

10TH INTERNATIONAL COMMAND AND CONTROL RESEARCH AND TECHNOLOGY SYMPOSIUM

THE FUTURE OF C2

DECISIONMAKING AND COGNITIVE ANALYSIS

Analysis of Competing Hypotheses using Subjective Logic

Simon Pope, Audun Jøsang

CRC for Enterprise Distributed Systems Technology (DSTC Pty Ltd)
Level 7, General Purpose South
The University of Queensland 4072 Australia
Tel: +61-7-3365-4310 *Fax:* +61-7-3365-4311

{simon.pope, ajosang}@dstc.edu.au

<http://www.dstc.edu.au>

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE JUN 2005		2. REPORT TYPE		3. DATES COVERED 00-00-2005 to 00-00-2005	
4. TITLE AND SUBTITLE Analysis of Competing Hypothesis using Subjective Logic				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) CRC for Enterprise Distributed Technology (DSTC Pty Ltd),Level 7, General Purpose South,The University of Queensland,Brisbane 4072 Australia, ,				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 29	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

This page is intentionally left blank.

Analysis of Competing Hypotheses using Subjective Logic

Simon Pope, Audun Jøsang

CRC for Enterprise Distributed Systems Technology (DSTC Pty Ltd)
Level 7, General Purpose South
The University of Queensland, 4072 Australia
Voice: +61 7 3365 4310 Fax: +61 7 3365 4311
{simon.pope, ajosang}@dstc.edu.au

Abstract

Intelligence analysis is a complicated task that requires a high degree of analytical judgement under conditions of considerable uncertainty. This judgement is used to fill in the gaps in knowledge and is the analyst's principal means of managing uncertainty.

Much of intelligence analysis includes judging the relevance and the value of evidence to determine the likelihood of competing hypotheses. The challenge is to create better formal methods of analysis that can be used under a wider variety of circumstances and which can handle both empirical data and formally-expressed beliefs as evidence for or against each hypothesis.

The authors have developed a formal approach to the evaluation of competing hypotheses that is based on the belief calculus known as Subjective Logic. The development of this formal approach allows for integration of empirical and statistical data, as well as for judgements made by analysts. Lastly, this formal approach makes redundant the separate analysis of "diagnosticity of evidence". Under this formal approach, diagnosticity is formally derived from the model and need not be considered as a separate input to the model, except as a means of limiting the initial set of evidence that should be formally considered.

I. INTRODUCTION

Intelligence deals with all the things which should be known
in advance of initiating a course of action.

– *Intelligence Activities* [1], in Warner [2]

Intelligence is a difficult term to define precisely, yet its role and importance can be both intuitively understood and appreciated. From one perspective it may be seen as an end product – 'information' that is used to enhance or aid a decision making process. From yet another perspective, it refers to the process that is applied to information, in order to transform it into a more useful product [2]. More importantly than what intelligence *is* perhaps, is what intelligence *does*.

Intelligence, both as a product and a process, is a means by which better decisions can be made, based on an increased understanding of likely courses of action, their influences and their consequences. In everyday personal affairs few of our decisions use any directed analytical processes – and even fewer of these require any sort of rigorous approach – arguably due to relatively minor consequences of the decisions we face.

The same is not true for large-scale human affairs – such as the business of nations and corporations – where the complexity of the environment and the relative consequences of decisions can have enormous

The work reported in this paper has been funded in part by the Co-operative Research Centre for Enterprise Distributed Systems Technology (DSTC) through the Australian Federal Government's CRC Programme (Department of Education, Science, and Training).

impact on the well-being and survival of a nation's citizenry or a corporation's ability to compete. This distinction in and of itself is cause enough to consider whether human 'everyday reasoning' is robust and reliable enough for use in these larger contexts.

As it happens, humans systematically make substantive errors in reasoning due to problems of framing, resistance of mental models to change, risk aversion, limitations of short-term memory, and other cognitive and perceptual biases [3], [4], [5], [6], [7]. This has severe implications for the process of intelligence analysis, and may lead to incorrect conclusions, especially in situations that appear familiar but which actually result in different outcomes; in situations where the gradual assimilation of information into established mental models results in the failure to detect 'weak signals' that should have triggered a major re-evaluation; and in situations where the complexity of the mental models are untenable due to human limitations of short-term memory [8], [7], [9], [10].

When applied to the business of nation states, the consequences of intelligence failure can be disastrous, so much so that the recorded history of the world – both ancient and modern – is replete with a litany of devastating intelligence failures too numerous to list. Examples of these are easily found in any period of history – such as the failure of the United States to perceive an impending attack on Pearl Harbor – and the failure of Japan to reason that Midway Island was a trap, with the consequent sinking of four Japanese aircraft carriers and the loss of all crews, aircrews and aircraft.

It is therefore foolhardy to believe that good intelligence can be developed by relying solely on human cognition without resort to products, methodologies or frameworks that attempt to augment human cognition while also mitigating its defects. The management of intelligence analysis should encourage the application of products that allow clear delineation of assumptions and chains of inference; the specification of the degree of uncertainty about the evidence and resultant conclusions; and the elaboration of alternative perspectives and conclusions [7].

This paper proposes that the use of *Subjective Logic* [11] should be coupled with the *Analysis of Competing Hypotheses (ACH)* approach, [7], [12] as a basis for reasoning about alternative hypotheses as part of the process of intelligence analysis – an approach that is referred to as *ACH-SL*. The paper also argues that the *ACH-SL* approach allows uncertainty of evidence, opinion about their influences, and the resultant conclusions to be expressed in both quantitative and qualitative forms that are applicable for human and machine interpretation. Lastly, the paper shows that 'diagnosticity of evidence' can be derived from this approach and need not be considered as a separate input to the process, except as a means of limiting the initial set of evidence to be considered for inclusion into the model.

II. ALTERNATIVE ANALYSIS

Intelligence analysis generally requires that analysts choose from among several alternative hypotheses in order to present the most plausible of these as likely explanations or outcomes for the evidence being analyzed. Analysts that do not use some rigorous methodology will often work intuitively to identify what they believe to be the most likely explanation and then work backwards, using a satisficing approach where the 'correct' explanation is the first one that is consistent with the evidence [7]. The major downfall of the satisficing approach is that there may be more than one explanation that is consistent with the evidence, and unless the analyst evaluates every reasonable alternative, they may arrive at an incorrect conclusion. Other common problems with using this strategy include the failure to generate appropriate alternative hypotheses; the propensity to filter and interpret the evidence to support the conclusions; and the failure to consider the diagnosticity of the evidence and how well it differentiates between hypotheses. The recognition of these problems with their disastrous consequences has led to the development of *Alternative Analysis* techniques that are widely employed within the intelligence services¹.

¹Other strategies that are less commonly used in intelligence analysis and are also ineffective are discussed in detail elsewhere [13]

Many alternative analysis techniques attempt to address the problems of fixed mind-sets and incomplete generation of alternative hypotheses, while still others attempt to address the problems of reasoning about the alternative hypotheses [10]. One way in which some of the problems of reasoning about alternative hypotheses can be overcome is to require the analyst to simultaneously evaluate all reasonable hypotheses and reach conclusions about their relative likelihood, based on the evidence provided. However, simultaneous evaluation of all non-trivial problems is a near-impossible feat for human cognition alone. Recent research suggests the number of individual variables we can mentally handle while trying to solve a problem is relatively small – four variables are difficult, while five are nearly impossible [14]. The Analysis of Competing Hypotheses (ACH) [7] was developed to provide a framework for assisted reasoning that would help overcome these limitations.

Alternative Analysis, and in particular ACH, is seen as so important that the CIA's Sherman Kent School for Intelligence Analysis runs a monthly Alternative Analysis Workshop and has introduced an Alternative Analysis unit into the Career Analyst Program (CAP) – the basic training program of the CIA Directorate of Intelligence (DI) which introduces all new employees to the basic thinking, writing and briefing skills that are needed as an analyst [15].

III. HEUER'S ANALYSIS OF COMPETING HYPOTHESES (ACH)

The ACH methodology was developed in the mid- to late-1970's by Richard Heuer, a former CIA Directorate of Intelligence methodology specialist, in response to his "never-ending quest for better analysis" [7]. His ACH methodology is still considered to be highly relevant today [12]. ACH consists of the following eight steps:

Step-by-Step Outline of Analysis of Competing Hypotheses

- 1) Identify the possible hypotheses to be considered. Use a group of analysts with different perspectives to brainstorm the possibilities.
- 2) Make a list of significant evidence and arguments for and against each hypothesis.
- 3) Prepare a matrix with hypotheses across the top and evidence down the side. Analyze the "diagnosticity" of the evidence and arguments—that is, identify which items are most helpful in judging the relative likelihood of the hypotheses.
- 4) Refine the matrix. Reconsider the hypotheses and delete evidence and arguments that have no diagnostic value.
- 5) Draw tentative conclusions about the relative likelihood of each hypothesis. Proceed by trying to disprove the hypotheses rather than prove them.
- 6) Analyze how sensitive your conclusion is to a few critical items of evidence. Consider the consequences for your analysis if that evidence were wrong, misleading, or subject to a different interpretation.
- 7) Report conclusions. Discuss the relative likelihood of all the hypotheses, not just the most likely one.
- 8) Identify milestones for future observation that may indicate events are taking a different course than expected.

– Heuer, *Psychology of Intelligence Analysis* [7]

These eight steps are intended to provide a basic framework for identification of assumptions, arguments and hypotheses; consideration of all evidence and hypotheses – including its value relative to the hypothe-

ses; a method of disconfirmation for identifying the most likely hypotheses; an approach to reporting the results of the analysis; and an approach to detecting future changes in the outcomes.

In simple terms, ACH requires the analyst to simultaneously evaluate all reasonable hypotheses and reach conclusions about their relative likelihood, based on the evidence provided. Heuer acknowledges that while this holistic approach will not always yield the right answer, it does provide some protection against cognitive biases and limitations [7].

Of particular interest is Step 5, which requires the analyst to draw tentative conclusions about the likelihood of each hypothesis. It has been argued that ACH recommends analysts consider the likelihood of each hypothesis h given the assertion of each item of evidence, e , i.e. $p(h|e)$ [12]. However, this can reasonably be interpreted to mean that the negation of each item of evidence, \bar{e} should also be considered ($p(h|\bar{e})$). Consideration of counterfactuals has the advantage that the model can be constructed independently of known facts and continually evaluated if the value of the evidence changes over time. The difference in interpretation lies in whether the evidence with respect to the hypotheses is considered *a priori* or *a posteriori*.

Evidence can be constructed *a posteriori* by the analyst from the ‘facts at hand’, where the evidence has already been measured and valued, rather than from a general examination of the possible signs for each hypothesis. While examination of available data is usually relevant, ‘hidden facts’ – conditions which are not observable, or conditions which have not yet taken place – are also likely to be relevant to the analysis. If reasoning is conducted *a priori*, then the value of the evidence is uncertain, and the analyst is more likely to consider the consequences of it being false as well as the consequences of it being true. If the reasoning is *a posteriori*, the analyst may know whether the evidence is true or false, and not consider its counterfactual to be relevant in determining the likelihood of the hypothesis. This is a mistake, since the analysis model will no longer be relevant if the value of the evidence changes, or there is uncertainty about its value.

Richard Heuer makes the excellent point that analysts should interpret ‘evidence’ in its broadest sense and not limit oneself just to current intelligence reporting. Indeed ACH is able to model the absence of evidence as well as its presence, and when done diligently presents no conceptual problem. However, ACH does not *require* analysts to consider both the assertion and negation of evidence, and this deficiency may lead them to frame the problem in terms of a single view of evidence – which often leads to incorrect conclusions, especially if deception or denial is being undertaken by an adversary [12].

IV. ANALYSIS OF COMPETING HYPOTHESES – COUNTER DECEPTION (ACH-CD)

ACH-CD was developed by Frank Stech and Christopher Elässer of the MITRE Corporation as a modified variant of ACH to account for cognitive factors that make people poor at detecting deception [12]. They correctly argue that the use of ACH can lead to greater susceptibility for deception, especially when reasoning about a single view of evidence, i.e. the likelihood of each hypothesis given the assertion of the evidence $p(h|e)$. Their argument is that this type of reasoning neglects the base rates both of the evidence $br(e)$ and of the hypothesis $br(h)$ which can result in reasoning errors that lead to incorrect conclusions [16]. More correctly it should be said that reasoning using only one of the logical conditionals (usually the positive conditional, $p(h|e)$) is more likely to produce reasoning flaws than when both are considered [17]. Stech and Elässer make the same point when they argue that analysts’ judgements are more susceptible to deception if they also do not take the false positive rate of the evidence into account [12].

An excellent example of this provided by Stech and Elässer is how the reasoning about the detection of Krypton gas in a middle-eastern country can lead to the erroneous conclusion that the country in question likely has a nuclear enrichment program. For clarity, their example has been reproduced below [12]:

Detect Krypton

- $p(\text{enrichment} \mid \text{Krypton}) = \text{high}$
- $p(\text{enrichment program}) = \text{high}$
- $p(\text{nuclear program}) = \text{high}$

They argue that the main problem with this reasoning is that it does not consider that Krypton gas is also used to test pipelines for leaks, and that being a middle-eastern country with oil pipelines, the probability of the gas being used outside of a nuclear program is also fairly high, i.e.

$$p(\text{Krypton} \mid \text{not enrichment}) = \text{medium to high}$$

This additional information should lead the analyst to the conclusion that there is a fair amount of uncertainty of a nuclear program given the detection of Krypton. The assignment of the ‘high’ value to $p(\text{enrichment} \mid \text{Krypton})$ neglects the fact that an oil-rich middle-eastern country is likely to use Krypton gas – regardless of whether they have a nuclear program.

However, it can be argued that Stech and Elässer have interpreted Step 5 of ACH more narrowly than perhaps was intended. Heuer makes no claim about which of $p(h|e)$ or $p(e|h)$ – and their corresponding counterfactual $p(h|\bar{e})$, $p(e|\bar{h})$ – should be used. Heuer describes the process in such general terms as to be consistent with either interpretation, although consideration of counterfactuals is essential if basic reasoning errors are to be avoided.

In any event, it can be shown² that $p(h|e)$ and $p(h|\bar{e})$ can be derived from knowledge of $p(e|h)$, $p(e|\bar{h})$ and the base rate of the hypothesis $br(h)$. Therefore the choice of which logical conditionals to use is less important than the soundness of the belief values assigned to them. The choice of logical conditionals becomes more important when the analyst considers whether the evidence is causal in nature with respect to the hypotheses, or is merely derivative. Section V-B discusses the problem of framing with respect to the causal or derivative nature of evidence and the implications for reasoning.

V. ANALYSIS OF COMPETING HYPOTHESES USING SUBJECTIVE LOGIC (ACH-SL)

Heuer’s ACH process describes a general process that is largely independent of technology and may be used by analysts that have little more than ‘pen and paper’ at their disposal. By contrast, the ACH-SL process is highly dependent on technology to perform Subjective Logic calculations that would be highly complex and time-consuming if they were performed manually. ACH-SL is not merely a theoretical concept – it is also a functioning, implemented DSTC technology known as *ShEBA*³, and provides a framework for the analysis of multiple hypotheses with multiple items of evidence. It was developed to address some of the key analytical issues within the defense, intelligence, and law enforcement communities. This section will outline the ACH-SL process and discuss some of its key features, including:

- compatibility with ‘fuzzy’ human representations of belief;
- interoperability with Bayesian systems;
- formalized abductive and deductive reasoning support; and,
- *a priori* derivation of diagnosticity from analyst judgements.

ACH-SL is not meant as a replacement of ACH, but instead is an elaboration of the basic ACH process that is consistent with Heuer’s intent. Within the original ACH process, Steps 3 through to 6 can be substituted with modified ACH-SL, described below.

² See Section V-F.2.

³*ShEBA* – Structured Evidence-Based Analysis (of hypotheses)

Step-By-Step Outline of Analysis of Competing Hypotheses – Subjective Logic

- 1) Identify the possible hypotheses to be considered. (*ACH Step 1*)
- 2) Make a list of significant evidence and arguments for and against each hypothesis. (*ACH Step 2*)
- 3) Prepare a model consisting of:
 - a) A set of exhaustive and exclusive hypotheses – where one and only one must be true.
 - b) A set of items of evidence that are relevant to one or more hypotheses; are influences that have a causal influence on one or more hypotheses; or, would disconfirm one or more hypotheses.
- 4) Consider the evidence with respect to the hypotheses:
 - a) For each hypothesis and item of evidence, assess its base rate.
 - b) Should the evidence be treated as *causal* or *derivative*? Decide and record for each item of evidence or evidence/hypothesis pair.
 - c) Make judgements for causal evidence as to the likelihood of each hypothesis if the evidence were true and if the evidence were false.
 - d) Make judgements for derivative evidence as to the likelihood that the evidence will be true if the hypothesis were true, and if the hypothesis were false.
 - e) From the judgements provided, compute the diagnosticity for each item of evidence.
- 5) Measure the evidence itself and decide the likelihood that the evidence is true. Supply the measured evidence as input into the constructed model, and use the Subjective Logic calculus to compute the overall likelihood of each hypothesis.
- 6) Analyze how sensitive the conclusion is to a few critical items of evidence. Changes in the value of evidence with high diagnosticity will alter the calculated likelihoods more than evidence with low diagnosticity. Consider the consequences for your analysis if that evidence were wrong, misleading, or subject to a different interpretation.
- 7) Record and report conclusions. Discuss the relative likelihood of all the hypotheses, not just the most likely one. (*ACH Step 7*)
- 8) Identify milestones for future observation that may indicate events are taking a different course than expected. (*ACH Step 8*)

A. Determining base rates of evidence and hypotheses

One of the main problems of applying probability theory and belief calculi to real world problems is determining the base rates for evidence and hypotheses. A distinction can be made between events that can be repeated many times and events that can only happen once. Events that can be repeated many times are frequentist events and the base rates for these can be derived from first principles, or reasonably approximated through empirical observation. For example, if an observer knows the exact proportions of the different colored balls in an urn, then the base rates will be equal to the probabilities of drawing each of the colors. For frequentist problems where base rates cannot be known with absolute certainty, then approximation through prior empirical observation is possible. For events that can only happen once, the observer must arbitrarily decide what the base rates should be, and are often elided as a consequence.

The difference between the concepts of subjective and frequentist probabilities is that the former can

be defined as subjective betting odds – and the latter as the relative frequency of empirically observed data, where the two collapse in the case where empirical data is available [18]. The concepts of *subjective* and *empirical* base rates can be defined in a similar manner where they also converge and merge into a single base rate when empirical data is available.

As an example, consider how a public health department establishes the base rate of some disease within a community. Typically, data is collected from hospitals, clinics and other sources where people diagnosed with the disease are treated. The amount of data that is required to calculate the base rate of the disease will be determined by some departmental guidelines, statistical analysis, and expert opinion about the data that it is truly reflective of the actual number of infections – which is itself a subjective assessment. After the guidelines, analysis and opinion are all satisfied, the base rate will be determined from the data, and can then be used in medical tests to provide a better indication of the likelihood of specific patients having contracted the disease.

1) *Base rates of hypotheses*: As a consequence of the ways in which base rates can be formed, there is an inherent danger in assigning base rates to hypotheses when dealing with events that can only happen once, or when a hypothesis does not have a strong relationship with a well-established model for the approximation of its base rate. Typically, events that have humans as causal influences are poor candidates for empirical base rates since they are highly contextual and may be subject to slight variation which cause large perturbations in their appearance. In addition, much of strategic intelligence deals with hypotheses that can only happen once and for which empirical data that can be used to approximate base rates simply does not exist. In these cases, the base rates for non-repeatable hypotheses should be evenly weighted since they form an exhaustive and exclusive state space – one and only one of the hypotheses is true. So, for a set of k hypotheses $\Phi = \{h_1, h_2, \dots, h_k\}$, the base rate of each hypothesis should be $\frac{1}{k}$, i.e.

$$\forall h_i \in \Phi, \text{br}(h_i) = \frac{1}{k} \quad (\text{V.1})$$

For three hypotheses, the base rate of each hypothesis should be $\frac{1}{3}$; for four hypotheses, the base rate should be $\frac{1}{4}$, and so on. This follows from the *Principle of Indifference* which states that if we are ignorant of the ways an event can occur, the event will occur equally likely in any way [19].

Setting the base rates to other than equal values for non-repeatable events is strongly discouraged. Doing so introduces an inherent bias in the model and may produce erroneous conclusions. Any direct reasoning about base rates under these conditions should be discarded in favor of consideration of the evidence. If reasoning is applied to subjectively determine base rates, then that reasoning needs to be explicitly stated in the model instead of implicitly included in the hypotheses' base rates. The evidence for the reasoning should be included as a standard part of the model and treated like all other evidence. Doing so reduces the likelihood of erroneous conclusions and eliminates the possibility of 'double counting' evidence – where the evidence has already been taken into account in setting the base rate but is also used as part of the model to reason about the hypotheses.

2) *Base rates of evidence*: Base rates for evidence should be considered in the same way as for hypotheses. The set of hypotheses form an exhaustive and exclusive state space, such that one and only one hypothesis is true. Similarly for each item of evidence, consideration must be given to the other elements of the state space in which the evidence is situated. When the base rate for a item of evidence can not be derived from first principles or approximated through empirical testing, then the base rate should be set according to proportion of the state space that the evidence consumes.

For example, the blood type being found at a crime scene to be AB might be considered evidence for or against certain competing hypotheses. The complete state space in which the evidence "Blood sample is type AB " allows for three other possibilities, namely that the blood type is A , B , or O . If the prevalence of the four different blood types within the community was known as a result of analysis of statistical

data, then the empirically approximated base rates should be used. If they are not known or cannot be reliably derived, then the base rate for a result of blood type AB should be set at $\frac{1}{4}$.

Typically, most evidence will have a base rate of $\frac{1}{2}$ when dealing with simple true/false statements. However, consideration of the complete state space in which the evidence is situated is important so as not to introduce base rate errors into the model.

B. Causal and Derivative Evidence

Abduction reasons about the likelihood of the hypothesis, given the likelihood of the assertion of the evidence under conditions when the hypothesis is true and false, i.e. using $p(e|h)$, $p(e|\bar{h})$. Physicians primarily use abduction for medical diagnosis when considering the likelihood of a patient having an particular disease, given that the patient presents specific symptoms that are associated with the disease.

By contrast, *deduction* directly reasons about the likelihood of the hypothesis, given the likelihoods of the hypothesis under conditions of the assertion and negation of the evidence, i.e. using $p(h|e)$, $p(h|\bar{e})$. Deduction is most often applied when there can be a causal link from the evidence to one or more hypotheses, such as when reasoning about the likelihood of a patient having a particular disease, given the possibility of recent exposure to the same infectious disease.

The original ACH describes a process that uses deductive reasoning [7], while Stech and Elässer explicitly use an abductive approach to reasoning with ACH-CD [12].

Both deductive and abductive reasoning have their uses and limitations. Deductive reasoning is best suited for reasoning about causal evidence, while abductive reasoning is best suited for reasoning about derivative evidence.

Causal evidence has a direct causal influence on a hypothesis – such as the presence of a persistent low pressure system is causal evidence for rain, since a low pressure system has a direct influence on precipitation. The ‘state of mind’ of an adversary is often regarded as causal evidence since it usually has a direct influence on their decision making processes.

Derivative evidence [20] – also known as *diagnostic evidence* [21] – is indirect secondary evidence – not causal in nature – and is usually observed in conjunction, or is closely associated with the hypothesis. For example, a soggy lawn should be considered derivative evidence for rain – but soggy lawns are also associated with the use of sprinklers, and recently-washed automobiles. In the nuclear enrichment example from Stech and Elässer (Section IV), the detection of Krypton gas would be considered derivative evidence, since the presence of Krypton gas does not causally influence the likelihood of a nuclear enrichment program.

In theory, both deductive and abductive reasoning can be used for analysis of competing hypotheses, providing the logical conditionals have suitable belief assignments. In practice though, there is an inherent danger in applying deductive reasoning to derivative evidence, just as there is a danger in applying abductive reasoning to causal evidence. The problem lies in how the questions about the evidence and the hypothesis are framed, and the nature of causality that can be inferred [20], [4], [21].

1) *Reasoning about causal evidence*: Using abductive reasoning to reason about causal evidence requires more cognitive effort than using deductive reasoning. It requires the analyst to suppose the assertion or negation of the consequent and reason about the likelihood of the antecedent, $p(e|h)$ and $p(e|\bar{h})$. At best, it is likely that the analyst is *actually* reasoning about the likelihood of the consequent given the assertion and negation of the antecedent, $p(h|e)$ and $p(h|\bar{e})$, and simply approximating $p(e|h)$ and $p(e|\bar{h})$ as a result. At worst, the analyst can draw completely different inferences that violate the causal nature of the evidence and lead to incorrect reasoning about the hypotheses.

For example, during the Cuban missile crisis of 1962, President Kennedy publicly warned that the United States would view any Soviet strategic missile placements in Cuba as a grave threat and would take appropriate action [20]. An abductive approach to this problem would require the analyst to ask of themselves:

$p(e|h)$ If the Soviets are shipping strategic missiles to Cuba, what is the likelihood that the President made a public statement that the U.S. would perceive strategic missile placement as a threat?

$p(e|\bar{h})$ If the Soviets are not shipping strategic missiles to Cuba, what is the likelihood that the President made a public statement that the U.S. would perceive strategic missile placement as a threat?

The way the question is framed suggests that the analyst should consider how likely there will be a public statement given the existence of strategic missiles in Cuba. If the analyst was unaware of the likely intention of the public statement, they might reasonably conclude that it is more likely that a public statement would be made if the Soviets are in the process of shipping strategic missiles to Cuba, thus increasing the likelihood of the ‘missiles’ hypothesis. However, this is almost certainly not as President Kennedy intended, instead likely reasoning that a public statement would serve to dissuade the Soviet leadership from shipping strategic missiles to Cuba – or at worst have no appreciable effect. In other words, he reasoned that the public statement would act in a causal manner to increase the likelihood of the ‘no missiles’ hypothesis [20].

If we apply a deductive approach to the same question, the analyst is required to ask of themselves different questions that preserve the causal nature of the evidence and appear less likely to facilitate the same framing errors:

$p(h|e)$ If the President made a public statement that the U.S. would perceive strategic missile placement as a threat, what is the likelihood that the Soviets are shipping strategic missiles to Cuba?

$p(h|\bar{e})$ If the President did not make a public statement that the U.S. would perceive strategic missile placement as a threat, what is the likelihood that the Soviets are shipping strategic missiles to Cuba?

Here the framing of the question suggests that the statement will influence the outcome, and one would probably conclude that the statement will serve to lessen the likelihood of missiles being shipped to Cuba – all other things being equal. This effect is similar to the paradoxical probability assessments problem of logically-equivalent pairs of conditional propositions, discussed by Tversky and Kahneman [21].

2) *Reasoning about derivative evidence:* The same paradox holds true for deductive reasoning about derivative evidence. Deductive reasoning requires the analyst to ask questions in a way that implies that the evidence is causal in nature. Deductive reasoning about derivative evidence also requires more cognitive effort than abductive reasoning, and can also lead to incorrect conclusions.

For example, consider the questions that an analyst might ask themselves if deductive reasoning were applied to the nuclear enrichment hypothesis (Section IV).

$p(h|e)$ If Krypton gas is detected in Iraq, what is the likelihood that the Iraqis have a nuclear enrichment program?

$p(h|\bar{e})$ If Krypton gas is not detected in Iraq, what is the likelihood that the Iraqis have a nuclear enrichment program?

The framing of the question does not prompt the analyst to consider the likelihood of Krypton gas when there is no nuclear enrichment program, i.e. $p(e|\bar{h})$. If the analyst is unaware that Krypton gas is also used for detecting leaks in oil pipelines, they will likely erroneously conclude that the likelihood of $p(h|e)$ is high. The analyst might reasonably infer from the framing of these questions that $p(e|h)$ may be a good approximation for $p(h|e)$ since Krypton gas is a by-product of nuclear enrichment – which may cause them to miss the fact that Krypton gas is merely diagnostic – and incorrectly conclude that its presence implies a nuclear enrichment program. When the problem is framed using abductive reasoning,

the analyst is prompted to consider the likelihood of Krypton gas in cases where there would be no nuclear enrichment program and the framing problems disappear.

$p(e|h)$ If the Iraqis have a nuclear enrichment program, what is the likelihood that Krypton gas will be detected?

$p(e|\bar{h})$ If the Iraqis do not have a nuclear enrichment program, what is the likelihood that Krypton gas will be detected?

Even though humans routinely attribute causality where none exists [21], framing the question to preserve the derivative nature of the evidence makes it less likely that the analyst will misattribute causality. Similar conclusions may be reached as for the use of abduction for causal evidence (Section V-B.1). While it is possible to use either style of reasoning for both types of evidence, less cognitive effort is required when deductive reasoning is used for causal evidence and abductive reasoning is used for derivative evidence, and therefore erroneous reasoning due to framing is less likely to occur.

C. Constructing the analytic model

Logical conditionals [22] are a pair of conditional beliefs that is used for reasoning about the likelihood of a hypothesis, given some evidence. The values of these conditionals constitute the judgements that the analyst supplies as part of the ACH-SL process to reason about each hypothesis in respect of each item of evidence.

Assigning belief values to the conditionals requires that the analyst answer certain questions of themselves, and probably others, including experts. The style of reasoning that is used for the evidence and hypothesis will largely determine the type of questions to be answered, which in turn will be strongly influenced by the causal or derivative nature of the evidence with respect to the hypothesis (see Section V-B). For *deductive* reasoning, the questions should use the following or an equivalent form:

$p(h|e)$ If [*the evidence is true*], what is the likelihood that [*the hypothesis is true*]?

$p(h|\bar{e})$ If [*the evidence is false*], what is the likelihood that [*the hypothesis is true*]?

By contrast, *abductive* reasoning should use questions in the following or an equivalent form:

$p(e|h)$ If [*the hypothesis is true*], what is the likelihood that [*the evidence is true*]?

$p(e|\bar{h})$ If [*the hypothesis is false*], what is the likelihood that [*the evidence is true*]?

For each question, the analyst must assign a value while considering *only* the evidence and the hypothesis to which it relates. They must assume when providing a judgement that no other information is known about other hypotheses or other items of evidence, other than the base assumptions on which the model is predicated. This is done for two reasons. Firstly, as a human being, the analyst will be incapable of weighing more than four factors simultaneously in supplying a judgement [14] – even though that they may believe that they are considering dozens [7]. Secondly, the analyst runs the risk of introducing bias if each judgement is not considered in isolation – which will also bias the overall conclusions about the likelihoods of the hypotheses and make it difficult for others, including policy makers, to understand the reasoning employed by the analyst.

1) *Constraints on logical conditionals*: The two types of logical conditionals used in inductive reasoning have different mathematical requirements placed on their values as a result. The logical conditionals used for deduction are termed *deductive logical conditionals*, i.e. $p(h_i|e)$ and $p(h_i|\bar{e})$. These conditionals must obey certain constraints in order to satisfy basic requirements of probability theory. For an exhaustive and exclusive set of hypotheses where one and only one hypothesis can be true, it logically follows that the sum of the probability expectations of the positive conditionals must be one; and similarly the probability

expectations of the negative conditionals must also sum to one, since

$$\begin{aligned} p(h_1) + p(h_2) + \dots + p(h_n) &= 1 \\ p(\bar{h}_1) + p(\bar{h}_2) + \dots + p(\bar{h}_n) &= 1 \end{aligned} \quad (\text{V.2})$$

The logical conditionals used for abduction are termed *abductive logical conditionals*, i.e. $p(e|h_i)$ and $p(e|\bar{h}_i)$. These have different constraints to deductive logical conditionals and there is no requirement that the probability expectations of their positive conditionals sum to one since the items of evidence are in separate state spaces. The probability expectation of the positive conditional $p(e|h_i)$ is unconstrained for each hypothesis. However, the probability expectation of each negative conditional $p(e|\bar{h}_i)$ must be equal to the average of the probability expectations of the other positive conditionals of the other hypotheses, i.e.

$$p(e|\bar{h}_1) = \frac{p(e|h_2) + p(e|h_3) + \dots + p(e|h_n)}{n - 1} \quad (\text{V.3})$$

Details on the constraints of logical conditionals are described in the *Appendix*.

2) *Model coherence*: Across multiple items of evidence, there are further constraints to ensure that the model is coherent. The requirements for coherence concern the assignment of belief values for different items of evidence for the same hypothesis. The constraints apply only to deductive logical conditionals $p(h_i|e_j)$ and $p(h_i|\bar{e}_j)$ which are either explicitly provided by the analyst when deductive reasoning is used, or are implicitly derived from knowledge of the abductive logical conditionals and base rate of the hypothesis⁴.

The belief values assigned to the deductive logical conditionals of different items of evidence for the same hypothesis must have overlapping or adjoining ranges for the model to be coherent. If the ranges of belief assigned to the logical conditionals do not overlap or join, then at least one of the logical conditionals of an item of evidence must be incorrect. Put simply, model coherence requires that the minimum likelihood of the hypothesis for one item of evidence can not be greater than the maximum likelihood for another item of evidence, since both judgements were supplied independently. If the minimum for one item of evidence were to be greater than the maximum for another item of evidence, then this indicates that at least one of the judgements is incorrect since it is a logical and a physical impossibility for non-quantum events. Model coherence is described further in the *Appendix*.

D. Subjective Logic Fundamentals

This section introduces Subjective Logic, which is extensively used within the ACH-SL approach to model the influence of evidence on hypotheses, and provide a calculus for the evaluation of the model when measurement of the evidence is provided.

Belief theory is a framework related to probability theory, but where the sum of probabilities over all possible outcomes not necessarily add up to 1, and the remaining probability is assigned to the union of possible outcomes. Belief calculus is suitable for approximate reasoning in situations where there is more or less uncertainty about whether a given proposition is true or false, and is ideally suited for both human and machine representations of belief.

Subjective logic[11] represents a specific belief calculus that uses a belief metric called *opinion* to express beliefs. An opinion denoted by $\omega_x^A = (b_x^A, d_x^A, u_x^A, a_x^A)$ expresses the relying party A 's belief in

⁴See Section V-F.2 for details on transforming abductive logical conditionals to deductive logical conditionals.

the truth of statement x . Here b , d , and u 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:

$$b_x^A + d_x^A + u_x^A = 1. \quad (\text{V.4})$$

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 taken⁵. 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:

$$E(\omega_x^A) = b_x^A + a_x^A u_x^A, \quad (\text{V.5})$$

meaning that a determines how uncertainty shall contribute to $E(\omega_x^A)$. When the statement x for example says "*Party B is honest and reliable*" then the opinion can be interpreted as trust in B , which can also be denoted as ω_B^A .

The opinion space can be mapped into the interior of an equal-sided triangle, where, for an opinion $\omega_x = (b_x, d_x, u_x, a_x)$, the three parameters b_x , d_x and u_x determine the position of the point in the triangle representing the opinion. Fig.1 illustrates an example where the opinion about a proposition x from a binary frame of discernment has the value $\omega_x = (0.7, 0.1, 0.2, 0.5)$.

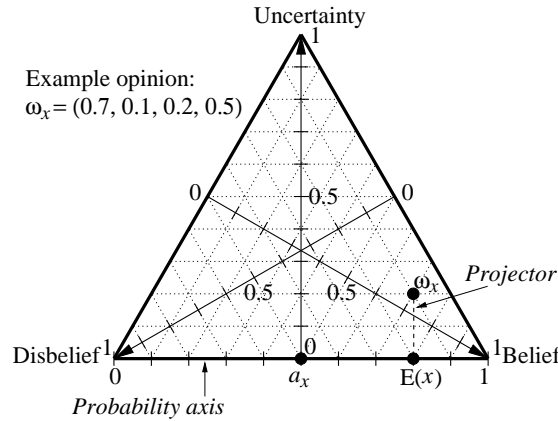


Fig. 1. Opinion triangle with example opinion

The top vertex of the triangle represents uncertainty, the bottom left vertex represents disbelief, and the bottom right vertex represents belief. The parameter b_x is the value of a linear function on the triangle which takes value 0 on the edge which joins the uncertainty and disbelief vertices and takes value 1 at the belief vertex. In other words, b_x is equal to the quotient when the perpendicular distance between the opinion point and the edge joining the uncertainty and disbelief vertices is divided by the perpendicular distance between the belief vertex and the same edge. The parameters d_x and u_x are determined similarly. The edge joining the disbelief and belief vertices is called the probability axis. The relative atomicity is indicated by a point on the probability axis, and the projector starting from the opinion point is parallel to the line that joins the uncertainty vertex and the relative atomicity point on the probability axis. The point at which the projector meets the probability axis determines the expectation value of the opinion, *i.e.* it coincides with the point corresponding to expectation value $b_x + a_x u_x$.

Opinions can be ordered according to probability expectation value, but additional criteria are needed in case of equal probability expectation values. We will use the following rules to determine the order of opinions[11]:

⁵See Section V-A for a discussion on the assignment of base rates.

Let ω_x and ω_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.

Opinions can be expressed as beta PDFs (probability density functions). The beta-family of distributions is a continuous family of distribution functions indexed by the two parameters α and β . The beta PDF denoted by $\text{beta}(\alpha, \beta)$ can be expressed using the gamma function Γ as:

$$\text{beta}(\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1} \quad (\text{V.6})$$

where $0 \leq p \leq 1$ and $\alpha, \beta > 0$, with the restriction that the probability variable $p \neq 0$ if $\alpha < 1$, and $p \neq 1$ if $\beta < 1$. The probability expectation value of the beta distribution is given by:

$$E(p) = \alpha / (\alpha + \beta). \quad (\text{V.7})$$

The following mapping defines how opinions can be represented as beta PDFs.

$$(b_x, d_x, u_x, a_x) \mapsto \text{beta} \left(\frac{2b_x}{u_x} + 2a_x, \frac{2d_x}{u_x} + 2(1 - a_x) \right). \quad (\text{V.8})$$

This means for example that an opinion with $u_x = 1$ and $a_x = 0.5$ which maps to $\text{beta}(1, 1)$ is equivalent to a uniform PDF. It also means that a dogmatic opinion with $u_x = 0$ which maps to $\text{beta}(b_x \eta, d_x \eta)$ where $\eta \rightarrow \infty$ is equivalent to a spike PDF with infinitesimal width and infinite height. Dogmatic opinions can thus be interpreted as being based on an infinite amount of evidence.

When nothing is known, the *a priori* distribution is the uniform beta with $\alpha = 1$ and $\beta = 1$ illustrated in Fig.2a. Then after r positive and s negative observations the *a posteriori* distribution is the beta PDF with the parameters $\alpha = r + 1$ and $\beta = s + 1$. For example the beta PDF after observing 7 positive and 1 negative outcomes is illustrated in Fig.2b below. This corresponds to the opinion of Fig.1 through the mapping of (V.8).

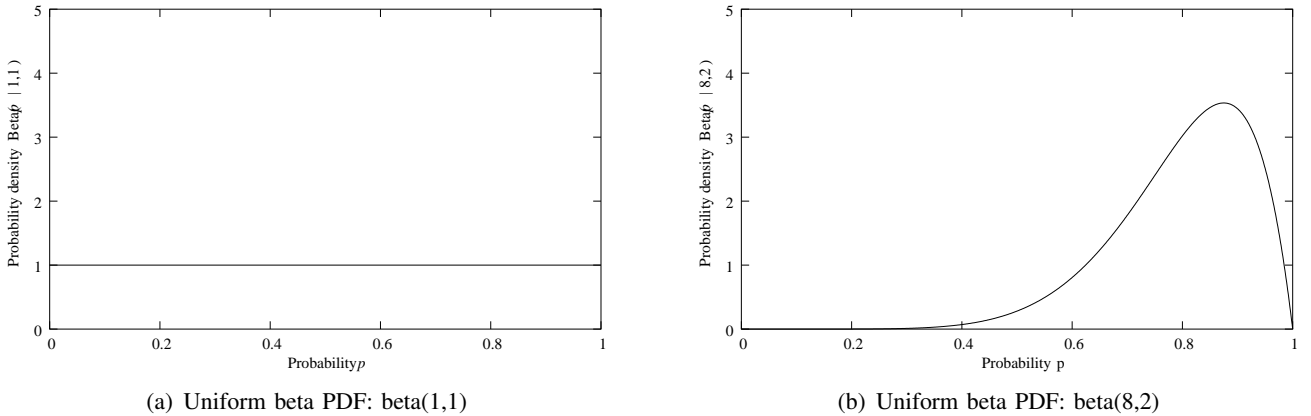


Fig. 2. Uniform beta PDF examples

A PDF of this type expresses the uncertain probability that a process will produce positive outcome during future observations. The probability expectation value of Fig.2b. is $E(p) = 0.8$. This can be interpreted as saying that the relative frequency of a positive outcome in the future is somewhat uncertain, and that the most likely value is 0.8.

The variable p in (V.6) is a probability variable, so that for a given p the probability density $\text{beta}(\alpha, \beta)$ represents second order probability. The first-order variable p represents the probability of an event, whereas the density $\text{beta}(\alpha, \beta)$ represents the probability that the first-order variable has a specific value.

By definition, the PDF and the corresponding opinion always have the same probability expectation value, and can be interpreted as equivalent. This makes it possible to fuse opinions using Bayesian updating of beta PDFs.

E. Representations of Subjective Opinions

Opinions that are expressed in Subjective Logic can be transformed to and from other belief representations and can be visualized in a number of different ways. This section describes techniques for visualization; mapping between verbal fuzzy adjectives and Subjective Logic opinions; and transformation opinions to and from Bayesian representations of belief.

Subjective Logic opinions, *Subjective Opinions*, can be transformed without loss to and from Bayesian belief representations⁶. This makes the Subjective Logic calculus ideal for reasoning about machine-supplied data that correspond to, or can be interpreted as Bayesian representations of belief. The immediate implications of this are that systems that use Subjective Logic – such DSTC’s *ShEBA* technology – are able to be interfaced with systems that use Bayesian representations of belief to provide data to other Bayesian systems, or take data from Bayesian systems to be used as input into an ACH-SL system.

In the earlier sections we have shown that bipolar beliefs in the form of opinions, as illustrated in Fig.1, can be mapped to and interpreted as beta PDFs. While these two graphical representations give a concise mathematical visualization of bipolar beliefs, people unfamiliar with the underlying mathematical concepts can have difficulty interpreting them. For this reason, more intuitive graphical and verbal representations can be used. This is shown in Fig.3, which is a screen capture of an online demonstration⁷.

The example visualizes bipolar beliefs about three different statements x , y and z . Each belief is visualized in different ways, i.e. in the form of 1) points in an opinion triangle, 2) beta density functions, 3) coloured/shaded bars, and 4) fuzzy verbal categories. The interpretation of the opinion triangle and the beta PDF need no further explanation, as they have been described in the previous sections.

The horizontal shaded bars are actually colored in the online demonstration, which makes them easier to interpret. The first horizontal bar, representing the belief in x , consists of a dark shaded area representing b_x , and a light shaded area representing $a_x u_x$ – i.e. the amount of uncertainty that contributes to $E(x)$ – so that the total length of the dark and light shaded areas together represent $E(x)$. The second horizontal bar, representing the belief in y , consists of a green (leftmost) area representing b_y , an amber (middle) area representing u_y , and a red (rightmost) area representing d_y , as well as a black vertical line within the amber area indicating $E(y)$. This uses a ‘traffic light’ metaphor, where green indicates “go”, red indicates “stop” and amber indicates “caution”. The third horizontal bar, representing the belief in z , simply has a single dark shaded area representing $E(z)$.

A fuzzy category is also indicated directly above each horizontal bar. The fuzzy verbal categories can be defined according to the need of the application. The example of Fig.3 (p.15) uses categories from the two-dimensional matrix defined in Fig.4 (p.16).

Note that the certainty category “Certain” is implicit, and need not be mentioned together with the applicable likelihood category. These fuzzy verbal categories can be mapped to areas in the opinion triangle as illustrated in Fig.5 (p.16). The mapping must be defined for combinations of ranges of expectation value and uncertainty. As a result, the mapping between a specific fuzzy category from Fig.4 (p.16) and specific geometric area in the opinion triangle depends on the base rate. Without specifying the exact underlying ranges, the visualization of Fig.3 and Fig.5 indicate the ranges approximately. The edge ranges

⁶See the *Appendix*.

⁷<http://security.dstc.edu.au/spectrum/beliefvisual/BVDDemo.html>

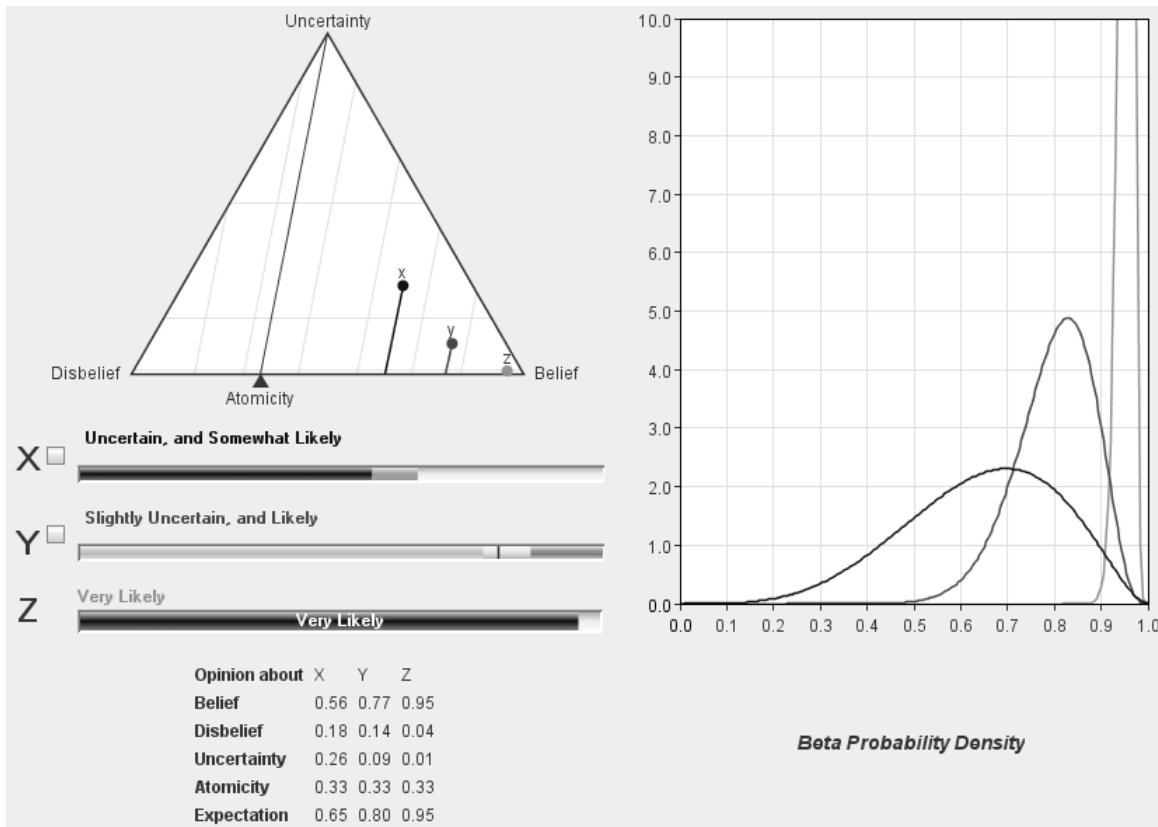


Fig. 3. Example visualizations of bipolar beliefs

are deliberately made narrow in order to have categories for near dogmatic and vacuous beliefs, as well as beliefs that express expectation values near absolute 0 or 1. The number of likelihood categories, and certainty categories, as well as the exact ranges for each, must be determined according to the need of each application, and the fuzzy categories defined here must be seen as an example. Real-world categories would likely be similar to those found in Sherman Kent's *Words of Estimated Probability* [23]; based on the *Admiralty Scale* as used within the UK National Intelligence Model⁸; or could be based on empirical results obtained from psychological experimentation.

Fig.5.a illustrates the case with base rate $a = \frac{1}{3}$, which was also the case in the visualization in Fig.3. Whenever a fuzzy category area overlaps, partly or completely, with the opinion triangle, that fuzzy category is possible.

The possible mappings depend on the base rate. For example, it can be seen that the category 7D: “Unlikely and Very Uncertain” is possible in case $a = \frac{1}{3}$, but not in case $a = \frac{2}{3}$. This is because the expectation of a state x is defined as $E(x) = b_x + a_x u_x$, so that when $a_x, u_x \rightarrow 1$, then $E(x) \rightarrow 1$, so that the likelihood category “Unlikely” would be impossible.

Mapping from fuzzy categories to Subjective Opinions is also straight-forward. Geometrically, the process involves mapping the fuzzy adjectives to the corresponding center of the portion of the grid cell contained within the opinion triangle (see Fig.5). Naturally, some mappings will always be impossible for a given base rate, but these are logically inconsistent and should be excluded from selection.

It is interesting to notice that although a specific fuzzy category maps to different geometric areas in the opinion triangle depending on the base rate, it will always correspond to the same range of beta PDFs.

⁸<http://www.policereform.gov.uk/implementation/natintellmodel.html>

		Likelihood Categories								
		Absolutely Not	Very Unlikely	Unlikely	Somewhat Unlikely	Chances about even	Somewhat Likely	Likely	Very Likely	Absolutely
Certainty Categories		9	8	7	6	5	4	3	2	1
Completely Uncertain	E	9E	8E	7E	6E	5E	4E	3E	2E	1E
Very Uncertain	D	9D	8D	7D	6D	5D	4D	3D	2D	1D
Uncertain	C	9C	8C	7C	6C	5C	4C	3C	2C	1C
Slightly Uncertain	B	9B	8B	7B	6B	5B	4B	3B	2B	1B
Completely Certain	A	9A	8A	7A	6A	5A	4A	3A	2A	1A

Fig. 4. Fuzzy Categories

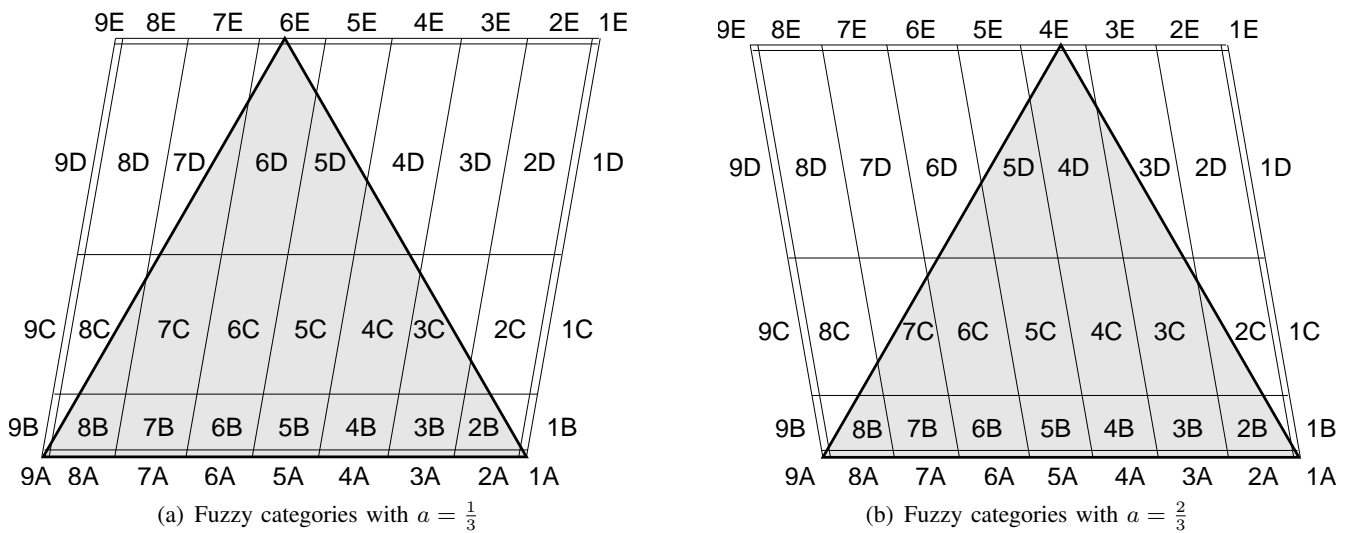


Fig. 5. Mapping of fuzzy categories to ranges of belief as a function of the base rate

It is simple to visualize ranges of bipolar beliefs with the opinion triangle, but it would not be easy to visualize ranges of beta PDFs. The mapping between bipolar beliefs and beta PDFs thereby provides a very powerful way of describing PDFs in terms of fuzzy categories, and vice versa.

F. Calculating opinions about the hypotheses

The likelihood of each hypothesis can be calculated from knowledge of the base rate for each hypothesis $br(h_i)$; the base rate for each item of evidence $br(e_j)$; the logical conditionals for each hypothesis/evidence pair (h_i, e_j) ; and measurements/opinions about the evidence $p(e_i)$. Two distinct steps are required:

- 1) The likelihood of each hypothesis, h_i , for each item of evidence, e_i is inferred using either *abduction* or *deduction*, depending on whether abductive or deductive logical conditionals that are provided. This produces knowledge of $p(h_i)$ for each e_j , i.e. $p(h_i|e_j)$.

- 2) The overall likelihood for each item of evidence, $p(h_i)$ for each h_i , is computed by fusing the separate $p(h_i|e_j)$ opinions using the *consensus* operator.

This section discusses these three basic Subjective Logic operators that are used within the ACH-SL system for inferring the likelihoods of the hypotheses from what is known about the evidence.

1) *Deduction*: Deduction reasons about the likelihood of a hypothesis, given knowledge of the likelihood of the hypothesis being true when some evidence is observed; and the likelihood of the hypothesis being true when the evidence is not observed (see Fig.6).

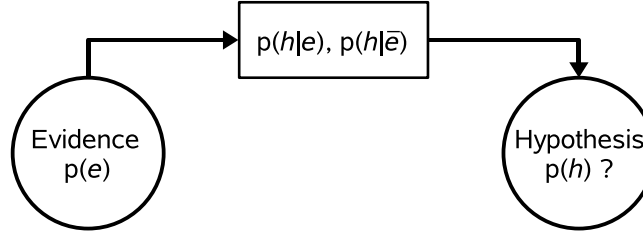


Fig. 6. Deductive reasoning from the evidence to the hypothesis.

Deductive reasoning is often used for reasoning when the appearance of the evidence temporally precedes the appearance of the hypothesis, where there is a causal link from the evidence to the hypothesis. As an example, suppose that a physician working in an area with a high rate of disease as the result of some natural disaster is trying to determine the likelihood of a patient having contracted an infectious disease. The patient has no visible signs or reported symptoms, but the physician knows that direct contact with a carrier of the disease results in an infection 95 percent (0.95) of the time. However, of those who are infected, about 10 percent (0.1) have contracted the disease without contact with disease carriers. If the likelihood of contact with a disease carrier for a particular patient is 10 percent (0.1), then the likelihood of infection is approximately 19 percent (0.19).

The details of the deduction operator, \odot , are described in [17]. The operator is written as $\omega_{h|e} = \omega_e \odot (\omega_{h|e}, \omega_{h|\bar{e}})$.

2) *Abduction*: Abduction reasons about the likelihood of a hypothesis, given knowledge of the likelihood of some evidence being observed when the hypothesis is true; the likelihood of the evidence being observed when the hypothesis is false; and the base rate of the hypothesis.

Abductive reasoning is often applied to interpret medical test results. As an example, suppose that a specific disease has a base rate of $\frac{1}{100}$ (i.e. one in every hundred people on average has the disease). A particular test for this disease has a false positive rate of $\frac{1}{100}$ and a false negative rate of $\frac{1}{50}$ (i.e. $\text{br}(h) = 0.01$, $p(e|h) = 0.99$, $p(e|\bar{h}) = 0.02$). The false negative rate means for two percent of those who have the disease and are tested, the test will erroneously report that they do not have the disease. Similarly, the false positive rate means that for one percent of those who do not have the disease and are tested, the test will erroneously report that they do have the disease. Under all other conditions, the test reports the results correctly. If the test is applied to a random person for whom it is not known if they have the disease, and the result is positive, then the likelihood that they actually have the disease is $\frac{1}{3}$ or 0.33 – and not 0.98 as might have been supposed if the base rate was ignored.

Deriving the likelihood of the hypothesis from the logical conditionals $p(h|e)$ and $p(h|\bar{e})$ and the likelihood of the evidence $p(e)$ is straight-forward. Likewise, deriving the likelihood of the evidence from the logical conditionals $p(e|h)$ and $p(e|\bar{h})$ and the likelihood of the hypothesis $p(h)$ is also straight-forward. However, deriving the likelihood of the hypothesis with only the logical conditionals, $p(e|h)$ and

$p(e|\bar{h})$, is not possible without knowledge of the base rate of the hypothesis $br(h)$ (see Fig.7).⁹

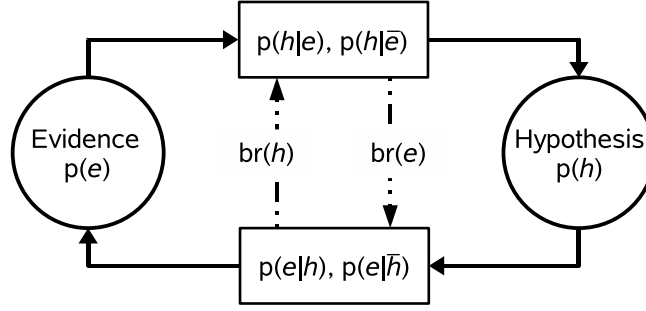


Fig. 7. Different conditionals are needed for reasoning in different directions.

Using the knowledge of these three pieces of information, the logical conditionals $p(h|e)$ and $p(h|\bar{e})$ can be derived, and the problem can be solved using deduction (see Fig.8).

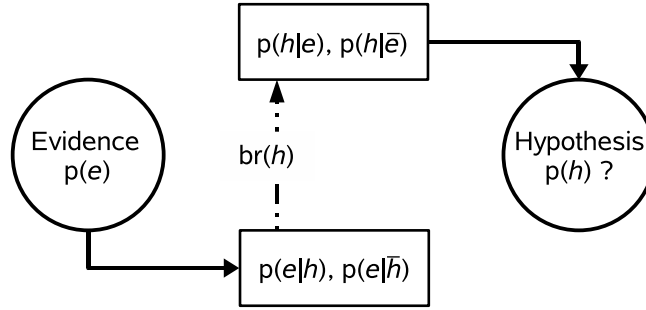


Fig. 8. Abductive reasoning from the evidence to the hypothesis.

Using Subjective Logic, the implicit logical conditionals that allow direct reasoning from the evidence to the hypothesis are derived and used with the deduction operator (Section V-F.1) to obtain the correct result. The logical conditionals used for deducing the hypothesis from knowledge of the evidence, $\omega_{e|h}$ and $\omega_{e|\bar{h}}$, can be derived from knowledge of the supplied conditionals, $\omega_{h|e}$ and $\omega_{h|\bar{e}}$, and knowledge of the base rate of the hypothesis, $br(h)$. While the abduction operator will be discussed in depth in a forthcoming paper, the general derivation of deductive logical conditionals from abductive logical conditionals is summarized below.¹⁰

Definition 5.1 (Abduction): Given knowledge of the base rate of the hypothesis $br(h)$ where the ω_h is a vacuous subjective opinion about the base rate of the hypothesis, defined as

$$\omega_h = (b_h, d_h, u_h, a_h) \begin{cases} b_h = 0 \\ d_h = 0 \\ u_h = 1 \\ a_h = br(h) \end{cases} \quad (\text{V.9})$$

and given the abductive logical conditionals about evidence e , expressed in Subjective Logic form $\omega_{h|e}, \omega_{h|\bar{e}}$, then the deductive logical conditionals $\omega_{e|h}, \omega_{e|\bar{h}}$ are derived using the following formula

⁹Similarly, with knowledge of the base rate of evidence $br(e)$ and $p(h|e)$ and $p(h|\bar{e})$, the conditionals, $p(e|h)$ and $p(e|\bar{h})$, could be derived.

¹⁰The authors wish to acknowledge the important contributions made by David McAnally (dsm@maths.uq.edu.au) in developing the abduction operator.

$$\begin{aligned}\omega_{h|e} &= \frac{\omega_h \cdot \omega_{e|h}}{\omega_h \odot (\omega_{e|h}, \omega_{e|\bar{h}})} \\ \omega_{h|\bar{e}} &= \frac{\omega_h \cdot \neg \omega_{e|h}}{\omega_h \odot (\neg \omega_{e|h}, \neg \omega_{e|\bar{h}})}\end{aligned}\tag{V.10}$$

and ω_h can be solved using the deduction operator (Section V-F.1) where

$$\omega_{h||e} = \omega_e \odot (\omega_{h|e}, \omega_{h|\bar{e}})\tag{V.11}$$

The abduction operator, $\bar{\odot}$, is written as $\omega_{h||e} = \omega_e \bar{\odot} (\omega_{e|h}, \omega_{e|\bar{h}}, \text{br}(h))$. Details on the multiplication and division operators can be found in [24].

3) *Consensus*: The consensus operator is used for belief fusion. It allows independent beliefs to be combined into a consensus opinion which reflects all opinions in a fair and equal way, i.e. when there are two or more beliefs about hypothesis h resulting from distinct items of evidence, the consensus operator produces a consensus belief that combines them into one.

For example, suppose that for the hypothesis, h , there two distinct items of evidence, e_1 and e_2 , that are either causal or derivative with respect to the hypothesis. Assume that for each of item of evidence some inference is drawn about the likelihood of h (using deduction or abduction), then the two independent opinions, $\omega_h^{e_1}$ and $\omega_h^{e_2}$, can be fused into a single consensus opinion ω_h which provides an overall assessment of the likelihood of the hypothesis.

The details of the consensus operator, \oplus , are described in [25] and discussed further in [26]. The operator is written as $\omega_h = \omega_h^{e_1} \oplus \omega_h^{e_2} \cdots \oplus \omega_h^{e_n}$.

G. Diagnosticity of Evidence

Not all evidence is created equal – some evidence is better for distinguishing between hypotheses than others. Evidence is considered to be diagnostic when it is indicative of the relative likelihood of the hypotheses being considered. If a item of evidence seems consistent with all the hypotheses, it will generally have little diagnostic value.

Heuer's *Analysis of Competing Hypotheses* [7] describes a process by which diagnosticity is indicated by the analyst for each evidence-hypothesis pair. Under a fairly narrow interpretation, such as given by Stech and Elässer [12], the ACH process appears to consider only the assertion of each item of evidence $h|e$ (i.e. the evidence is true), and does not consider its negation $h|\bar{e}$ (i.e. the evidence is false). Under a broader interpretation, consideration of both logical conditionals is implied.

Consider Heuer's medical analogy as an illustration of this point. A high temperature, e , might have little diagnostic value in determining which illness a person is suffering from – yet the absence of a high temperature, \bar{e} , may be more significant for distinguishing between possible illnesses.

In the ACH process, diagnosticity is explicitly provided by the analyst as an input [7], and it is used both to eliminate evidence from the model that does not distinguish well between hypotheses, and to provide a means of eliminating hypotheses based on the relative weight of disconfirming evidence. These inputs are 'second-order' judgements, since the analyst must first consider the relationship between the item of evidence and the hypotheses in order to determine the diagnosticity of the item of evidence. This reasoning is usually hidden from the final analysis and may be subject to the cognitive limitations and biases which significantly contribute to errors in reasoning.

In the modified ACH-SL system, diagnosticity is *not* explicitly provided by the analyst. Instead, it is derived from the ‘first-order’ values that the analyst assigns to the logical conditionals, independently of the actual value of the evidence. This allows analysts to concentrate on the judgements they make, rather than requiring them to consider diagnosticity as a separate, ‘second-order’ measure of the evidence.

Diagnosticity is represented as a real number between 0 and 1 – with a value of 0 indicating that the evidence does not distinguish between the hypotheses in any way; and with a value of 1 indicating that the evidence is capable of completely distinguishing between the hypotheses.

Diagnosticity can also be derived for any subset of hypotheses and provides the analyst with detail as to how well the evidence distinguishes between the members of the subset. For example, the overall diagnosticity of a item of evidence may be poor in distinguishing between a set of six hypotheses, yet it may be very good at distinguishing between just two of those six.

Diagnosticity is derived using the logical conditionals $p(h|e)$ and $p(h|\bar{e})$. If these conditionals are not known, then they can be derived from knowledge of the $p(e|h)$ and $p(e|\bar{h})$, and from the base rate of the hypothesis $br(h)$ (see Section V-F.2). Details on how diagnosticity is derived, including examples, are described in the *Appendix*.

VI. CONCLUSION

The approach of Analysis of Competing Hypotheses using Subjective Logic (ACH-SL) has been developed by the authors to address some of the key analytical issues within the defense, intelligence, and law enforcement communities. ACH-SL is not meant as a replacement of ACH, but instead is an elaboration of the basic ACH process that is consistent with the original intent. More than just a theoretical concept, ACH-SL is also a functioning, implemented technology known as *ShEBA* – developed by DSTC – which provides a framework for the analysis of multiple hypotheses with multiple items of evidence.

The Subjective Logic calculus used by the system provides a means of translating between both terms used by human agents, and bayesian data used by other systems. The ability to express calculated opinions in everyday human terms allows the model and results to be more easily translated into appropriate language for presentation to policy makers and other non-analysts.

ACH-SL uses a formal calculus, known as Subjective Logic to make recommendations about the likelihoods of the hypotheses, given individual analyst judgements, uncertain knowledge about the value of the evidence, and multiple items of evidence. In addition, ACH-SL derives measures of diagnosticity for each item of evidence with respect to the hypotheses directly from the judgements that analysts make, rather than requiring them to consider diagnosticity as a separate, second-order measure of the evidence.

ACH-SL allows analysts to apply both deductive or abductive reasoning, and minimize the errors that often occur as a result of consideration of only one logical conditional. The ability of ACH-SL to allow both approaches helps to ensure that the individual tasks undertaken by analysts require less possible cognitive effort. This in turn allows analysts to focus on the judgements they make to produce higher-quality analysis as a result.

REFERENCES

- [1] M. Clark, "Intelligence activities," Commission on Organization of the Executive Branch of the Government [the Hoover Commission], Tech. Rep., June 1955, interim report to Congress prepared by a team under Gen. Mark Clark.
- [2] M. Warner, "Wanted: A definition of "intelligence"," *Studies in Intelligence*, vol. 46, no. 3, 2002. [Online]. Available: <http://www.cia.gov/csi/studies/vol46no3/article02.html>
- [3] A. Tversky and D. Kahneman, "Judgement under uncertainty: Heuristics and biases," *Science*, vol. 125, pp. 1124–1131, 1974.
- [4] —, "The framing of decisions and the psychology of choice," *Science*, vol. 211, no. 4481, pp. 453–458, 1981.
- [5] M. Thürling and H. Jungermann, "Constructing and running mental models for inferences about the future," in *New Directions in Research on Decision Making*, B. Brehmer, H. Jungermann, P. Lourens, and G. Sevón, Eds. Elsevier, 1986, pp. 163–174.
- [6] K. J. Gilhooly, *Thinking: Directed, Undirected and Creative*. Academic Press, 1996.
- [7] R. J. Heuer, *Psychology of Intelligence Analysis*. Washington, D.C.: Central Intelligence Agency Center for the Study of Intelligence, 1999. [Online]. Available: <http://www.cia.gov/csi/books/19104>
- [8] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *The Psychological Review*, vol. 63, pp. 81–97, 1956. [Online]. Available: <http://www.well.com/user/smalin/miller.html>
- [9] W. Fishbein and G. Treverton, "Making sense of transnational threats," CIA Sherman Kent School for Intelligence Analysis, Tech. Rep. 1, 2004. [Online]. Available: http://www.odci.gov/cia/publications/Kent_Papers/pdf/OPV3No1.pdf
- [10] R. Z. George, "Fixing the problem of analytical mind-sets: Alternative analysis," *International Journal of Intelligence and Counter Intelligence*, vol. 17, no. 3, pp. 385–405, Fall 2004.
- [11] A. Jøsang, "A Logic for Uncertain Probabilities," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, no. 3, pp. 279–311, June 2001.
- [12] F. J. Stech and C. Elässer, "Midway revisited: Deception by analysis of competing hypothesis," MITRE Corporation, Tech. Rep., 2004. [Online]. Available: http://www.mitre.org/work/tech_papers/tech_papers_04/stech_deception
- [13] A. George, *Presidential Decisionmaking in Foreign Policy: The Effective Use of Information and Advice*. Boulder CO, USA: Westview Press, 1980.
- [14] G. S. Halford, R. Baker, J. E. McCredden, and J. D. Bain, "How many variables can humans process?" *Psychological Science*, vol. 16, no. 1, pp. 70–76, January 2005.
- [15] J. Davis, "Improving CIA analytic performance: Strategic warning," CIA Sherman Kent School for Intelligence Analysis, Tech. Rep. 1, 2002. [Online]. Available: http://www.odci.gov/cia/publications/Kent_Papers/pdf/OPNo1.pdf
- [16] K. Burns, *Mental Models and Normal Errors*. Lawrence Erlbaum Associates, 2004, ch. How Professionals Make Decisions. [Online]. Available: http://mentalmodels.mitre.org/Contents/NDM5_Chapter.pdf
- [17] A. Jøsang and S. Pope, "Conditional deduction under uncertainty," in *Proceedings of the 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2005)*, 2005.
- [18] B. de Finetti, "The true subjective probability problem," in *The concept of probability in psychological experiments*, C.-A. Staël von Holstein, Ed. Dordrecht, Holland: D.Reidel Publishing Company, 1974, pp. 15–23.
- [19] J. M. Keynes, *A Treatise on Probability*. Macmillan, 1921, ch. 4 "Fundamental Ideas".
- [20] J. Zlotnick, "Bayes' theorem for intelligence analysis," *Studies in Intelligence*, vol. 16, no. 2, Spring 1972. [Online]. Available: http://www.odci.gov/csi/kent_csi/pdf/v16i2a03d.pdf
- [21] A. Tversky and D. Kahneman, *Judgment under Uncertainty: Heuristics and Biases*. Press Syndicate of the University of Cambridge, 1982, ch. Causal schemas in judgments under uncertainty, pp. 117–128.
- [22] M. Diaz, *Topics in the Logic of Relevance*. München: Philosophia Verlag, 1981.
- [23] S. Kent, *Sherman Kent and the Board of National Estimates: Collected Essays*. CIA, Center for the Study of Intelligence, 1994, ch. Words of Estimated Probability. [Online]. Available: <http://www.cia.gov/csi/books/shermankent/6words.html>
- [24] A. Jøsang and D. McAnally, "Multiplication and Comultiplication of Beliefs," *International Journal of Approximate Reasoning*, vol. 38, no. 1, pp. 19–51, 2004.
- [25] A. Jøsang, "The Consensus Operator for Combining Beliefs," *Artificial Intelligence Journal*, vol. 142, no. 1–2, pp. 157–170, October 2002.
- [26] A. Jøsang, M. Daniel, and P. Vannoorenberghe, "Strategies for Combining Conflicting Dogmatic Beliefs," in *Proceedings of the 6th International Conference on Information Fusion*, X. Wang, Ed., 2003.

APPENDIX
DEFINITIONS, THEOREMS, PROOFS AND EXAMPLES

A. *Equivalence of Bayesian and Subjective Opinion*

Definition 1.1: Any pure bayesian opinion $\varphi = (r, s)$ with base rate $\text{br}(\varphi) = a^\varphi$ can be transformed into a subjective opinion $\omega = (b, d, u)$ with base rate $\text{br}(\omega) = a^\omega$ using the transformation function $\mathcal{F}(\varphi) = \omega$ [11].

$$\mathcal{F}(\varphi) = \omega = (b, d, u) \left\{ \begin{array}{l} b = \frac{r}{r+s+2} \\ d = \frac{s}{r+s+2} \\ u = \frac{2}{r+s+2} \\ a^\omega = a^\varphi \end{array} \right. \quad (\text{A-1})$$

Corollary 1.2: Any subjective opinion $\omega = (b, d, u)$ can be transformed into a pure bayesian $\varphi = (r, s)$ using the transformation function $\mathcal{F}'(\omega) = \varphi$.

$$\mathcal{F}'(\omega) = \varphi = (r, s) \left\{ \begin{array}{l} r = \frac{2b}{u} \\ s = \frac{2d}{u} \\ a^\varphi = a^\omega \end{array} \right. \quad (\text{A-2})$$

B. *Logical Conditionals*

Logical conditionals are pairs of conditional beliefs that are used for reasoning about the likelihood of a hypothesis, given some evidence. There are two types of logical conditionals used in inductive reasoning – *abductive* and *deductive* conditionals – and they have different mathematical requirements placed on their values as a result. This section will define the mathematical requirements and constraints for both types of logical conditionals.

1) *Deductive Logical Conditionals:* The logical conditionals used for deduction, $\text{p}(h_i|e)$ and $\text{p}(h_i|\bar{e})$, must obey certain mathematical constraints in order to satisfy basic requirements of probability theory. For a complete set of hypotheses where one and only one hypothesis can be true, it logically follows that the sum of the probability expectations of the positive conditionals must be one; and similarly the probability expectations of the negative conditionals must also sum to one, since

$$\begin{aligned} \text{p}(h_1) + \text{p}(h_2) + \dots + \text{p}(h_n) &= 1 \\ \text{p}(\bar{h}_1) + \text{p}(\bar{h}_2) + \dots + \text{p}(\bar{h}_n) &= 1 \end{aligned}$$

Definition 1.3 (Deductive Logical Conditionals): Let $\Phi = \{h_1, h_2, \dots, h_k\}$ be a complete set of k hypotheses where one and only one $h_i \in \Phi$ is true. Then the logical conditionals used for deduction on a

single item of evidence, e ($\omega_{h_i|e}$ and $\omega_{h_i|\bar{e}}$), must obey the following mathematical constraints

$$\begin{aligned} \sum_{n=1}^{n=k} \mathbb{E}(\omega_{h_n|e}) &= 1 \\ \sum_{n=1}^{n=k} \mathbb{E}(\omega_{h_n|\bar{e}}) &= 1 \end{aligned} \tag{A-3}$$

2) *Abductive Logical Conditionals*: The logical conditionals used for abduction, $p(e|h_i)$ and $p(e|\bar{h}_i)$, have different constraints to deductive logical conditionals and there is no requirement that the probability expectations of their positive conditionals sum to one, so the probability expectation of the positive conditional $p(e|h_i)$ may be unconstrained for each hypothesis. However, the probability expectation of each negative conditional $p(e|\bar{h}_i)$ must be equal to the average of the probability expectations of the other positive conditionals of the other hypotheses, i.e.

$$p(e|\bar{h}_1) = \frac{p(e|h_2) + p(e|h_3) + \dots + p(e|h_n)}{n-1} \tag{A-4}$$

Definition 1.4 (Abductive Logical Conditionals): Let $\Phi = \{h_1, h_2, \dots, h_k\}$ be a complete set of k , ($k > 1$) hypotheses where one and only one $h_i \in \Phi$ is true. Then the logical conditionals used for abduction on a single item of evidence, e ($\omega_{e|h_i}$ and $\omega_{e|\bar{h}_i}$), must obey the following mathematical constraints

$$\forall h_i, h_i \in \Phi, \mathbb{E}(\omega_{e|\bar{h}_i}) = \frac{\sum_{n=1}^{n=k} \mathbb{E}(\omega_{e|h_n}) - \mathbb{E}(\omega_{e|h_i})}{k-1} \tag{A-5}$$

and consequently, the following two conditions must also hold true

$$\mathbb{E}(\omega_{e|h_i}) = \sum_{n=1}^{n=k} \mathbb{E}(\omega_{e|\bar{h}_n}) - (k-1) \cdot \mathbb{E}(\omega_{e|\bar{h}_i}) \tag{A-6}$$

$$\sum_{n=1}^{n=k} \mathbb{E}(\omega_{e|h_n}) = \sum_{n=1}^{n=k} \mathbb{E}(\omega_{e|\bar{h}_n}) \tag{A-7}$$

C. Model coherence

The belief values assigned to the deductive logical conditionals of different items of evidence for the same hypothesis must share a common overlapping range for the model to be coherent. If there is no common overlap in the ranges of belief assigned to the logical conditionals, then at least one of the logical conditionals of either item of evidence must be incorrect.

Definition 1.5 (Model Coherence): Let $\Phi = \{h_1, h_2, \dots, h_k\}$ be a state space for a set k hypotheses where one and only one $h_i \in \Phi$ is true. Let $\xi = \{e_1, e_2, \dots, e_n\}$ be a set of n items of evidence. Then a model is defined to be *coherent* for a hypothesis, h_i , if and only if

$$\text{lower}(h_i) \leq \text{upper}(h_i) \tag{A-8}$$

where the upper bound, $\text{upper}(h_i)$ and the lower bound, $\text{lower}(h_i)$ are defined as

$$\begin{aligned} \text{lower}(h_i) &= \max(\forall e_j \in \xi, \min(p(h_i|e_j), p(h_i|\bar{e}_j))) \\ \text{upper}(h_i) &= \min(\forall e_j \in \xi, \max(p(h_i|e_j), p(h_i|\bar{e}_j))) \end{aligned} \tag{A-9}$$

D. Diagnosticity

Diagnosticity is a measure of how well evidence distinguishes between hypotheses, based on knowledge of the logical conditionals $p(h_i|e)$ and $p(h_i|\bar{e})$.

Definition 1.6 (Diagnosticity of evidence): Let $\Phi = \{h_1, h_2, \dots, h_k\}$ be a state space for a set k hypotheses where one and only one $h_i \in \Phi$ is true. Let $\Omega^\Phi = \{\Theta_1, \Theta_2, \dots, \Theta_m\}$ be the corresponding set of m state spaces for a single item of evidence, e (where $\Theta_i = \{e_i, \bar{e}_i\}, \Theta_i \in \Omega^\Phi$) that represent the conditionals $\omega_{h_i|e}, \omega_{h_i|\bar{e}}$ for each hypothesis $h_i \in \Phi$. Then we define the *diagnosticity* of the evidence e with respect to an arbitrary subset of hypotheses $H \subseteq \Phi$ and $H \not\subseteq \Phi$ with elements $k > 0$ to be

$$D(e, H) = \begin{cases} E_{\text{total}}(e, H) = 0, & 0; \\ E_{\text{total}}(e, H) > 0, & \frac{\sum_{n=1}^{n=k} |E(\omega_{h_n|e}) - E(\omega_{h_n|\bar{e}}) - D_{\text{mean}}(e, H)|}{E_{\text{total}}(e, H)} \end{cases} \quad (\text{A-10})$$

where $D_{\text{mean}}(e, H)$ is the mean of the sum of the differences, and $E_{\text{total}}(e, H)$ is the sum of their expectations, defined respectively as

$$D_{\text{mean}}(e, H) = \frac{\sum_{n=1}^{n=k} [E(\omega_{h_n|e}) - E(\omega_{h_n|\bar{e}})]}{k} \quad (\text{A-11})$$

$$E_{\text{total}}(e, H) = \sum_{n=1}^{n=k} [E(\omega_{h_n|e}) + E(\omega_{h_n|\bar{e}})] \quad (\text{A-12})$$

Then the *diagnosticity* of the evidence e with respect to an arbitrary subset of hypotheses H can be rewritten as (substituting A-11 and A-12 into A-10):

$$D(e, H) = \frac{\sum_{n=1}^{n=k} \left| E(\omega_{h_n|e_n}) - E(\omega_{h_n|\bar{e}_n}) - \sum_{n=1}^{n=k} \left[\frac{E(\omega_{h_n|e_n}) - E(\omega_{h_n|\bar{e}_n})}{k} \right] \right|}{\sum_{n=1}^{n=k} [E(\omega_{h_n|e_n}) + E(\omega_{h_n|\bar{e}_n})]} \quad (\text{A-13})$$

Remark 1.7: It can be seen that $D(e, H) \in [0..1]$ where a value of zero indicates that the evidence lends no weight to any of the hypotheses, while a value of 1 indicates that at extreme values for the evidence (i.e. $E(\omega_e) = 0 \vee E(\omega_e) = 1$), one of the hypotheses, $h_i \in H$, will be absolutely true and for the other extreme, one or more will be absolutely false.

Lemma 1.8 (Diagnosticity of evidence for a complete set of hypotheses): The diagnosticity of the evidence e for a complete set of hypotheses Φ can be expressed in a simplified form as

$$D(e, \Phi) = \frac{\sum_{n=1}^{n=m} |E(\omega_{h_n|e}) - E(\omega_{h_n|\bar{e}})|}{2} \quad (\text{A-14})$$

Proof: For a complete set of hypotheses, where $k = m, P = \Phi$, the sum of the expectations of the conditionals will be exactly two, (i.e. $E_{\text{total}}(x, \Phi) = 2$) since

$$\sum_{n=1}^{n=m} E(\omega_e) = 1, \quad \sum_{n=1}^{n=m} E(\omega_{\bar{e}}) = 1 \quad (\text{A-15})$$

and also $D_{\text{mean}}(e, \Phi) = 0$, since

$$\sum_{n=1}^{n=m} [E(\omega_{h_n|e}) - E(\omega_{h_n|\bar{e}})] = 0$$

So the diagnosticity of the evidence e for a complete set of hypotheses Φ can be simplified from A-13:

$$D(e, \Phi) = \frac{\sum_{n=1}^{n=m} |E(\omega_{h_n|e}) - E(\omega_{h_n|\bar{e}})|}{2}$$

■

1) *Example 1:* Consider an exhaustive set of hypotheses $H = \{h_1, h_2, h_3\}$, and the respective expectations of their conditionals with respect to a item of evidence, e to be:

$$\begin{aligned} E(\omega_{h_1|e_1}) &= 1.0 & E(\omega_{h_1|\bar{e}}) &= 0.0 \\ E(\omega_{h_2|e_2}) &= 0.0 & E(\omega_{h_2|\bar{e}}) &= 0.35 \\ E(\omega_{h_3|e_3}) &= 0.0 & E(\omega_{h_3|\bar{e}}) &= 0.65 \end{aligned}$$

Since both the sums of the positive conditionals and the sums of the negative conditionals of H add up to 1, i.e.:

$$\sum_{n=1}^{n=3} E(\omega_{h_n|e}) = 1, \quad \sum_{n=1}^{n=3} E(\omega_{h_n|\bar{e}}) = 1$$

and therefore the mean of the difference between the positive and negative conditionals is zero, and the sum of all conditionals is exactly 2, i.e.

$$D_{\text{mean}}(e, H) = 0, \quad E_{\text{total}}(e, H) = 2$$

it follows that the simplified form (A-14) can be used to obtain:

$$D(e, H) = \frac{|1.0 - 0| + |0 - 0.35| + |0 - 0.65|}{2} = 1.0$$

The diagnosticity of the evidence e in respect of all hypotheses is $D(e, H) = 1.0$ – since when the evidence is true, h_1 must be true and all other hypotheses (h_2, h_3) must be false. Furthermore, the diagnosticities with respect to the subsets of H (using A-13) are:

$$\begin{aligned} D(e, \{h_1, h_2\}) &= 1.0 \\ D(e, \{h_2, h_3\}) &= 0.3 \\ D(e, \{h_1, h_3\}) &= 1.0 \end{aligned}$$

so that it can be seen that the evidence is capable of distinguishing perfectly between h_1 and h_2 , and between h_1 and h_3 , but cannot distinguish well between h_2 and h_3 , (assuming h_1 is false).

2) *Example 2:* Consider another exhaustive set of hypotheses $H = \{h_1, h_2, h_3, h_4\}$, and the respective expectations of their conditionals with respect to a item of evidence, x to be:

$$\begin{aligned} E(\omega_{h_1|x}) &= 0.5 & E(\omega_{h_1|\bar{x}}) &= 0.1 \\ E(\omega_{h_2|x}) &= 0.2 & E(\omega_{h_2|\bar{x}}) &= 0.0 \\ E(\omega_{h_3|x}) &= 0.15 & E(\omega_{h_3|\bar{x}}) &= 0.5 \\ E(\omega_{h_4|x}) &= 0.15 & E(\omega_{h_4|\bar{x}}) &= 0.4 \end{aligned}$$

The diagnosticity of the evidence e with respect to a subset of the hypotheses $H' \subset H$, $H' = \{h_1, h_2, h_3\}$ will be $D(e, H')$ and since both the sums of the positive conditionals and the sums of the negative conditionals of H' do *not* add up to 1, i.e.

$$\sum_{n=1}^{n=3} E(\omega_{h_n|e}) = 0.85, \quad \sum_{n=1}^{n=3} E(\omega_{h_n|\bar{e}}) = 0.6$$

and therefore

$$D_{\text{mean}}(e, H') = 0.25, \quad E_{\text{total}}(e, H') = 1.45$$

it follows that the more usual form (A-13) must be used to obtain $D(e, H')$:

$$D(e, H') = \frac{|(0.5 - 0.1 - 0.083)| + |0.2 - 0.0 - 0.083| + |0.15 - 0.5 - 0.083|}{1.45} = 0.6$$

with the diagnosticities with respect to the subsets of J' being:

$$\begin{aligned} D(e, \{h_1, h_2\}) &= 0.25 \\ D(e, \{h_2, h_3\}) &= 0.65 \\ D(e, \{h_1, h_3\}) &= 0.3 \end{aligned}$$

E. Relevance

Relevance is a measure of how relevant the evidence is for determining the likelihood to a hypothesis, based on knowledge of the logical conditionals $p(h_i|e)$ and $p(h_i|\bar{e})$.

Theorem 1.9 (Relevance of evidence): The relevance of evidence e with respect to a single hypothesis h is defined as the difference between its conditionals (e, \bar{e})

$$R(e, h) = |p(h|e) - p(h|\bar{e})|.$$

It can be seen that $R(e, h) \in [0, 1]$, where $R(e, h) = 0$ expresses total irrelevance/independence, and $R(e, h) = 1$ expresses total relevance/dependence between e and h . For belief conditionals, the same type of relevance can be defined being equivalent to the diagnosticity of evidence e with respect to a single hypothesis h_i and its complement \bar{h}_i .

$$R(e, h_i) = D(e, \{h_i, \bar{h}_i\}) = |E(\omega_{h_i|e}) - E(\omega_{h_i|\bar{e}})| \quad (\text{A-16})$$

Proof: Let $\Phi = \{h_1, h_2, \dots, h_m\}$ be a state space for a set m hypotheses. Let $\Omega^\Phi = \{\Theta_1, \Theta_2, \dots, \Theta_m\}$ be the corresponding set of m state spaces for a single item of evidence, x (where $\Theta_i = \{x, \bar{x}\}$) that represent the conditionals $\omega_{h_i|e}, \omega_{h_i|\bar{e}}$ for each hypothesis h_i .

Then, the diagnosticity for a single hypothesis $h_i \in \Phi$ and its complement $\bar{h}_i = \Phi \setminus \{h_i\} = \{h_1, h_2, \dots, h_m\} \setminus \{h_i\}$ (with $k = m - 1$ elements), can be defined by the diagnosticity of the set $Q = \{h_i, \bar{h}_i\}$ where the conditionals for \bar{h}_i are defined as

$$E(\omega_{\bar{h}_i|e}) = \sum_{n=1}^{n=k} E(\omega_{h_n|e}), \quad \text{and} \quad E(\omega_{\bar{h}_i|\bar{e}}) = \sum_{n=1}^{n=k} E(\omega_{h_n|\bar{e}})$$

It can be shown from A-14 that $Q = \{h_i, \bar{h}_i\}$ form a complete set of hypotheses since both $D_{\text{mean}}(e, Q) = 0$ (A-11) and $E_{\text{total}}(e, Q) = 2$. Therefore, the simplified form for $D(e, Q)$ can be used (A-14), i.e.

$$\begin{aligned}
D(e, Q) = D(e, \{h_i, \bar{h}_i\}) &= \frac{\sum_{n=1}^{n=k} |\mathbb{E}(\omega_{h_n|e}) - \mathbb{E}(\omega_{h_n|\bar{e}})|}{2} \\
&= \frac{|\mathbb{E}(\omega_{h_i|e}) - \mathbb{E}(\omega_{h_i|\bar{e}})| - |\mathbb{E}(\omega_{\bar{h}_i|e}) - \mathbb{E}(\omega_{\bar{h}_i|\bar{e}})|}{2} \\
&= \frac{2 |\mathbb{E}(\omega_{h_i|e}) - \mathbb{E}(\omega_{h_i|\bar{e}})|}{2} \\
&= |\mathbb{E}(\omega_{h_i|e}) - \mathbb{E}(\omega_{h_i|\bar{e}})| \\
&= R(e, h_i)
\end{aligned}$$

■

F. Relative Diagnosticity

When considering a multi-dimensional problem with multiple items of evidence and multiple hypotheses, it is useful to consider the diagnosticity of evidence *relative* to the diagnosticity of all other evidence, rather than its *actual* diagnosticity (see 1.6).

Definition 1.10 (Relative Diagnosticity): Let the set of k items of evidence be $\Psi = \{e_1, e_2, \dots, e_k\}$, and a set of m hypotheses $\Phi = \{h_1, h_2, \dots, h_m\}$ and where the maximum of the evidence diagnosticities D_{max} is defined as:

$$D_{\text{max}}(\Psi, \Phi) = \max_{n=1}^{n=k} [D(e_n, \Phi)] \quad (\text{A-17})$$

Then the *relative diagnosticity* D_{rel} for each item of evidence $e_i \in \Psi$ be defined as the the diagnosticity $D(e_i, \Phi)$ divided by the maximum of the evidence diagnosticities, D_{max} i.e:

$$D_{\text{rel}}(e_i, \Phi) = \begin{cases} D_{\text{max}}(\Psi, \Phi) = 0, & 0; \\ D_{\text{max}}(\Psi, \Phi) > 0, & \frac{D(e_i, \Phi)}{D_{\text{max}}(\Psi, \Phi)}. \end{cases} \quad (\text{A-18})$$