest on doubtful mathematical foundations to the extent that they would not be advisable to use for defence related decision making.

RELEASE LIMITATION

Approved for public release

UNCLASSIFIED

UNCLASSIFIED

Published by

*National Security and ISR Division
DSTO Defence Science and Technology Organisation
PO Box 1500
Edinburgh South Australia 5111 Australia*

*Telephone: 1300 333 362
Fax: (08) 7389 6567*

*© Commonwealth of Australia 2014
AR-015-838
January 2014*

UNCLASSIFIED

UNCLASSIFIED

A Critique of Three Decision Support Techniques

Executive Summary

In recent years a wide variety of decision support tools have been developed many of which incorporate non-standard mathematical methods, especially for the modelling of uncertainty with methods other than probability theory. Although some of these non-standard methods have been published in refereed journals, there is still a need for further in-depth critical analysis beyond the critical review of the referee. This need becomes more imperative if these tools are to provide a basis for critical or strategic decision making.

The three techniques reviewed in this report are the *Causal Influence Logic* incorporated in the SIAM tool developed by the Science Applications International Corporation (SAIC), the *Subjective Logic* of Audun Josang, and the *Recursive Noisy OR Operator* of John Lemmer and Don Gossink. These three techniques have been selected for critical analysis because they have been proposed for application in the military and intelligence arenas where decisions may have great strategic importance.

The following analysis of those techniques indicates that many of the claimed benefits are non-existent, and moreover, the presence of various arbitrary equations can only lead to arbitrary and non-rigorous results. Consequently, we conclude that it would not be advisable to apply these techniques, or software based on them, to defence decision making. Generally speaking, it would be prudent to examine carefully any new algorithms or techniques proposed for defence decision making before committing to software investments.

UNCLASSIFIED

UNCLASSIFIED

This page is intentionally blank

UNCLASSIFIED

Contents

1. INTRODUCTION.....	1
2. UNCERTAINTIES IN DECISION SUPPORT MODELS	1
2.1 Levels of Uncertainty	1
2.2 Higher-order Uncertainty Forms in Input Information to Models.....	3
2.3 Interdependent Model Elements	3
2.4 References on Uncertainty in Models.....	3
3. DETAILED CRITIQUE OF <i>CAUSAL STRENGTH LOGIC</i>	4
3.1 Introduction	4
3.2 The CAST Causal Inferencing Logic.....	4
3.3 Some Questionable Features of CAST.....	8
3.3.1 Ambiguous meaning of Causal Strengths	8
3.3.2 Overall Influence from Combination of C+ and C-.....	9
3.3.3 Combining Map and Non-map Causes on a Child Node	10
3.3.4 Dangers of the Independent Parent Assumption	10
3.4 Conclusions of the CAST Analysis	11
3.5 References for CAST Analysis	11
4. DETAILED CRITIQUE OF THE <i>RECURSIVE NOISY OR OPERATOR</i>	11
4.1 Introduction	11
4.2 Summary of the Derivation of the Generalised OR Operator	12
4.3 Some Comments on the RNOR Derivation.....	13
4.4 Conclusions of the RNOR Analysis.....	13
4.5 References for RNOR Analysis.....	14
5. DETAILED CRITIQUE OF <i>SUBJECTIVE LOGIC</i>	14
5.1 Introduction	14
5.2 The basic rationale of Josang's Subjective Logic.....	14
5.3 The two foundation equations of Subjective Logic	15
5.4 Analysis of the two foundation equations.....	16
5.4.1 Analysis of Equation 1	16
5.4.2 Analysis of Equation 2.....	16
5.5 Josang's Proposed Consensus Operator	18
5.6 Conclusions of the Subjective Logic Analysis	18
5.7 References for Subjective Logic.....	19
6. OVERALL CONCLUSION OF THE THREE CRITIQUES.....	20

UNCLASSIFIED

DSTO-TN-1254

This page is intentionally blank

UNCLASSIFIED

1. Introduction

In recent years a wide variety of decision support tools have been developed many of which incorporate non-standard mathematical methods, especially for the modelling of uncertainty with methods other than probability theory. Although some of these non-standard methods have been published in refereed journals, there is still a need for further in-depth critical analysis beyond the critical review of the referee. This need becomes greater if these tools are to provide a basis for making strategically important defence decisions. The objective of this report is to determine if an *ad hoc*, obtuse or non-standard mathematical technique exceeds the limits of what could be called reasonable mathematical reasoning, while not necessarily adhering strictly to pure mathematical logic. The three techniques reviewed in this report are the *Causal Influence Logic* incorporated in the SIAM tool developed by the Science Applications International Corporation (SAIC), the *Recursive Noisy OR Operator* of John Lemmer and Don Gossink, and *Subjective Logic* of Audun Josang.

2. Uncertainties in Decision Support Models

2.1 Levels of Uncertainty

Conceptual models and their embedded techniques may contain many different types of uncertainty (U) that limit their fitness to the real world and can compromise their application. This author has previously described [1] several levels of such uncertainties that may exist in analytical models. These levels are briefly described below.

These levels of potential U are decomposed in the following schema where higher numerical values actually represent lower levels of uncertainty. The U levels are grouped into two main classes A and B : the first class A is associated with the inherent model structure and its mechanisms, and the second class B concerns the U introduced by the sets of data/information that is processed by the model.

Class A Uncertainties

Level 1: Uncertainty in Objective or Problem Definition --

Uncertainty of the purpose of the analysis may be related to mistaken perceptions, confusion, lack of information, or complexity. Personal factors such as experience, skill and bias can influence the cognisance of a problem and some analytical methods such as personal construct theory, cognitive mapping or the soft systems methodology, may assist in consolidating the problem definition and variable identification. This level of U may also be invoked when a concept used to describe a problem is given a different interpretation by different individuals.

Level 2: Uncertainty in Model Conceptualisation --

At this level a conceptual model is adopted as the computational framework. A range of questions must be answered to define the broad characteristics of the model and fit it to reality in an adequate manner. Such considerations are:

What paradigm : Single or multiple models? Random variables or not?

What structure: Where are the boundaries? How detailed and what model granularity? Hierarchical or complex system inductive model?

Uncertainty can arise at this level if the choices made to answer these questions are inept and do not describe the problem characteristics with an adequate degree of verisimilitude. An example of U introduced at this level is when a probabilistic inferencing model is adopted where the characteristics of the problem would suggest that an inductive type information integration model would be more appropriate for decision making. Although it may not always be clear as to what represents an "adequate" fit, and because there may also be multiple adequate fits, there can still be cases of greatly oversimplified models based on unreasonable assumptions.

Level 3: Uncertainty in Computational Macro-Structure --

This level refers to the U introduced by model components such as the inferencing or clustering techniques, or type of nodal squashing function in neural nets, or information aggregation procedures. Overall, this level addresses uncertainty as the level of validity or rigour in the mathematical equations used in a model.

Level 4: Uncertainty in Computational Micro-Structure and Parameters --

This level refers to the U introduced by analytical method components such as arbitrary parameters as in squashing function gain and bias in the fuzzy cognitive mapping technique, aggregation optimism/pessimism degree parameters, or the prior conditional probabilities in a Bayes Net.

Class B Uncertainties

Two additional levels of U pertain to the individual information elements themselves and the set of information elements available.

Level 5: Uncertainty in Sample Evidence --

Quantity - the amount of information affects measure estimation.

Quality - conflicting information in a sample also affects estimation.

Level 6: Intrinsic Uncertainty within Information Elements --

This is U inherited from variable definition and measurement.

A qualitative concept associated with a variable may be inherently vague, and measurements pertaining to any type of variable may be approximate, subjective, or indirect.

The uncertainty of interest in the three techniques analysed in this report belongs mainly to Level 3 uncertainty, concerning the rigour of the computational structure. The

uncertainty in the information that these techniques actually process would be Level 6 uncertainty.

2.2 Higher-order Uncertainty Forms in Input Information to Models

Input information to models may also be qualitative and be based on subjective opinion of experts, for example subjective beliefs; this Level 6 U may then require special equations to cater for the lower input information content. This report will not describe the many methods that have been proposed for representing this type of elemental information but a detailed description of the development of belief functions theory can be found in [2]. The focus of this analysis will be mainly on the equations in these three particular techniques used to process this type of elemental information. However, a separate report by this author [3] does investigate different ways to distinguish and represent hybrid U forms in elemental information inputs.

2.3 Interdependent Model Elements

A real world characteristic that presents significant challenges to analytical techniques in models is the presence of interactions between aspects of the real world that need to be captured in the model. Modelling such interactions can then introduce Level 2 U which is associated with the adequacy of the fit of a model to the real world characteristics of a problem. The Recursive Noisy OR is one technique that has been proposed to address situations where input information elements may not be independent and are synergistically interactive.

2.4 References on Uncertainty in Models

- [1] Warren, L (2006) *Structural Uncertainties in Numerical Induction Models*, DSTO Technical Report, DSTO-TR-1895.
- [2] Vastava, S. and Mock, T. (2002) *Belief Functions in Business Decisions*, New York: Physica-Verlag.
- [3] Warren, L (2007) *On Modelling Hybrid Uncertainty in Information*, DSTO Research Report DSTO-RR-0325.

3. Detailed Critique of *Causal Strength Logic*

3.1 Introduction

SIAM is a computer tool developed by Science Applications International Corporation (SAIC) and based on some work published in 1994 by a group of people at the C3I Center at George Mason University [4]. The underlying model is a type of causal influence diagram called an "Influence Net" to which an analytical process called Causal Strengths Logic (CAST) is applied. The following comments are based on the original report [4] as well as a report by Rosen and Smith [5] which is said by SAIC to provide the technical details of the method. The tool adopts a 5 step algorithm without the need to elicit sets of local conditional probabilities at each node as are required in pure probabilistic Bayes Nets. For simplicity, "causal strengths" [0,1] between parent and child pairs of binary state nodes are elicited from domain experts, instead of conditional probabilities for the complete set of different parent states for each child node. This report examines the foundations of the CAST algorithm within the SIAM tool.

3.2 The CAST Causal Inferencing Logic

For models that are composed of inter-related sets of binary (ON/OFF) random variables, n parent nodes require conditional probabilities for 2^n parent state combinations for each child state, if probabilistic inferencing is to be applied as in a Bayes Net. However, in many real world problems with minimal data, it is difficult to get meaningful values for all these probabilities rather than simply being guesses. For this reason, various simplified approaches have been proposed, such as the CAST algorithm, to sidestep the need to use the complete set of conditional probabilities. Thus, CAST only requires estimates of parent to child node causal influences, and the probabilities that the leaf nodes in the model are ON or OFF.

The CAST algorithm "causal influences" (C) need to be elicited for each parent/child pair. These may be promoting (+) or inhibiting (-) influences on a child state when a parent node is either TRUE (ON) or FALSE (OFF), yielding ON and OFF C values (h,g,) for each parent/child pair (as in Figure 1), i.e. a parent concept can still influence a child node state by promoting or inhibiting it even when the parent is in an OFF state. A probability called "b" also needs to be elicited for the algorithm, which is the probability that the child occurs due to a cause *other than* any of the parent nodes depicted.

$C_{X/A} = 0$ Means Parent A has no influence on Child X

$C_{X/A} = 1$ Means Parent A totally determines or promotes the Child X

$C_{X/A} = -1$ Means Parent A totally inhibits Child X

The exact meaning of causal influence between a Parent A and Child X is difficult to ascertain from the CAST literature but it is hinted at in an early paper [5, p.4]. In the following quotation from that reference P(C) is the prior probability that child (C) is ON.

Although computationally equivalent to the conditional probabilities of a child node given a parent node, the interpretation is somewhat different... The value assigned to $C_{C/A}$ indicates that if A is valid then the probability of C increases by $(C_{C/A}) (1-P(C))$.

In this expression it is difficult to understand why the amount the $P(\text{Child}=\text{ON})$ is updated by should be a proportion (the causal strength) of the prior probability that the child is OFF. This seems to be an attempt to compute a "turning-on" strength or probability as a function of the degree of probability that the child is OFF. However, this concept and function does not appear in the equations presented in [6] so this early explanation of causal strength in the CAST literature may have been discarded, although this is unclear because of the incomplete nature of the CAST literature.

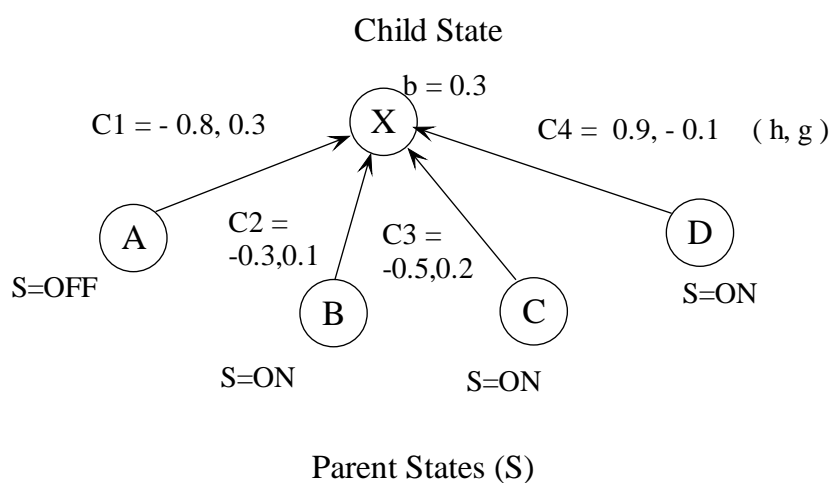


Figure 1: Causal Strengths (h, g)

There are five main steps in the CAST algorithm which lead to an estimate for $P(X=\text{ON})$ from the conditional probability of the Child occurring for each of the 2^n Parent node ON or OFF sets (as determined by equations 4.1 and 4.2).

Step 1: Aggregate promoting parent influences strengths

$$\text{Promoting state influences (+)} \quad C_+ = 1 - \prod (1 - C_+) \quad (1)$$

Step 2: Aggregate inhibiting parent influence strengths

$$\text{Inhibiting state influences (-)} \quad C_- = 1 - \prod (1 - |C_i|) \quad (2)$$

Step 3: Aggregate overall influence (O) as the net causal influence, from the promoting and inhibiting parents in that parent state set, using equations below.

$$1 - O = \frac{1 - C_+}{1 - C_-} \quad \text{when } C_+ \geq C_- \quad (3)$$

$$1 - O = \frac{1 - C_-}{1 - C_+} \quad \text{when } C_- > C_+ \quad (3)$$

These can be rearranged for explanatory purposes as below.

$$O = \frac{C_+ - C_-}{1 - C_-} \quad \text{when } C_+ \geq C_- \quad (3.1)$$

$$O = \frac{C_- - C_+}{1 - C_+} \quad \text{when } C_- > C_+ \quad (3.2)$$

Rosen and Smith [6] explain the above equations as follows.

In order to combine the aggregated positive and negative causal strengths, we introduce the following axiom:

Cancellation Axiom:

Let $(1-C_+)$ denote the potential of a child's occurrence being promoted due to a set of parents, and let $(1-C_-)$ denote the potential of a child's occurrence being inhibited by a set of parents. Then, there is an overall influence, O , that represents the net influence of the set of parents. The overall influence is given by the ratio of the aggregated Promoting and Inhibiting influences. Heuristically, this axiom asserts that the accumulated influence of all parents (specified in the conditioning case) is partitioned into: a portion that balances out the "opposing side;" and the remaining overall influence.

The meanings of (3), (3.1) and (3.2) are not clear and the question could be asked of the above explanation: Why should $(1-C)$ be the potential of one type of influence (promoting or inhibiting) when the aggregate influence of that type is C ?

Obviously, the complement is referring to the amount of one type of influence not yet activated, and the ratio of these two unactivated types of influences as in (3), equals the unactivated overall net influence. Although this is termed an axiom, the meaning and rationale is questionable. What equation (3.1) effectively computes is O , as the fraction of the net degree of promoting influence over the degree that the child state will not be inhibited, i.e. the difference $(C_+ - C_-)$ as a fraction of the degree child is not inhibited.

And similarly for (3.2). This is somewhat similar to the conditional probability:

$P(\text{Promoted instantiation} | \text{Not Inhibited}) = [P(\text{Promoted}) \cap P(\text{Not Inhibited})] / P(\text{Not Inhibited})$.

But C s are explicitly stated in [6] not to be probabilities. Furthermore, the numerator is incorrect if they are assumed to be probabilities since the subtraction of two unrelated

probabilities in (3.1) and (3.2) is inadmissible. So the axiomatic meaning of (3) is questionable, as is the O value so computed.

Example O computation by (3.1):

For Figure 1: $C_- = 0.65$, $C_+ = 0.93$ by (1) and (2)

$$\begin{aligned} O &= \frac{C_+ - C_-}{1 - C_-} \\ &= \frac{0.93 - 0.65}{1 - 0.65} \\ &= \frac{0.28}{0.35} \\ &= 0.8 \end{aligned}$$

Step 4: Compute probability of child for the j th set of parent states:

$P(X/\text{parent state set})$ by equations (4)

The power set of all parent ON and OFF combinations represents all possible combined parent state sets. For each node in each set take the "h" or "g" value as determined by the node's ON or OFF state. For each string of + and - values compute C_+ from (1) and C_- from (2). Then determine overall influence (O) from (3.1) or (3.2) i.e. 2^n values for O .

Then substitute O and "b" into (4) where b_x is an input called the *baseline probability* of child which is the probability that X occurs through some cause *other than* by any of the parent nodes on the map. Thus a probability b is combined with a non-probability when determining $P(X)$ for every Parent State Set.

$$\begin{aligned} P(X)/j\text{th Parent State Set} &= b_x + (1 - b_x)O_j && \text{for } O_j \geq 0 \\ &= b_x + O_j - b_x O_j && \end{aligned} \quad (4.1)$$

$$\begin{aligned} \text{And} &= b_x (1 - |O_j|) && \text{for } O_j < 0 \\ &= b_x - b_x |O_j| && \end{aligned} \quad (4.2)$$

Step 5: Compute $P(X= ON)$ by Aggregating All Parental State Set Influences

After using (1) to (4) to compute $P(X)$ /each Parent State Set for all 2^n Parent State Sets with combinations of ON and OFF activations, the overall $P(X)$ at each node is computed by (5). This is Jeffrey's rule for mutually exclusive sets of parent states (but not necessarily mutually exclusive parent states).

$$P(X) = \sum_{j=1}^{2^n} P(X | j^{\text{th}} \text{ Parent State Set}) \times P(j^{\text{th}} \text{ Parent State Set}) \quad (5)$$

As Rosen and Smith state in [6], in real world situations it would probably not be feasible to calculate the complete joint probability matrix. So for simplicity the joint probability of Parent states in the set are computed assuming the *Parent states are unconditionally independent* by using the multiplication rule (as in the example below in [6] for $Y = \text{OFF}$, and $Z, A, B, C, D = \text{ON}$).

$$P(j^{\text{th}} \text{ Parent State Set}) = P(\neg Y, Z, A, B, C, D) = P(\neg Y)P(Z)P(A)P(B)P(C)P(D) \quad (6)$$

Evidence Propagation

Observation and evidence will determine that some nodes are ON and some are OFF. Probabilities are assigned to these nodes which then enable the parent state set probabilities to be updated using (6). Subsequently, $P(X)$ of all directly affected children and their children are updated using equations (1-4) as previously described and then (5).

3.3 Some Questionable Features of CAST

Some questionable features of the causal inferencing logic in SIAM follow.

3.3.1 Ambiguous meaning of Causal Strengths

Although Chang [4], one of the original developers of what is now called CAST logic, and also Rosan and Smith [6], state that causal strengths “ C ” are *not* probabilities, they are also described in [5] as being “*computationally equivalent to conditional probabilities*”. However, in equations (3.1) and (3.2), C cannot be interpreted as probabilities because the numerator would be an inadmissible probability expression as the subtraction of two unrelated probabilities, because probabilities can only be subtracted when one is a subset of the other, and this is not the case for C_+ and C_- . Furthermore, C is treated as a probability in equations (1) and (2) which are equivalent to the probabilistic Noisy OR expression.

In other types of strictly deterministic causal inferencing mechanisms, such non-probabilistic causal strengths have a clear meaning, which determine the strength of a parent’s stimulus as a proportion of its state activation level [0,1]. But those types of mechanisms (neural nets, system dynamics and fuzzy cognitive maps) require deterministic mechanisms for causal inferencing that are very different to the CAST logic. Thus, the latent ambiguity of the CAST causal strengths means that the expressions that use them will also be ambiguous and of limited meaning.

However, operations performed on causal strengths when they are probabilities of events or states, by implication must conform to probability laws. For example, they must be combined with logical meanings such as AND, OR, EXOR etc and cannot be used in algebraic expressions that are inadmissible probabilistic equations. So if C were to be interpreted as elicited probabilities then:

- C_+ represents probability of X due to at least one of the promoting parents.

- C_- represents probability of NOT X (inhibited) due to at least one of the inhibiting parents.

3.3.2 Overall Influence from Combination of C_+ and C_-

The next step in CAST is to combine C_+ and C_- to get a net influence by equations (3). Chang [4, p.9] explains the rationale for equations (3) thus:

Basically this procedure partitions the stronger influence into two parts. The first part is set equivalent to the weaker influence of opposite sign, and the second part is the net remaining causal influence. If $C_+ > C_-$ then $C_+ = 1 - (1 - O)(1 - C_-)$, where O is the overall influence.

After rearranging, this becomes (3.1):
$$O = \frac{C_+ - C_-}{1 - C_-}$$

Chang's explanation is also difficult to understand, nor does the **Cancellation Axiom** of Rosen and Smith [6] clarify this equation.

But whichever ambiguous meaning is ascribed to C , equations (3) are of little meaning for the following reasons:

If C s are considered as probabilities as derived by equation (2), the numerators in (3.1) and (3.2) are inadmissible by probability laws, so the expressions would be meaningless and cannot be interpreted as conditional probabilities.

If C s are not probabilities, equations (3), as well as Chang's and Rosen and Smith's explanations, are still questionable.

Thus aggregate O influence computation at Step 3 is very questionable which renders all subsequent computations based on it of dubious merit.

Alternatively, a purely probabilistic approach could be applied to compute O as follows.

Assume that for each parent state set we compute $P(X=ON)$ and $P(X=OFF)$ from the appropriate $\{h,g\}$ values for the promoting and inhibiting parent state values (as follows). Then the probability expression for computing a probabilistic overall O would be as follows.

When $C_+ > C_-$:

The aggregate influence O is the $P(X = ON)$ from Promotors and Inhibitors

$$\begin{aligned} &= P(\text{Caused by Promotors}) \text{ AND } P(\text{Not inhibited by Inhibitors}) \\ &= C_+ (1 - C_-) \end{aligned} \tag{7}$$

For example, for parent state set in Figure 1, the appropriate causal influence string is $\{0.3, -0.3, -0.5, 0.9\}$. Then with C_+ and C_- in (7):

$$\begin{aligned} C_+ &= 1 - (1-0.3)(1-0.9) = 1 - (0.7)(0.1) = 1 - 0.07 = 0.93 \\ C_- &= 1 - (1-0.3)(1-0.5) = 1 - (0.7)(0.5) = 1 - 0.35 = 0.65 \\ \text{Thus } O &= C_+ (1 - C_-) = 0.93 (1 - 0.65) = 0.325 \end{aligned}$$

Hence the recommendation would be to replace (3) by (7).

3.3.3 Combining Map and Non-map Causes on a Child Node

When $O_j \geq 0$, the probabilistic meaning of (4) for aggregating map and non-map causes for each 2^n state sets is that $P(X)$ is the probability that X is caused by non-map causes "b", OR is caused by map causes, AND NOT by non-map causes (i.e. only by aggregate map influences O). Thus, O is assumed to be a probability measure in expression (4) which applies for non-mutually exclusive events.

$$\text{i.e. } P(X | j^{\text{th}} \text{ Parent State Set}) = P(\text{caused by one or more Parent State sets}) \\ + P(\text{due to some external cause}) - P(\text{caused by both})$$

Then $P(X)$ is aggregated across all Parent State Sets by (5),

$$P(X) = \sum_{j=1}^{2^n} P(X | j^{\text{th}} \text{ Parent State Set}) \times P(j^{\text{th}} \text{ Parent State Set}) \quad \text{as by (5)}$$

Unfortunately, this summation across 2^n Parent State sets by (5) includes "b" too many times and would be an excessive estimate of $P(X)$. The non-map causal probability "b" should only be included once as shown in (9) below, with the j^{th} Parent State set probability being determined simply as in (8) as follows.

$$P(X | j^{\text{th}} \text{ Parent State Set}) = O_j \quad (8)$$

Then, still adopting the assumption that the parent states are unconditionally independent, the $P(j^{\text{th}} \text{ Parent State Set})$ would be determined by multiplication of individual Parent state probabilities. For the previous example as appears in [6]:

$$P(\neg Y, Z, A, B, C, D) = P(\neg Y)P(Z)P(A)P(B)P(C)P(D)$$

Next,

$$P(X) = \sum_{j=1}^{2^n} P(X | j^{\text{th}} \text{ Parent State Set}) \times P(j^{\text{th}} \text{ Parent State Set}) + b - b \Sigma(.) \quad (9)$$

Thus, if we applied standard probabilistic reasoning, then (4) with the spurious fractional probabilities, and (5) with the excessive inclusion of "b", could be replaced by the above two new equations (8) and (9) respectively.

3.3.4 Dangers of the Independent Parent Assumption

While parents are conditionally independent, they are not unconditionally independent since they may have one or more common parents themselves. This means that knowledge of the state of one parent of a child may be derived through knowledge of another parent of the same child, by indirect inferencing on what the state of the common parent to those two parents must be. In other words, knowledge of the two parent states may be linked through knowledge of a common parent state, i.e. not independent states.

As Rosen and Smith clearly state in [6], it is not realistic to try to elicit the full set of parent child conditional probabilities in real problems, so they adopt the independence assumption. Unfortunately, it is impossible to gauge the effect of this assumption since there would exist no true set of conditional probabilities to compare the values so computed.

3.4 Conclusions of the CAST Analysis

Chang and his colleagues have attempted to combine some non-probabilistic inferencing measures, as used in some expert systems, with probabilistic inferencing measures resulting in some dubious expressions with no clear meaning. And even though the numerical results of CAST are constrained to $[0,1]$, as probabilities must be, the computed result of equation (3) for overall influence O does seem to be highly questionable. Nevertheless, the first problem in computing the overall influence O could be remedied by replacing equations (3) with (7). Another problematic feature is the combination of map and non-map causes with the excessive use of "b". However, this could also be remedied by replacing (4) by (8) and (5) by (9).

3.5 References for CAST Analysis

- [4] Chang, K., Lehner, P., Levis, A., Zaidi, S., Zhao, X. (1994) *On Causal Influence Logic*. Technical Report of George Mason University (GMU/C3I-154-R).
- [5] Rosen, J. and Smith, W. (1994) Cooperative Security Issues: Assessing the effect of situational influence measures, *Proc. 1994 Symposium on C2 Research and Decision Aids*, Monterey CA , pp 510-524.
- [6] Rosen, J. and Smith, W. (1996) Influence Net modeling with causal strengths: An evolutionary approach, *Proc. 1996 Command and Control Research and Technology Symposium*.

4. Detailed Critique of the *Recursive Noisy OR Operator*

4.1 Introduction

This critique examines a paper by Lemmer and Gossink [7] that proposes a generalised version of the probabilistic disjunctive OR operator, called the Recursive Noisy OR (RNOR), which avoids the independent event requirement of the traditional OR operator. The assertion is that the recursive RNOR operator can be applied to estimate the probability of an event which can have either independent causes, or interdependent and interactive causes. To commence, a summary of the method used to develop the generalised operator will be described and examined to determine how the independency constraint has been relaxed.

4.2 Summary of the Derivation of the Generalised OR Operator

The steps in the RNOR derivation are described below and demonstrated for the case of three events (x_1, x_2, x_3) . The same logic then applies to RNOR derivations with higher numbers of events.

1. The standard Noisy OR or NOR is:

$$NOR(x_1, x_2, \dots, x_n) = 1 - \prod_{i=1}^n (1 - p(x_i)) \text{ where } p(x_i) \text{ probability of event } x_i \text{ cause}$$

Then for three causal events x_1, x_2, x_3 :

$$NOR(x_1, x_2, x_3) = 1 - (1 - p(x_1)) (1 - p(x_2)) (1 - p(x_3)) \quad (1)$$

2. Next create a fraction by raising right side of (1) rearranged to the power $(n-1)$ and let the denominator be right side of (1) rearranged raised to power $(n-2)$. Thus, right side of (1) rearranged is unchanged and this step yields the general equation A1 in Appendix A of [7].

$$1 - NOR(x_1, x_2, x_3) = \frac{[(1 - p(x_1)) (1 - p(x_2)) (1 - p(x_3))]^2}{[(1 - p(x_1)) (1 - p(x_2)) (1 - p(x_3))]} \quad (2)$$

3. Next rearrange the numerator of right side of (2) as follows.

$$1 - NOR(x_1, x_2, x_3) = \frac{(1 - p(x_1)) (1 - p(x_2))}{(1 - p(x_1))} \cdot \frac{(1 - p(x_1)) (1 - p(x_3))}{(1 - p(x_3))} \cdot \frac{(1 - p(x_2)) (1 - p(x_3))}{(1 - p(x_2))} \quad (3)$$

4. In numerator on right side of (3) insert $(1 - \text{standard NOR})$ for the relevant event pairs.

$$1 - NOR(x_1, x_2, x_3) = \frac{(1 - NOR(x_1, x_2))}{(1 - p(x_1))} \cdot \frac{(1 - NOR(x_1, x_3))}{(1 - p(x_3))} \cdot \frac{(1 - NOR(x_2, x_3))}{(1 - p(x_2))} \quad (4)$$

5. The format of this equation is then called the RNOR whereby the disjunctive OR for higher numbers of events can be recursively determined from the disjunctive OR combinations of lower numbers of events.

$$RNOR(x_1, x_2, x_3) = 1 - \frac{(1 - NOR(x_1, x_2))}{(1 - p(x_1))} \cdot \frac{(1 - NOR(x_1, x_3))}{(1 - p(x_3))} \cdot \frac{(1 - NOR(x_2, x_3))}{(1 - p(x_2))} \quad (5)$$

Thus, in (5) the numerator NOR values can be computed for independent events, or alternatively NOR values inserted for interactive combinations if they are known, as was implemented in the example in [7].

In Example 2 in [7] there are four causal events. Some causal event combinations act independently, while other causal event combinations are synergistically interactive with those greater NOR values being given in the input data. Table 1 in [7] lists all recursive computations in the example and here we show how RNOR is computed for one such combination of causal events $\{x_1, x_2, x_3\}$.

Input Data: x_1 and x_2 are independent with $p(x_1) = 0.25$, and $p(x_2) = 0.3$
 x_1 and x_3 are independent with $p(x_1) = 0.25$, and $p(x_3) = 0.35$
 x_2 and x_3 are interactive with given NOR = 0.68 (> independent event NOR = 0.545)

$$1 - \text{NOR}(x_1, x_2) = (1 - 0.25)(1 - 0.3) ;$$

$$= 0.5250$$

$$1 - \text{NOR}(x_1, x_3) = (1 - 0.25)(1 - 0.35)$$

$$= 0.4875$$

$$1 - \text{NOR}(x_2, x_3) = (1 - 0.68)$$

$$= 0.3200$$

$$\text{RNOR}(x_1, x_2, x_3) = 1 - \frac{0.5250}{(1 - 0.25)} \cdot \frac{0.4875}{(1 - 0.35)} \cdot \frac{0.3200}{(1 - 0.30)}$$

$$= 1 - (0.700)(0.750)(0.457)$$

$$= 1 - 0.240$$

$$= 0.760 \text{ (as in Table 1 in [7])}$$

4.3 Some Comments on the RNOR Derivation

1. Step 2 in the derivation only raises the standard NOR to the power of one, thus leaving it unaltered.
2. Neither does rearranging the order of terms in the numerator as in (3) change the standard NOR or its fundamental requirement for independent events.
3. The requirement for independent events in RNOR as in (5) can be understood if the denominator for the three events case is expanded as follows.

$$(1 - p(x_1))(1 - p(x_2))(1 - p(x_3)) = 1 - p(x_1) - p(x_2) - p(x_3) + p(x_1)p(x_2) + p(x_1)p(x_3) + p(x_2)p(x_3) - p(x_1)p(x_2)p(x_3)$$

Thus x_2 and x_3 are considered as independent in the denominator of RNOR, while for the same event combination an interactive value is concurrently inserted in the numerator (as in the example above).
4. Not surprisingly, when all events are independent the RNOR computation results in the same value as the standard NOR equation (1).

4.4 Conclusions of the RNOR Analysis

When events are all independent the RNOR equation results in the same output values as the traditional NOR equation. However, the RNOR is also claimed to overcome the independent event constraint of the NOR equation. But in [7] there is no formal proof or informal reasoning to show why RNOR can be used with non-independent events. And as has been demonstrated in the example from [7], the RNOR equation does combine independent event values together with interactive values for the same pairs, and this does seem incoherent and inadmissible. Thus, we conclude that RNOR is only valid for independent events and will yield false output values if interactive event values are inserted.

4.5 References for RNOR Analysis

- [7] Lemmer, J. and Gossink, D. (2004) Recursive Noisy OR-A Rule for Estimating Complex Probabilistic Interactions, *IEEE Trans. Systems, Man and Cybernetics-Part B*; 34(6), pp. 2252-2261.

5. Detailed Critique of *Subjective Logic*

5.1 Introduction

Subjective Logic (SL) theory was developed by Audun Josang [8,9] commencing around 1997. A large number of developments occurred while Josang was at the DSTC commercial arm of the Queensland University of Technology. The SL theory is based on two equations intended to convert subjective estimates of the likelihoods of the truth of propositions or hypotheses into probabilities, which can then be manipulated by the standard probability operations in an inferencing process or model. This critique will first examine the foundational equations, and various problematic theoretical issues will be itemised. Subsequently, the proposed operator for combining expert opinions will be examined. The central theory of SL will be examined from the descriptions of SL in the source publications and documents listed in the References.

5.2 The basic rationale of Josang's Subjective Logic

Josang describes the benefits of SL in the abstract of [10]:

The limitation of traditional probabilistic logic is that it is unable to express uncertainty about the probability values themselves. This paper provides a brief overview (of) subjective logic which is a probabilistic logic that explicitly takes uncertainty about probability values into account...

...Subjective logic is directly compatible with binary logic, probability calculus and classical probabilistic logic. The advantage of using subjective logic is that real world situations can be more realistically modelled, and that conclusions more correctly reflect the ignorance and uncertainties about the input arguments.

Josang also proposes that SL is an extension of the Dempster-Shafer (DS) Belief theory [10, p.1].

Belief theory represents an extension of classical probability by allowing explicit expression of ignorance. Belief Theory had its origin in a model for upper and lower probabilities proposed by Dempster in 1960. Shafer later proposed a model for expressing beliefs. The main idea behind belief theory is to abandon the additivity principle of probability theory, i.e. that the sum of probabilities on all pairwise disjoint states must add up to one. Instead belief theory gives observers the ability to assign so-called belief mass to any subset of the state space, i.e. to non-exclusive possibilities including the whole state space itself. The main advantage of this approach is that ignorance, i.e. lack of information, can be explicitly expressed e.g. by assigning belief mass to the whole state space.

5.3 The two foundation equations of Subjective Logic

The rationale for the two foundation equations of SL is described in [11, p.11-12]:

Subjective logic represents a specific belief calculus that uses a belief metric called opinion to express beliefs. An opinion (regarding an hypothesis x) denoted by $w_x^A = (b_x^A, d_x^A, u_x^A, a_x^A)$ expresses the relying party A 's belief in the truth of a statement. Here b , d , and u (and a omitted in original) represent belief, disbelief and uncertainty, and relative atomicity respectively where $b_x^A, d_x^A, u_x^A, a_x^A \in [0,1]$ and the following equation holds:

$$b_x^A + d_x^A + u_x^A = 1 \quad (1)$$

In this context "disbelief" means the degree of belief that a hypothesis or proposition is not true, and in [10, p.1] Josang describes the meaning of "uncertainty" as :

...the term uncertainty will be used in the sense of "uncertainty about the probability values".

Also, "relative atomicity", which is later called "baseline probability" or "base rate", is the a priori probability according to the principle of insufficient reason, which simply assigns uniform probabilities across the possible states or hypotheses, e.g. for two states $a = 0.5$, and for three states or hypotheses $a = 0.333$.

The meaning of atomicity or base rate is also described in [11, p.12] :

The parameter a_x^A represents the base rate of x and reflects the size of the state space from which the statement x is taken. In most cases the state space is binary, in which case $a_x^A = 0.5$. The relative atomicity is used for computing an opinion's probability expectation value expressed by:

$$E(w_x^A) = b_x^A + a_x^A u_x^A \quad (2)$$

In [9, p.299] and [10, p.3] Josang states:

It can be shown that E_x satisfies the additivity principle:

$$E_x(0) = 0 \quad \text{and} \quad \sum_{x \in X} E_x(x) = 1$$

The base rate vector of Def.3 (in [10]) expresses non-informative a priori probability, whereas the probability expectation function of Eq.xx (Eqn 2 above) expresses informative a posteriori probability.

Also in [11, p.13]:

Let w_x and w_y be two opinions. They can be ordered according to the following rules by priority:

- 1) *The opinion with the greatest probability expectation is the greatest opinion.*
- 2) *The opinion with the least uncertainty is the greatest opinion.*

The notions and equations above outline the rationale of SL which aims to convert input subjective belief values and uncertainty estimates into probabilities (called "expectations" in SL) upon which decisions can be based. All subsequent developments in SL theory are based on, and utilise, the above two equations. This parallels the same process in Smets' Transferable Belief Model which converts input subjective beliefs and likelihood values into what Smets calls pignistic probabilities for decision making.

5.4 Analysis of the two foundation equations

5.4.1 Analysis of Equation 1

As Josang has indicated, SL is an extension from DS belief theory so first we will look at the DS basic equations for the two basic DS measures of Belief (Bel) and Plausibility (PL). Here PL is a measure of what could feasibly be true based on an overlapping inconclusive set of data, and Bel is based on a smaller set of hypothesis supporting data but still with embedded uncertainty. In this way, PL is an upper or higher value than Bel, which is based on a smaller set of evidence. Thus, in DS belief or evidence theory:

$$\begin{aligned} \text{Bel}(X) &< \text{PL}(X) \\ \text{Bel}(X) &< 1 - \text{Bel}(\text{Not } X) \quad \text{or} \quad \text{Bel}(X) + \text{Bel}(\text{Not } x) < 1 \quad (\text{i.e. non-additivity}) \end{aligned}$$

Josang in SL adds some uncertainty (U) to the second DS equation above to make the equality:

$$\text{Bel}(X) + \text{Bel}(\text{Not } X) + \text{Uncertainty} = 1 \quad (\text{the previous equation 1})$$

In this fundamental SL equation (where Bel (Not X) = Disbelief), it can be seen that U is the amount of uncertainty in the aggregate of both the belief and disbelief estimates, which are not exactly complementary because they are non-additive.

5.4.2 Analysis of Equation 2

SL equation 2 below will be examined from 5 viewpoints with associated issues.

$$E(w_x^A) = b_x^A + a_x^A u_x^A \quad (2)$$

Viewpoint 1:

As Josang states, the subjective *uncertainty* estimate in this equation only relates to the subjective belief estimate Bel(X) of the likelihood of X being true. And by being constrained to the [0,1] interval this represents the proportion of Bel(X) that is uncertain, e.g. Somewhat Uncertain = 35% Uncertain, in magnitude of Bel(x) only.

Issue 1:

It is notable that this is a different meaning of uncertainty to that in equation 1, and for this reason, these two equations should not be combined in inferencing operations.

Viewpoint 2:

Since both b and u can be anywhere along the [0,1] interval, some combinations will result in the dependent probability variable (E) being greater than one.

For example: b = very likely = 90% , and u = 55%

$$E = 0.9 + (0.5)(.55) = 0.9 + 0.275 = 1.175, \text{ for a baseline probability } a = 0.5.$$

Issue 2:

Thus, with some belief and uncertainty subjective inputs (b and u) one requirement of probability measures may be violated by this equation in many instances (about 40% of combinations). Also, if the subjective belief likelihood inputs are erroneously considered to

be E , the dependent belief values b then calculated from equation 2 will be less than zero, which are also inadmissible values for probability measures.

Issue 3:

Then, if a single pair of subjective inputs can be greater than one, the sum of several E estimates for different hypotheses may also be greater than one. Thus, Josang's unproven claim that, $\sum_{x \in X} E_X(x) = 1$ may definitely be invalid in many instances.

Viewpoint 3:

What is the rationale behind the product 'au' in equation 2?

Issue 4:

The product 'au' in equation 2 has no meaning because the meaning of the subjective estimate 'u' is an uncertainty degree in the belief input likelihood 'b', and not in the *a priori* value 'a'.

Viewpoint 4:

If 'u' is to be used to add a tolerance degree (or error) to the input subjective belief estimate, the equation should be: Probability = $b \pm 0.5 u$.

Thus, equation 2 effectively inflates the subjective belief estimate on the high side only, while it could equally well be that the input belief estimate is an excessive estimate itself and thus should be decreased.

Issue 5:

In this manner, atomicity should not be used at all because it does not relate to the meaning of 'b' and 'u'. The uncertainty degree should always be equally split (0.5) to increase and decrease the input subjective belief estimate by equal amounts. Thus the + sign in equation 2 is invalid from a margin of error (or tolerance) viewpoint, according to Josang's definition of the meaning of uncertainty in these equations.

Viewpoint 5:

But if we do assume the 'au' product to be valid, we can further examine the possible meaning of equation 2 according to the meaning of probabilistic operators. The 'au' product is effectively a proportion of the baseline *a priori* probability value. The addition then adds this probability *portion* to the input subjective likelihood, which itself is simply a subjective probability estimate. In probability theory, when two probabilities are combined by a simple addition they must be for mutually exclusive events or states. But these two represent the *a priori* and *a posteriori* estimates for the same event, state, or hypothesis. Thus, this equation is fundamentally different to Bayes equation for estimating an *a posteriori* value updated from evidence, and which does not assume mutually exclusive prior and posterior states.

Issue 6:

The addition operation is invalid according to the meaning ascribed to these variables, because only a single truth value or likelihood is being estimated.

Based on all the issues identified from different viewpoints, the overall conclusion is that the fundamental Equation 2 of SL theory is questionable. Furthermore, it should not be combined with equation 1 in inferential mathematics because the uncertainty measure has a different meaning in each equation.

5.5 Josang's Proposed Consensus Operator

In [13] Josang defined the Consensus Operator for combining two expert opinions (A,B) about a single hypothesis as follows:

$$\text{Consensus}(A,B) = \text{Bel}(A,B) = [\text{Bel}(A) U(B) + \text{Bel}(B) U(A)] / k \quad (3)$$

where $k = U(A) + U(B) - U(A)U(B)$

More than two input beliefs would then be combined iteratively.

Issue 7:

No derivation (or meaning) of this equation is presented in the original paper [13] in 2002. Another recent paper that is purported to generalise the Consensus expression, states [12, p. 196] that it can be derived using a Dirichelet distribution. However, this author has not been able to locate such a derivation in any publication.

Issue 8:

Equation (3) is associative such that the order of combining elements does not affect the combined value. However, it is not idempotent such that $a \cup a \neq a$ and $a \cap a \neq a$. This property can then yield combined belief values very different to the input belief values as shown below.

Inputs: $b_1=0.8, u_1=0.8; b_2=0.8, u_2=0.8$ >> Output combined belief = 1.33
 $b_1=0.6, u_1=0.4; b_2=0.6, u_2=0.4$ >> Output combined belief = 0.75
 $b_1=0.5, u_1=0.9; b_2=0.5, u_2=0.9$ >> Output combined belief = 0.91

These examples illustrate that the combined value:

- can be significantly different to the inputs.
- can be > 1 (which exceeds the [0,1] constraint for belief values)

For example, the last set of inputs represents the maximum uncertainty ($b=0.5$) of the experts committing to a binary belief value (a proposition being either true or false), and additionally there is also a very high degree of uncertainty about being so undecided ($u=0.9$). Nevertheless, the combined belief value is almost certainty (0.91) which is very different to the input states of belief. This result would seem to be very hard to justify.

5.6 Conclusions of the Subjective Logic Analysis

Subjective Logic has been proposed for modelling and reasoning with higher-order uncertainty as inherent in subjective likelihood estimates for truth values of propositions. Also, associated with those subjective likelihood estimates are uncertainty estimates, as

subjective degrees of certainty in the likelihood estimates themselves. The foundations of Subjective Logic are two simple linear equations about which we conclude the following.

- They are incompatible because of the different meaning of uncertainty in each.
- Equation 2 is invalid for estimating probabilities because it can violate a basic probability requirement in many instances.
- Equation 2 is also undesirable from several other theoretical viewpoints.

As with other systems for processing belief measures in logical models, the rough input beliefs or truth likelihoods are converted to probabilities, called “expectations” in Subjective Logic, which are then processed through a proposition tree using probability operators.

Since we have concluded that the two fundamental equations for generating probabilities from the input belief measures are invalid, the probability values computed from those belief inputs will also be questionable. Hence, the final output probability for the head proposition in the inference tree will also be invalid. Furthermore, the Consensus operator proposed by Josang for combining multiple expert beliefs yields results for many combinations of inputs that are significantly different to the inputs, and for which no logical or mathematical reason has been presented. Thus, the validity and usefulness of that expression is also very questionable.

For the various reasons above, overall we conclude that the mathematics of Subjective Logic rest on very questionable foundations, and if any computational results are used for decision making there will be substantial risk as to their validity.

5.7 References for Subjective Logic

- [8] Josang, A. (1997) Artificial reasoning with subjective logic, in *Proc. 2nd Australian Workshop on Commonsense Reasoning*, (Eds. Nayak and Pagnucco), Perth, December 1997.
- [9] Josang, A. (2001) A Logic for Uncertain Probabilities. *I.J. Uncertainty, Fuzziness, and Knowledge-based Systems*, **9**(3), pp. 279-311.
- [10] Josang, A. (2007) Probabilistic Logic Under Uncertainty, in *Proc. Computing: The Australian Theory Symposium (CATS2007)*, CPRT Vol 65, Ballarat 2007.
- [11] Pope, S. and Josang, A. (2005) Analysis of Competing Hypotheses using Subjective Logic, In *Proc. 10th ICCRT Symposium*, 2005.
- [12] Josang, A., Diaz, J. and Rifqi, M. (2010) Cumulative and averaging fusion of beliefs, *Information Fusion*, **11**(2), pp. 192-200.
- [13] Josang, A. (2002) The Consensus Operator for Combining beliefs, *Artificial Intelligence*, **141**, pp.157-170.

6. Overall Conclusion of the Three Critiques

When proposing non-standard analytical methods that deviate significantly from conventional mathematical logic the need for careful evaluation by multiple experts becomes imperative. Although this may seem obvious it does not always occur due to enthusiasm and various other reasons.

This type of in-depth analysis is necessary even though the methods may have been published in conference proceedings or in refereed journals. In other words, for non-standard techniques and algorithms, especially those to support important decision making for defence, the referees' opinions and approvals for publication may not be sufficient.

Our analysis of three such techniques has indicated that many of the claimed benefits are non-existent, and moreover, the presence of the proposed non-standard mathematical expressions would most likely lead to arbitrary and non-rigorous results. Needless to say, the outputs of these techniques would not then provide a basis for sound decision making.

Thus, a rigorous theoretical analysis of proposed techniques and algorithms by multiple experts may alert to significant and insurmountable weaknesses that would indicate that investments in software development would not be wise.

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA				1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)	
2. TITLE A Critique of Three Decision Support Techniques				3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION) Document (U) Title (U) Abstract (U)	
4. AUTHOR(S) Lewis Warren			5. CORPORATE AUTHOR DSTO Defence Science and Technology Organisation PO Box 1500 Edinburgh South Australia 5111 Australia		
6a. DSTO NUMBER DSTO-TN-1254	6b. AR NUMBER AR-015-838	6c. TYPE OF REPORT Technical Note		7. DOCUMENT DATE January 2014	
8. FILE NUMBER -	9. TASK NUMBER JNT07/329	10. TASK SPONSOR Wayne Hobbs	11. NO. OF PAGES 20	12. NO. OF REFERENCES 13	
13. URL on the World Wide Web http://dspace.dsto.defence.gov.au/dspace/			14. RELEASE AUTHORITY Chief, National Security and ISR Division		
15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT <i>Approved for public release</i>					
OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111					
16. DELIBERATE ANNOUNCEMENT No Limitations					
17. CITATION IN OTHER DOCUMENTS Yes					
18. DSTO RESEARCH LIBRARY THESAURUS http://web-vic.dsto.defence.gov.au/workareas/library/resources/dsto_thesaurus.shtml Decision support algorithms, Probability theory, Belief measures, Causal logic					
19. ABSTRACT This document provides an in-depth critical analysis of three recent non-standard analytical techniques and algorithms. This analysis continues this author's ongoing study of the foundations of extant decision analysis techniques for the purpose of identifying those that are mathematically sound and robust. These techniques have been selected because they have been proposed for application in the military and intelligence arenas where decisions can have great strategic importance. Our analysis indicates that all rest on doubtful mathematical foundations to the extent that they would not be advisable to use for defence related decision making.					