11TH ICCRTS

COALITION COMMAND AND CONTROL IN THE NETWORKED ERA

Formal Methods of Countering Deception and Misperception in Intelligence Analysis

COGNITIVE DOMAIN ISSUES

C2 MODELING AND SIMULATION

C2 ANALYSIS

Simon Pope
Defence Science and Technology Organisation (DSTO)
Edinburgh, South Australia.
Tel: +61 423 783 015
skjpope@gmail.com

Audun Jøsang
Queensland University of Technology
Brisbane, Australia.
a.josang@qut.edu.au

David McAnally
University of Queensland
Brisbane, Australia.
dsm@maths.uq.edu.au

This page is intentionally left blank.

Formal Methods of Countering Deception and Misperception in Intelligence Analysis

Simon Pope
Defence Science and
Technology Organisation (DSTO)
Edinburgh, Australia.
simon.pope@dsto.defence.gov.au

Audun Jøsang
Queensland University of Technology
Brisbane, Australia.
a.josang@qut.edu.au

David McAnally
University of Queensland
Brisbane, Australia.
dsm@maths.uq.edu.au

Abstract— **The development of formal approaches to intelligence analysis has a wide range of application to both strategic and tactical intelligence analysis within law enforcement, defence, and intelligence communities. The ability of these formal models to mitigate attempted deception by an adversary is affected by many factors, including the choice of analytical model, the type of formal representation used, and the ability to address issues of source reliability and information credibility. This paper discusses how the use of Subjective Logic and the modelling approach known as the Analysis of Competing Hypotheses using Subjective Logic (ACH-SL) can provide a level of protection against attempted deception and misperception.**

I. Introduction

O, what a tangled web we weave,
When first we practise to deceive!

– Sir Walter Scott, *Marmion*, 1808

In many ways, intelligence analysts are very similar to physical scientists. They both study aspects of the world in order to understand its causes and effects, and to make predictions about their future states. However, the natures of their domains of enquiry are vastly different. Physical processes are not perverse in their behaviour – they do not attempt to deceive their observers – and they are neither arbitrary nor capricious. Their laws are assumed to be universal and constant, even if the theories that attempt to describe these laws are not. The same cannot be said for the domain of intelligence analysts. For the most part, they study aspects of human behaviour 'in the wild', where their subjects exhibit complex and perverse behaviours – including attempts to deceive those observing them.

Despite the differences in their domains of enquiry, both physical scientists and intelligence analysts need to apply a rigourous methodological approach in studying their subjects. Naturally, scientists will use scientific method to understand the physical world – producing very valuable knowledge as a result – and intelligence analysts likewise should apply some suitable methodology.

Both science and intelligence analysis require the enquirer to choose from among several alternative hypotheses in order to present the most plausible of these as likely explanations for what they observe. Scientists who do not use a rigourous methodology risk their work being scorned, along with their professional reputations. Intelligence analysts who do not use a rigourous methodology risk something far greater – catastrophic intelligence failure.

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.

Analysts that do not use some rigourous 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 [1]. The single greatest flaw with this approach is that the evidence may be consistent with more than one hypothesis, and unless the analyst evaluates every reasonable alternative, they may arrive at an incorrect conclusion. Worse still, if an adversary is undertaking deception, then the evidence may have been suppressed or manipulated so as to lead the analyst to false conclusions.

It is therefore foolhardy to believe that good intelligence can be developed by relying solely on human cognition without resort to methodologies or frameworks that attempt to augment human cognitive capacity while also mitigating its defects. The following sections will discuss formal approaches based on the Analysis of Competing Hypotheses (ACH) [1] – and in particular the variant known as ACH-SL [2], and some of the ways in which the application of ACH-SL can serve to mitigate or detect adversarial deception and non-adversarial misperception.

II. Deception and Misperception

Deception is the act of deceiving. It is an intentional action that requires both a deceiver and someone to be deceived. We say that someone is deceived when they subjectively believe an aspect of the world to be in some state other than it objectively is, as a result of the deliberate actions of an adversary.

Planned deception by an adversary can be categorised into two general types – *simulation* and *dissimulation*. Dissimulation is the act of hiding or obscuring, while simulation attempts to show the false [3]. These can be further categorised into the practices of *masking*, *repackaging*, *dazzling*, *mimicking*, *inventing* and *decoying* [3], [4], [5], [6].

Deception works because of human perceptual and cognitive biases [7]. Our expectations and our experience have a lasting and direct influence on our perceptions. We fail to correctly reason about alternatives that do not align with our expectations, and we assign and judge evidence according to our expectations and our experiences [6], [7]. Consequently, we miss important events, discount information that is not consistent with the expected outcome, and do not consider alternative outcomes. Stech and Elässer note that people tend to be poor at detecting deception since its occurrence is relatively rare. They categorise four types of analytic errors that hinder detection of deception [8], [9]:

- *Poor anomaly detection*: Analysts miss indicators of anomalies or discount their importance as being either irrelevant or inconsistent with other information.
- *Misattribution*: Analysts attribute inconsistencies or anomalies to collection gaps or processing errors, rather than to deception.
- *Failure to link deception tactics to deception hypotheses*: Analysts fail to recognise anomalies as possible indications of attempted deception.
- *Inadequate support for deception hypotheses*: Analysts fail to consider the likelihood of deception with respect to an adversary's strategic goals.

In practice, simply being able to counter deception is insufficient, since it is the *consequences* of deception and misperception that are of immediate importance – not just its causes. For example, if an analyst were to misinterpret the data that is available to them and therefore misread an adversary's intention, then the consequences could be just as dire as if the adversary deliberately engaged in some form of deception. The problem of misperception – and deception – is a direct consequence of the limitations of our cognitive faculties.

Our limited mental capacity cannot cope with the enormous complexity of the 'real world' so instead we create simplified mental models of reality that approximate what we perceive to be the 'real world', and reason about those models instead. This creates a bounded rationality [10], where each person behaves rationally according to their own simplified model, but not necessarily from any objective perspective. The sufficiency of these mental models as approximate models of the world varies with the risks and rewards of their application. For frequently experienced events, basic human reasoning is usually sufficient [6]. In everyday personal affairs few of our decisions use any directed analytical processes – and even fewer of these require any sort of rigourous approach – due to relatively minor consequences. The same is not true for large-scale human affairs – such as the business of nations and corporations – where the relative consequences of decisions can have enormous impact. This distinction in and of itself is cause enough to consider whether human 'everyday reasoning' is robust and reliable enough for use in these contexts.

Unfortunately as humans, we 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 [7], [1], [11], [12], [13], [14]. 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 [1], [15], [16], [17]. Readers looking for a good discussion of the cognitive biases of humans and their impact on intelligence analysis – and deception in particular – should consult Heuer's *Strategic Deception and Counterdeception* [7].

III. Analysis of Competing Hypotheses

Intelligence analysis generally requires that analysts choose from among several alternative hypotheses in order to present the most plausible of these as likely explanations for the evidence being analyzed. One way in which some of the inherent cognitive limitations 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, the simultaneous evaluation of non-trivial problems is a near-impossible feat for human cognition alone. While the limitations of short term memory appear to be around seven items [15], 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 [18]. This implies that for any problem with more than three possible hypotheses or three items of evidence, the ability of humans to reason correctly diminishes rapidly with an increase in the number of items of evidence or hypotheses.

The Analysis of Competing Hypotheses (ACH) approach [1] was developed to provide a framework for assisted reasoning that would help overcome these limitations. ACH was developed in the mid- to late-1970's by Richards Heuer, a former CIA Directorate of Intelligence methodology specialist, in response to his "never-ending quest for better analysis" [1]. His eight-step ACH methodology provides a basic framework for the identification of assumptions, arguments and hypotheses; consideration of all evidence and hypotheses – including its value relative to the hypotheses; 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 [1]. While this original ACH approach is fundamentally sound, it suffers from a number of significant but correctable problems.

1) Base rate errors due to framing and other causes: ACH recommends that analysts consider how consistent each item of evidence is with each possible hypothesis. This can be reasonably interpreted to mean that for each hypothesis, one should consider the likelihood that the evidence is true $\mathrm{p}(e_j|h_i)$ – and this will likely be the interpretation for *derivative* evidence. However, for *causal* evidence[1], a different and possibly erroneous interpretation is likely.

Causal evidence is perceived to have a direct causal influence on a hypothesis, and typically reflects reasoning from cause to effect. An example of this is the presence of a persistent low pressure system being causal evidence for rain, since a low pressure system appears to influence precipitation. The 'state of mind' of an adversary is often regarded as causal evidence since it usually presumed to have direct influence on their decision making processes.

Derivative evidence [19] – also known as *diagnostic evidence* [20] – is indirect intermediate evidence – not usually perceived as being causal in nature – and is usually observed in conjunction or contemporaneous with the occurrence of one or more hypotheses. Derivative evidence typically reflects reasoning from effect back to cause – or where no causal link seems suitable. For example, a soggy lawn would likely be considered derivative evidence for rain since soggy

[1]See *Causal and Derivative Evidence* in [2]

lawns are also associated with the use of sprinklers, and recently-washed automobiles.

The particular problem arises when analysts attempt to make judgements about the likelihood of causal evidence being true when the hypothesis is true $\mathrm{p}(e_j|h_i)$. Since there is an apparent causal relationship between the evidence and the hypothesis, the analyst is more likely to reason from the cause to the effect – from the evidence to the hypothesis. The danger lies in the the analyst reasoning about the likelihood of the hypothesis, given that the evidence is true $\mathrm{p}(h_i|e_j)$ and taking this as an approximation of $\mathrm{p}(e_j|h_i)$. This type of reasoning tends to ignore the likelihood of of the hypothesis being true when the evidence is false $\mathrm{p}(h_i|\bar{e}_j)$, and can produce very misleading results. Stech and Elässer [8] make a similar 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. They developed ACH-CD[2] as a modified variant of ACH to account for cognitive factors that make people poor at detecting deception [8]. Stech and Elässer 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 $\mathrm{p}(h_i|e_j)$. Their argument is that this type of reasoning neglects the base rates both of the evidence $\mathrm{br}(e_j)$ and of the hypothesis $\mathrm{br}(h_i)$ which can result in reasoning errors that lead to incorrect conclusions [21], and increase susceptibility to deception.

Stech and Elässer demonstrate this with an excellent example of how 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 [8]:

Detect Krypton
p(enrichment | Krypton) = high
→ p(enrichment program) = high
→ p(nuclear program) = 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(Krypton | not enrichment) = medium to high

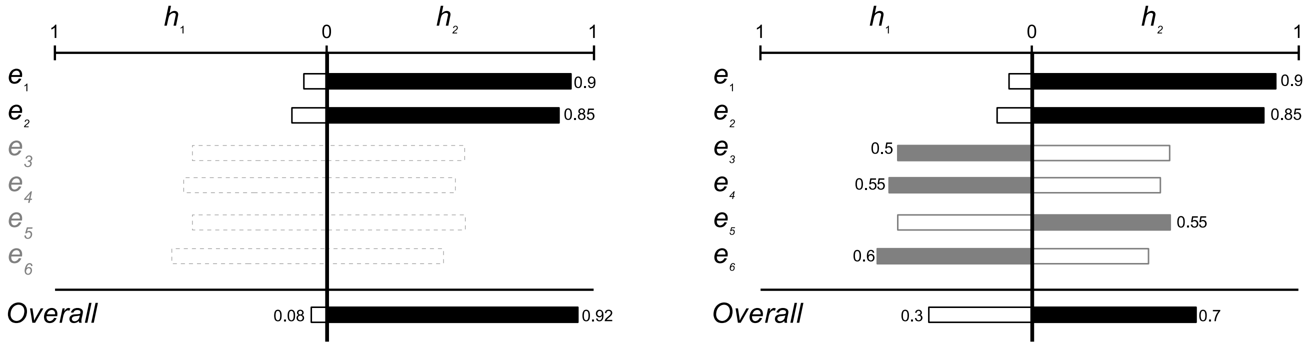
This additional information should lead the analyst to conclude that there is a fair amount of uncertainty of a nuclear program given the detection of Krypton. The assignment of the 'high' value to p(enrichment | Krypton) neglects the fact that an oil-rich middle-eastern country is likely to use Krypton gas – regardless of whether they have a nuclear program.

2) Problems of discarding of weakly diagnostic evidence: Another problem with the original ACH approach is the discarding of evidence that has little ability to distinguish between hypotheses [1]. While this process of discarding weakly diagnostic evidence is intended to mitigate some cognitive biases, it may actually lead to *greater* susceptibility to deception. If an adversary is planning deception, then they might simulate strong indicators of an alternative hypothesis, or dissimulate (i.e. suppress) indicators which have strong diagnostic value for the correct hypothesis. If the analyst has no way to know which evidence, if any, is being dissimulated or simulated, then sole reliance on evidence that has relatively strong diagnostic value may lead the analyst to an incorrect conclusion.

In all forms of ACH, analysts are required to consider how well each item of evidence is capable of distinguishing between the hypotheses. From this, the analyst must arrive at measures of diagnosticity that are consistent across each hypothesis for an item of evidence, and also across the items of evidence. This is a high-load cognitive task that is likely to result in inconsistencies or very coarse measures of diagnosticity that may not provide a clear differentiation between hypotheses. As a consequence, the diagnosticity of an item of evidence may be so coarse that the evidence appears to have no diagnostic value. Discarding many items of evidence that appear to have no or little diagnosticity may remove noise from the analysis but may result in the loss of 'weak signals' that contradict the strong indicators and could be suggestive of attempted deception.

The problem arises since a combination of weak diagnostic evidence can significantly affect the overall results. As an illustration of this point, compare the results of Fig. 1(a), where weak diagnostic evidence was excluded from the model, with the results of

[2] Analysis of Competing Hypotheses – Counter Deception (ACH-CD)



(a) A two-hypothesis analytical model without consideration of weak diagnostic evidence $e_3 \ldots e_6$.

(b) The same analytical model but with the inclusion of the weak diagnostic evidence.

Fig. 1. Example of the possible effect on the overall result from including or discarding weak diagnostic evidence in an abstract, hypothetical problem. The weak diagnostic evidence $e_3 \ldots e_6$ is missing in Fig 1(a), and shown in grey in Fig 1(b). This example has been greatly simplified for illustrative purposes. The overall conclusions would be calculated using the *consensus* of the subjective opinions $\omega_{h_i||e_j}$ – and not an average simple scalar probabilities $\mathrm{p}(h_i||e_j)$ – which provide overall beliefs in ω_{h_1} and ω_{h_2}.

Fig. 1(b), where the weak diagnostic evidence was not excluded. In Fig. 1(a), the conclusions strongly support Hypothesis B, while in Fig. 1(b), the additional weak diagnostic evidence decreases its likelihood – and by implication might be indicative of deception or some other type of misperception.

Often the presence of weak diagnostic evidence allows the calculation of interesting metrics that may serve as good indicators of deception or misperception, especially in cases where it would be difficult for an adversary to simulate or dissimulate the weak diagnostic evidence. These metrics will be discussed in detail later.

3) Limitations of disconfirming evidence: Lastly, the original ACH approach suggests that the process of analysis should attempt to disconfirm each hypothesis – where the hypothesis with the least amount of disconfirming evidence is likely to be the correct one [1]. While this approach is appropriate for scientific enquiry, it is insufficient for intelligence analysis. As already noted, scientific method assumes a world that is not perverse in its behaviour, where its inhabitants are not arbitrary nor capricious in their actions, and where there are assumed to be fundamental laws operating behind its phenomena. The same does not hold true for intelligence analysis, where nations and other actors behave in ways that may appear rational to the participants, but are not always rational from an objective point of view; are subject to rich and dynamic social, political and cultural interests and effects; and where the very evidence itself is likely to be uncertain and may not be from trustworthy sources. As such, it is easy to show that under these conditions of uncertainty, the most likely hypothesis may well be the one with the most amount of evidence against; the most in favour; or neither.

IV. Analysis of Competing Hypotheses using Subjective Logic (ACH-SL)

The *Analysis of Competing Hypotheses using Subjective Logic* (ACH-SL) [2] was developed to address a range of intelligence analysis issues, including susceptibility to deception, and is an elaboration of the original ACH approach. It is more than just a theoretical concept, and has been implemented in analytical technology called S*h*EBA[3] that provides a framework for the analysis of multiple hypotheses with multiple items of evidence.

The ACH-SL process can be described in terms of information transformation, as visualised in Fig. 2. Knowledge or belief about the collected evidence is transformed by the analytical model (the relationships between hypotheses and evidence) to produce the analytical results of the beliefs in the hypotheses. The approach uses the formal calculus known as *Subjective Logic* [22] to make recommendations about the likelihoods of the hypotheses, given individual analyst judgments, uncertain knowledge about the value of the evidence, and multiple items of evidence.

[3] S*h*EBA – Structured Evidence Based Analysis of Hypotheses

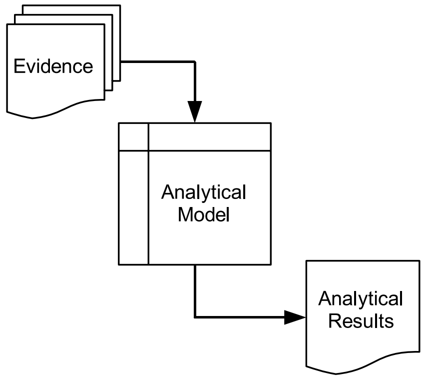


Fig. 2. ACH-SL as an information transformation

The objective of this process is to produce accurate analysis – but not necessarily certain results. In cases where the evidence is clear and the analytical model perfectly distinguishes between the hypotheses, the results should be equally clear and certain. However, even with unreliable or sketchy information and an imperfect model, the analytical results should produce accurate results that reflect this lack of certainty. The accuracy of the analytical results therefore depends on both accurate assessment of the evidence, and the quality of the analytical model.

Accurate assessment of the evidence requires that the analyst correctly judges for each item of collected information, its currency, the reliability of its sources, and how well it supports the evidence. A high quality analytical model requires that analyst has included all plausible hypotheses, has considered all causal factors and diagnostic indicators, and made sound judgements as to how the evidence relates to the hypotheses.

However, even with accurate evidence assessment and a high-quality model, the possibility of deception or misperception still remains. It is still possible that information has been simulated or dissimulated by an adversary in order to mislead the analysis, and it is still possible that the analyst has not generated a suitable set of hypotheses. Fortunately, ACH-SL includes formal metrics which, used correctly, can measure the quality of the analytical model and its results, and provide an indication of possible deception or misperception.

A related issue concerns the discarding of unobservable but relevant evidence. If unobservable yet relevant evidence is not included in the analytical model, then the analyst increases their susceptibility to deception since the evidence may be unobservable due to dissimulation by an adversary. Therefore it is important to accurately model the analyst's ignorance about the value of the evidence, otherwise the lack of evidence can lead to erroneous conclusions, i.e. "absence of evidence is not evidence of absence."[4]

Some knowledge of Subjective Logic is needed to understand ACH-SL. Readers that are not familiar with Subjective Logic are invited to review *Subjective Logic Fundamentals* in Appendix A. Since ACH-SL is an elaboration of the original ACH approach, it follows the same basic approach as the original ACH method [1], except where noted.

A. Outline of the ACH-SL process

The analyst starts by deciding on an exhaustive and exclusive set of hypotheses to be considered. The term 'exhaustive and exclusive' refers to the condition that one of the hypotheses – and only one – must be true. The competing hypotheses may be alternative courses of action, adversarial intentions, force strength estimations, etc. Deciding on what hypotheses to include is extremely important. If the correct hypotheses are not included in the analysis, then the analyst will not get the correct answer, no matter how good the evidence. However, the issue of hypothesis generation lies outside the scope of this document is discussed further in [1].

Next, the analyst considers each hypothesis and lists its possible causal influences and diagnostic indicators. These form the items of evidence that will be considered in the analysis. In deciding what to include, it is important not to limit the evidence to what is already known or believed to be discoverable. Heuer makes the excellent point that analysts should interpret 'evidence' in its broadest sense and should not be limited just to current intelligence reporting [1].

Since ACH-SL requires calculations that are generally too complex and time consuming to perform manually, an appropriate analytical framework that embodies ACH-SL – such as S*h*EBA – is highly recommended. Using this framework the analyst constructs an analytical model of how the evidence relates to the hypotheses. Not all evidence is created equal – some evidence is better for distinguishing between

[4]Dr. Carl Sagan, astronomer, writer and scientist, 1934-1996.

hypotheses than others. The degree to which evidence is considered diagnostic is the degree to which its presence or absence is indicative of one or more hypotheses. If an item of evidence seems consistent with all the hypotheses, it will generally have weak diagnostic value. In the original ACH method, diagnosticity is explicitly provided by the analyst as an input [1], 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.

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 a set of conditional beliefs for each combination of hypothesis h_i and item of evidence e_j. For *causal* evidence, the conditionals $\omega_{h_i|e_j}, \omega_{h_i|\bar{e}_j}$ represent the beliefs that the hypothesis will be true, assuming that the item of evidence is true – and the belief that the hypothesis will be true, assuming the item of evidence is false. For *derivative* evidence, $\omega_{e_j|h_i}$ represents the belief that the item of evidence will be true, assuming that the hypothesis is true. In forming the conditionals for an item of evidence, the analyst must separate out their understanding of the item of evidence under enquiry from the general set of evidence to be considered, i.e. the analyst must not consider the significance of other evidence when forming the conditionals. Failure to do so can bias the analysis.

The choice of whether an item of evidence should be treated in a causal or derivative manner is immaterial to the calculations – the style of reasoning that produces the least cognitive load should be the primary consideration[5]. Analysts can choose to make statements about combinations of hypotheses such as $\omega_{(h_2 \vee h_3)|e_2}$, but not for combinations of evidence since this would likely introduce bias.

It is important to note that a distinction can be made between events that can be repeated many times and events that can only happen once. Conditionals for events that can be repeated many times are frequentist events and can be expressed as simple Bayesian beliefs[6] *if* there can be absolute certainty regarding their values. However, expressing a conditional as a frequentist probability seems to be a meaningless notion when the consequent is an event that can only happen once. Even when the conditionals are calculated from a very large set of data, the possibility remains that the evidence at hand does not provide complete information about these conditionals, and can not be expressed on a purely frequentist form[7]. For events that can only happen once – including almost all problems of intelligence analysis – the observer must arbitrarily decide what the conditionals should be, and

[5]See *Causal and Derivative Evidence* in [2]

[6]Which correspond to zero-uncertainty beliefs in Subjective Logic. See Appendix A.

[7]See [24] for a discussion about the problems of induction and causation

	h_1 br(h_1)	h_2 br(h_2)	h_3 br(h_3)
e_1	$\omega_{e_1\|h_1}$	$\omega_{e_1\|h_2}$	$\omega_{e_1\|h_3}$
e_2	$\omega_{h_1\|e_2}$ $\omega_{h_1\|\bar{e}_2}$	$\omega_{(h_2 \vee h_3)\|e_2}$ $\omega_{(h_2 \vee h_3)\|\bar{e}_2}$	
e_3	$\omega_{h_1\|e_3}$ $\omega_{h_1\|\bar{e}_3}$	$\omega_{h_2\|e_3}$ $\omega_{h_2\|\bar{e}_3}$	$\omega_{h_3\|e_3}$ $\omega_{h_3\|\bar{e}_3}$

(a) Model defined by user using a mixture of causal and derivative conditionals

	h_1 br(h_3)	h_2 br(h_2)	h_3 br(h_3)
e_1	$\omega_{e_1\|h_1}$ $\omega_{e_1\|\bar{h}_1}$	$\omega_{e_1\|h_2}$ $\omega_{e_1\|\bar{h}_2}$	$\omega_{e_1\|h_3}$ $\omega_{e_1\|\bar{h}_3}$
e_2	$\omega_{h_1\|e_2}$ $\omega_{h_1\|\bar{e}_2}$	$\omega_{h_2\|e_2}$ $\omega_{h_2\|\bar{e}_2}$	$\omega_{h_3\|e_2}$ $\omega_{h_3\|\bar{e}_2}$
e_3	$\omega_{h_1\|e_3}$ $\omega_{h_1\|\bar{e}_3}$	$\omega_{h_2\|e_3}$ $\omega_{h_2\|\bar{e}_3}$	$\omega_{h_3\|e_3}$ $\omega_{h_3\|\bar{e}_3}$

(b) Same model after transformation into a normalised form

Fig. 3. Example ACH-SL model construction. Note that in Fig. 3(a), e_1 has derivative conditionals and that for e_2, a single pair of conditionals has been specified for the combination of $h_2 \vee h_3$. Fig. 3(b) is the equivalent representation of Fig. 3(a) with all individual conditionals generated by a normalisation transform. The base rate of each hypothesis is represented by $\mathrm{br}(h_i)$.

consequently there can be a great deal of uncertainty about their values. For these non-frequentist problems, each conditional is usually expressed as a *maximized-uncertainty* belief [22], where the uncertainty of the belief is set to the maximum allowable amount for the desired probability expectation value. Therefore, the conditionals for any problem can usually be provided by the analyst as simple scalar probabilities – i.e. $\mathrm{p}(h_i|e_j), \mathrm{p}(h_i|\bar{e}_j)$ – and the uncertainty maximization can be handled by the analytical system.

Regardless of how the conditionals for the hypotheses are specified – derivative or causal, single- or multi-hypothesis, zero-uncertainty or uncertainty-maximized – the conditionals can be expressed (within the internals of the analytical framework) as a 'normalised' set of $\omega_{h_i|e_j}, \omega_{h_i|\bar{e}_j}$ conditionals (see Fig. 3 for an example). The complete set of conditionals for all items of evidence and all hypotheses constitutes the analytical model.

After completion of the analytical model, it can be used to evaluate a complete or incomplete set of evidence. The inputs to the analytical model are a set of observable evidence $\mathcal{E} = \{\omega_{e_1}, \omega_{e_2}, \ldots \omega_{e_k}\}$ The value of each item of evidence can be highly certain or uncertain, with varying degrees of likelihood, depending on the reliability of the sources of information for the evidence, the presumed accuracy of the sources' observations, the currency of the information, and other factors which will be discussed later. Evidence for which no data is available is expressed as a vacuous belief that is completely uncertain.

The primary output of the analysis is a set of n beliefs $\mathcal{H} = \{\omega_{h_1}, \omega_{h_2}, \ldots \omega_{h_n}\}$, representing the certainty and likelihood of each hypothesis. In addition, intermediate analytical items are available, including separate analytical results for each combination of n hypotheses and m items of evidence, i.e.

$$\mathcal{H}||\mathcal{E} = \{\omega_{h_1||e_1}, \omega_{h_1||e_2}, \ldots \omega_{h_1||e_m}, \omega_{h_2||e_1}, \omega_{h_2||e_2}, \ldots \omega_{h_n||e_m}\} \quad \text{(IV.1)}$$

Deriving the primary results from an analytical model and a set of evidence is a two-step process. Fig. 4 shows how both the intermediate and primary results are derived from the combination of a set of evidence and the analytical model. First, a set of intermediate results $\omega_{h_i||e_j}$ are calculated for each pair of conditionals in the analytic model. Different but equivalent formulas for calculation are used depending on whether the conditionals are causal or or derivative. For causal conditionals, $\omega_{h_i||e_j}$ is calculated using the *deduction* operator $\circledcirc$ [25]. For derivative conditionals,

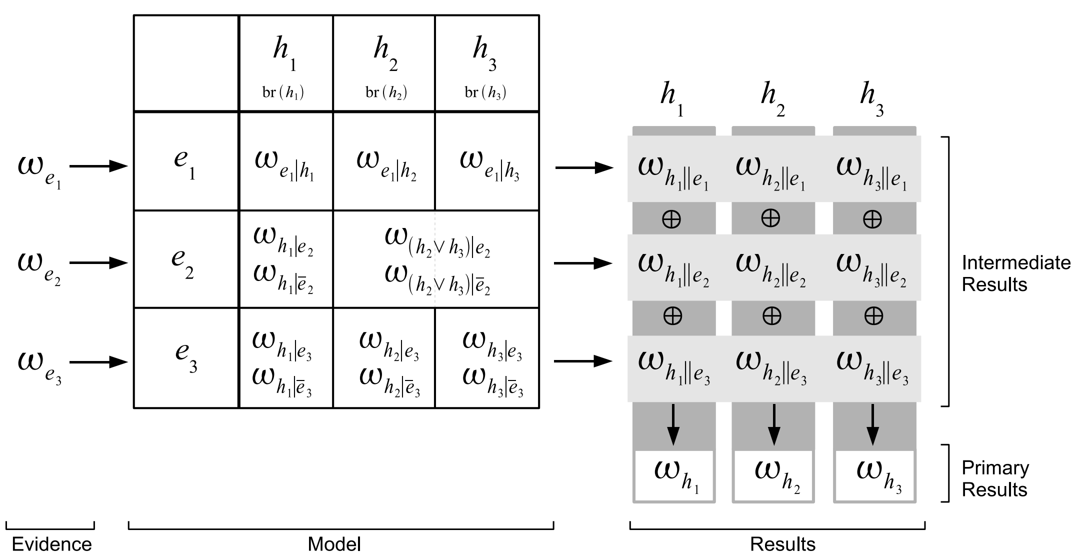


Fig. 4. Generation of intermediate and primary results from the ACH-SL model in Fig. 3(b), and values for the evidence e_1, e_2, e_3. The primary results are a result of belief fusion $\oplus$ [23] of the intermediate results for each hypothesis. Note the method by which the intermediate results are obtained is not shown. For details, please refer to the text.

$\omega_{h_i||e_j}$ is calculated using the *abduction* operator $\bar{\circledcirc}$ [8].

$$\omega_{h_i||e_j} = \omega_{e_j} \circledcirc (\omega_{h_i|e_j}, \omega_{h_i|\bar{e}_j}) \quad \text{(IV.2)}$$

$$\omega_{h_i||e_j} = \omega_{e_j} \bar{\circledcirc} (\omega_{e_j|h_i}, \omega_{e_j|\bar{h}_i}, \mathrm{br}(h_i)) \quad \text{(IV.3)}$$

The second and last step of the process involves fusing the intermediate results for each hypothesis h_i using the consensus operator $\oplus$ [23] to obtain the overall belief in each hypothesis ω_{h_i}.

$$\omega_{h_i} = \omega_{h_i||e_1} \oplus \omega_{h_i||e_2} \cdots \oplus \omega_{h_i||e_m} \quad \text{(IV.4)}$$

V. ACH-SL METRICS AND THEIR USES

In the original ACH approach, after the primary analytical results are calculated, it is recommended that the analyst perform additional analysis to determine how sensitive the results are to a few pieces of crucial evidence. While the basic ACH approach does not elaborate on the methods for doing this, the ACH-CD approach provides computed indicators of possible vulnerabilities for deception [9]. Similarly, ACH-SL provides a number of metrics that are useful in indicating the sensitivity of the conclusions, the possibility of misperception or deception, and the explanatory power of the chosen hypotheses.

Two key metrics that can be derived from the analytical model without the need to consider either the beliefs of either the items of evidence or the analytical results, are *diagnosticity* and *sensitivity*:

- **Diagnosticity** is a measure of how well an item of evidence is capable of distinguishing between a chosen subset of hypotheses. As an aid for intelligence collection, diagnosticity is most useful as a guide for which evidence would be useful for analysing a particular problem.
- **Sensitivity** is most useful for analysing the sensitivity of the results. It is a measure of the relative influence of a single item of evidence on the primary results for a subset of hypotheses $\{\omega_{h_1}, \omega_{h_2}, \ldots \omega_{h_n}\}$. It provides an indication of the degree to which the value of the calculated beliefs could change if the item of evidence e_j were to alter in value.

[8]The abduction operator is described briefly in [2] and will be detailed in a forthcoming paper.

In addition, here are a number of useful metrics that can be derived after the primary and intermediate analytical results are calculated, specifically *support*, *concordance*, and *consistency*:

- **Support** is a measure of the degree to which an intermediate analytical result $\omega_{h_i||e_j}$ supports or opposes the primary result for the hypothesis ω_{h_i}. A positive measure indicates support while a negative result indicates opposition.
- **Concordance** is a measure of the current and potential similarity of a set of beliefs. Beliefs that are completely concordant are exactly the same expectations and base rates, while those with partial or no concordance have different expectations, certainties, or base rates.
- **Consistency** is a measure of how well the primary result for a hypothesis ω_{h_i} is supported by the intermediate results for each item of evidence $h_i||\mathcal{E} = \{\omega_{h_i||e_1}, \omega_{h_i||e_2}, \ldots \omega_{h_i||e_k}\}$.

The following sections describe these metrics in further detail.

A. Diagnosticity

Diagnosticity provides a measure of how well an item of evidence is capable of distinguishing between a chosen subset of hypotheses. For example, the overall diagnosticity of a item of evidence may be poor in distinguishing between six hypotheses, yet it may be very good at distinguishing between just two of those six.

Diagnosticity is derived from the logical conditionals $\mathrm{p}(h_i|e_j)$ and $\mathrm{p}(h_i|\bar{e}_j)$. If these conditionals are not known, then they can be derived from knowledge of the $\mathrm{p}(e_j|h_i)$ and $\mathrm{p}(e_j|\bar{h}_i)$, and from the base rate of the hypothesis $\mathrm{br}(h_i)$[9]. 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. The diagnosticity is a useful measure for intelligence collection purposes, as it indicates the relative importance of a piece of evidence for a particular analytical purpose.

[9]see *Calculating opinions about the hypotheses* in [2]

Details on how diagnosticity is derived are found in [2], and reproduced in Appendix E.

As an aid for intelligence collection, diagnosticity can be used as a guide to determine what items of evidence would be most useful for analysing a particular problem. A high degree of certainty is needed for items of evidence that are diagnostic of high-risk or high-value hypotheses, and should therefore be sourced from multiple channels if possible.

B. Sensitivity

While the diagnosticity metric describes how well a single item of evidence is capable of distinguishing between competing hypotheses, the *sensitivity* metric provides an indication of the degree to which the results could change if the item of evidence were to alter in value. This makes the metric particularly useful for analysing the sensitivity of the primary results. Fig. 5 shows the key difference in the way in which diagnosticity and sensitivity are derived.

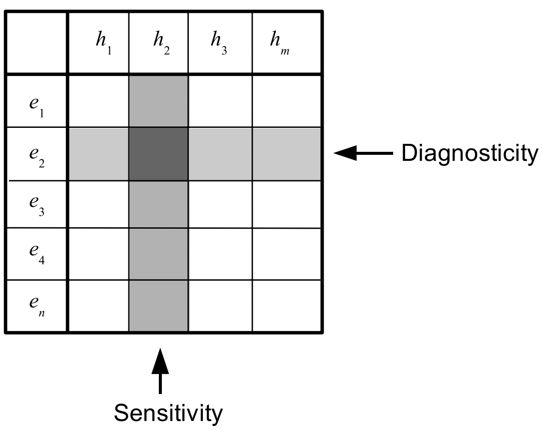


Fig. 5. Diagnosticity is calculated relative to the conditionals in the same row, while sensitivity is calculated relative to the conditionals in the same column.

The calculated beliefs in the hypotheses $\omega_{h_1}, \omega_{h_2}, \ldots \omega_{h_k}$ will vary with changes in the belief in the evidence $\omega_{e_1}, \omega_{e_2}, \ldots \omega_{e_n}$, that form the inputs to the analytical model. Some items of evidence have greater potential than others to alter the results – some disproportionately so. This is due to both the difference in the expectations of the causal logical conditionals $|\mathrm{E}(\omega_{h_i|e_j}) - \mathrm{E}(\omega_{h_i|\bar{e}_j})|$, and their uncertainty values $u_{h_i|e_j}, u_{h_i|\bar{e}_j}$. A large difference in the expectation of the conditionals has the potential to produce a large fluctuation in the expectation of the intermediate result $\mathrm{E}(\omega_{h_i||e_j})$, as the belief in the item of evidence ω_{e_j} changes (IV.2). However, the degree of certainty of the conditionals determines the minimum certainty of each intermediate result $h_i||e_j$ that is obtained using the deduction or deduction operators (IV.2). The behaviour of the consensus operator is to weight the input intermediate results according to the square of their relative certainty, and therefore the certainty greatly influences the calculation of the primary result ω_{h_i} (IV.4).

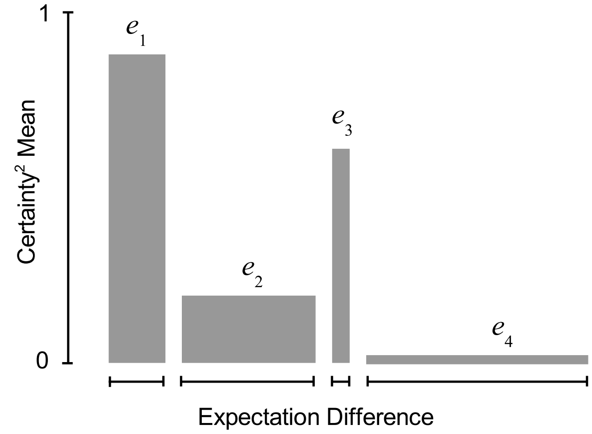


Fig. 6. The sensitivity of a hypothesis with respect to an item of evidence is the relative product of the mean of the squares of their certainties and expectation difference of their causal logical conditionals

The relative degree to which the the value of a hypothesis (or a subset of hypotheses) will be determined by the value of an item of evidence is known as its *sensitivity*, and is useful in sensitivity analysis for quickly identifying the items of evidence that will likely have the greatest impact. It is calculated for a hypothesis (or a subset of hypotheses) and an item of evidence as the magnitude of the product of the difference in the conditionals and their mean certainty, relative to that for all evidence. This is demonstrated in Fig. 6, where the sensitivity of a hypothesis with respect to an item of evidence is the relative area of each.

Sensitivity is a value in the range $\mathrm{S_{rel}} \in [0, 1]$, where a value of zero indicates that the value of the primary result for the subset of hypotheses H is completely independent of the evidence e_j, and conversely, where a value of one indicates complete dependence. The mathematical definition of sensitivity can be found in Appendix F on page 26.

C. Support

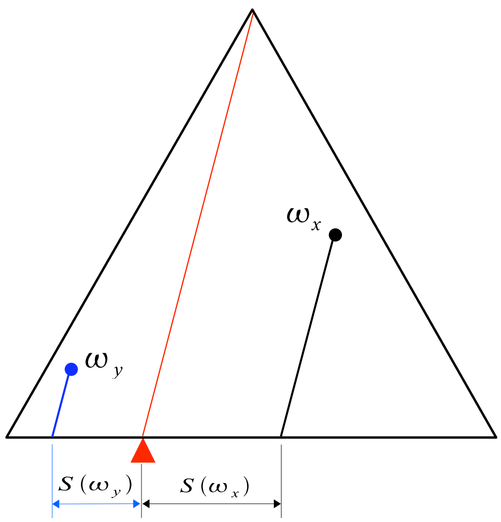


Fig. 7. Support of an opinion is calculated as the difference between its expectation and base rate

Support is a measure of the difference between the expectation and base rate for a belief. When applied to an intermediate analytic result $\omega_{h_i||e_j}$, it indicates the degree to which the result supports or opposes the hypothesis h_i. Support is a value in the range $\mathcal{S}_{\omega_x} \in [-a_x, 1-a_x]$, with a a positive measure indicates support while a negative result indicates opposition. In Fig. 7, the opinion ω_x has positive support, while ω_y has negative support. Generally, evidence that is highly diagnostic will have a greater absolute magnitude of support than for weakly diagnostic evidence. The mathematical definition of support can be found in Appendix G.

D. Concordance

Concordance is a measure of similarity of a set of beliefs, and indicates their degree of agreement. Concordance is a value in the range $\mathcal{K} \in [0,1]$, with a value of $\mathcal{K} = 1$ indicating perfect agreement (i.e. the beliefs are exactly the same), and a value of $\mathcal{K} = 0$ indicating completely antithetical beliefs. It is closely related to but not the same as the difference of their expectations. For beliefs, it is necessary also to account for the projected convergence or divergence as a result of possible past or future decay, or improvement or degradation of the reliability of the source of the belief. For a set of beliefs, concordance is derived from a measure of the mean difference of their expectation values, and the mean difference of the mean of the expectation values of the beliefs that are projected to the same uncertainty.

Concordance is particularly useful as an indication of how closely the intermediate analytical results for a hypothesis agree with each other. In Fig. 8, opinions ω_{x_1} and ω_{x_2} lie on the same line of projection – and have a higher concordance than that of ω_{y_1} and ω_{y_2}. The opinions ω_{y_1} and ω_{y_2} lie on different lines of projection, and therefore have less concordance, even though the expectation differences the two pairs of opinions are the same, and the apparent distance between $\omega_{y_1}, \omega_{y_2}$ is less than that of $\omega_{x_1}, \omega_{x_2}$.

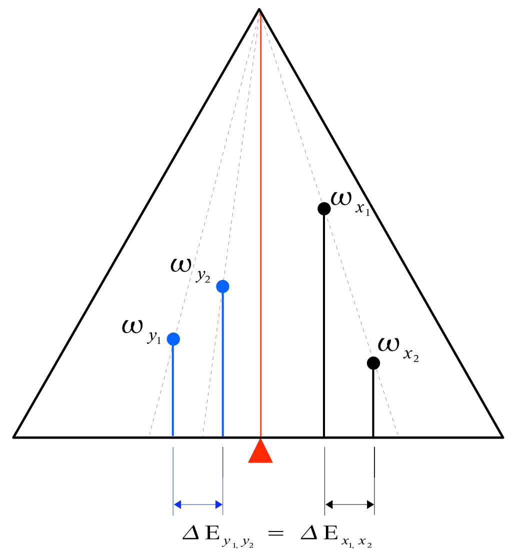


Fig. 8. The concordances of two sets of opinions can be different, even when their expectation differences are the same.

Concordance is calculated with respect to a common reference point, usually the mean opinion ω_μ[10]. In Fig. 9, the concordance is calculated as the complement of half of the sum of the difference in expectations from the mean opinion of both the original opinions and of the opinions projected to the same uncertainty as the mean opinion, i.e.

$$\mathcal{K}_{\omega_x,\omega_y} = 1 - \frac{\Delta E_{x,\mu} + \Delta E_{y,\mu} + \Delta E_{x',\mu} + \Delta E_{y',\mu}}{2}$$

The mathematical definition of concordance can be found in Appendix H.

The concordance of an intermediate analytical result $\omega_{h_i||e_j}$ with respect to the primary result for a hypothesis ω_{h_i} indicates their degree of agreement, taking into

[10]See *Mean of a set of opinions* in Appendix D

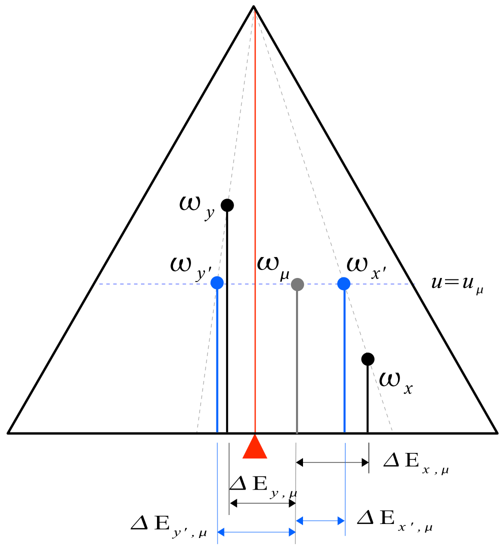


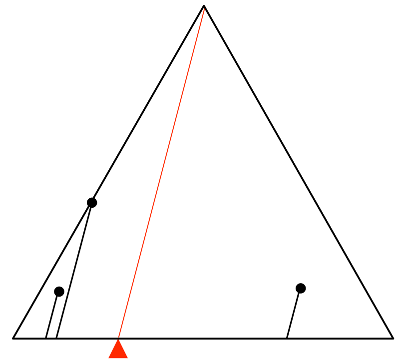
Fig. 9. Concordance is calculated with respect to the mean of the opinions, ω_μ

account the certainty of the beliefs. In general, each intermediate result should have a high concordance with its associated primary result. A large difference in their uncertainty acts to reduce the degree of difference in their expectations – providing the beliefs tend to project along similar projection paths.

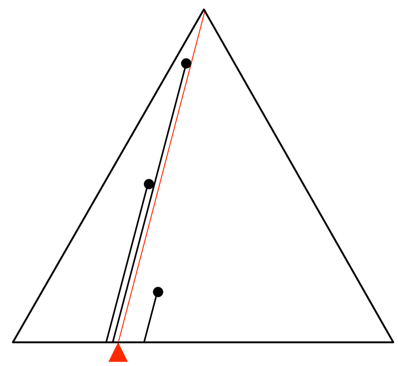
Within the broader context of intelligence analysis, concordance is a useful measure of the agreement between multiple information sources, concerning a single piece of evidence (i.e. $\omega^A_{e_j}, \omega^B_{e_j}, \ldots \omega^C_{e_j}$). The measure of their concordance provides an indication of their agreement, despite potentially having very different expectation values. Discussion and examples of the use of the concordance metric for this purpose will be discussed in later papers.

E. Consistency

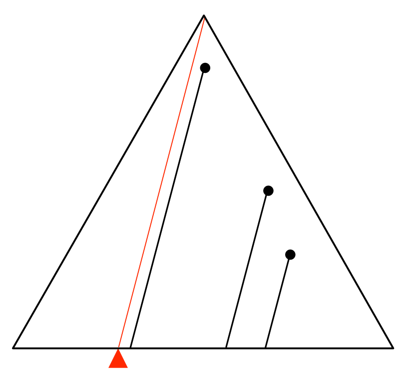
Consistency is a measure of the degree to which all of the intermediate analytical results for a single hypothesis either support or oppose the hypothesis. If all results support a hypothesis – or all oppose – the results are said to be completely consistent. If there is both support[11] and opposition, then there is at best only partial consistency. If the mean difference of support is low, then any consistency will tend to be somewhat high. Conversely, if the mean difference is high, then there will likely be low consistency.



(a) Extremely low consistency



(b) Medium consistency



(c) Complete consistency

Fig. 10. Consistency of a set of opinions depends on the tendency to support or oppose the hypothesis

The amount of support or opposition contributed by each of the intermediate analytic results $\omega_{h||e_i}$ will roughly be proportional to the relevance[12] of the evidence with respect to the hypothesis. Therefore, evidence with low relevance for the hypothesis will, at best, exhibit low support or opposition, while evidence with high relevance may exhibit high support or opposition – depending on the certainty and value

[11] see Section V-C

[12] See *Relevance* in the *Appendix* of [2]

of the evidence. Since the relevance of the logical conditionals $\omega_{h|e_i}$ and $\omega_{h|\bar{e}_i}$ will likely vary for each item of evidence, the amount of individual support or opposition will also likely vary.

Fig. 10 shows three sets of three opinions, each with different consistency. In Fig. 10(a), one opinion strongly supports the hypothesis, while two fairly strongly oppose it, leading to extremely low consistency. Fig. 10(b) exhibits a medium level of consistency – while the opinions are inconsistent in their support or opposition of the hypothesis, the mean difference of their expectations is very low. Complete consistency can be seen in Fig. 10(c), where all of the opinions support the hypothesis.

Consistency is a value in the range $\mathcal{C} \in [0, 1]$, with a value of $\mathcal{C} = 1$ indicating complete consistency (i.e. the beliefs all support or all oppose the hypothesis), and a value of $\mathcal{C} = 0$ indicating extreme inconsistency. The mathematical definition of consistency can be found in Appendix I.

Since for any non-trivial ACH problem, there are likely to be many items of evidence, an indication of how consistently the evidence supports or opposes the analytical result is very useful. In 'real world' problems with evidence of uncertain quality, unclear links between evidence and hypotheses, and many weakly diagnostic factors at play, the consistency of the intermediate analytic results will likely be suboptimal, with some small amount of inconsistency as a result. A high degree of inconsistency in the results suggests at least two interesting possibilities:

1) A high degree of inconsistency in combination with low concordance for one or more items of evidence could indicate possible deception or misperception. Examination of the sensitivity of the items of evidence – particularly items of evidence with low concordance – may reveal items of evidence of questionable reliability.

2) If more than one hypothesis has a high degree of inconsistency, this may also indicate that the set of chosen hypotheses has poor explanatory value for the observed evidence. The inconsistency may be due to an inadequate set of hypotheses, or a weak analytical model of the relationships between the hypotheses and evidence. Possible remedies might include considering a different or expanded set of hypotheses and/or evidence, and a careful examination of the conditionals within the analytical model that specify their relationship.

VI. Increasing the accuracy of evidence assessment

The objective of the ACH-SL process is to produce accurate analysis that is reflective of accuracy and certainty of the evidence provided. Accurate assessment of the evidence requires that the analyst correctly judges for each item of collected information, its currency, the reliability of its sources, and how well it supports the evidence.

A. Reliability of multiple or intermediary sources

The reliability of different information sources can vary widely, can change over time, and can also vary with the type of information that they provide. For example, a former mafia hitman-turned-informer may be considered a very reliable source of information for certain topics of interest; will be less reliable for others; and will be completely untrustworthy/unreliable for topics that might implicate a member of his family. Over time, the reliability of the informer to provide information on current topics will decay, until at some time – weeks or years – the informer is considered to not be a reliable source of current information on any topics of interest.

Information can also come from a variety of different types of sources. The simplest of these is a single source who is considered to be an original source of the information. For example, a witness to a robbery would be considered to be an original source of information about certain aspects of the robbery, as would others who were present at the time and place the robbery took place – including the robber. Even though the witnesses and the robber may all disagree about the details of the event, their knowledge or belief about the robbery is first-hand and not obtained through intermediaries. More complex types of sources include those who are intermediate sources, who are by definition not original sources. To greater or lesser degrees they filter, interpret, repackage and disseminate information from other sources – sometimes other intermediate sources – and the information relayed by them can sometimes be very different from that provided by their originators. It is important to note that a single entity can be both an original source for some information, and an intermediary source for other information.

In practice, discovering the original information and the original sources from the intermediate sources can be very difficult, if not impossible. The problem of using intermediate sources is the uncertainty introduced by their reliability. As the number of intermediate sources in the information chain grows, the greater becomes the potential for uncertainty in the information that is received. Fig. 11 shows how an item of information that originated with D is passed through intermediate sources of until it is delivered to the analyst A. If the information is transmitted via word-of-mouth, or filtered and repackaged like a report or news article, then the certainty of the information will decrease unless every intermediate source is completely reliable. Even if there is no obvious repackaging or filtering, and the information is passed though apparently untouched (as with paper-based documents) the possibility of deliberate tampering still remains.

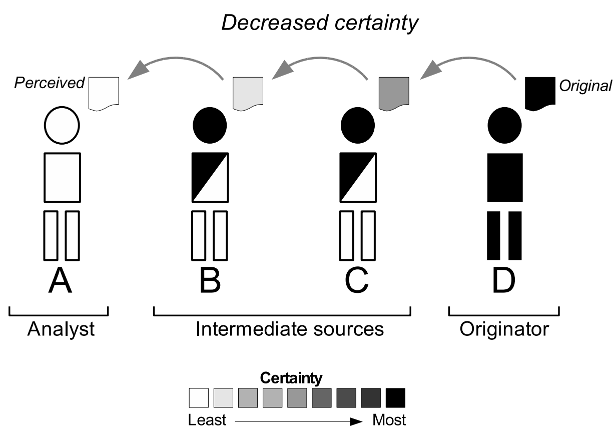


Fig. 11. Transitive information chains decreases the certainty of the received information

Even worse is the problem of a hidden common source

Ideally then, each item of evidence needs multiple independent sources of information. Having multiple sources increases the overall certainty of the received information, so long as the intermediate sources are not completely unreliable. Multiple opinions about an item of evidence can also be fused to provide a consensus opinion which reflects all of the opinions in a fair and equal way, and which serves to decrease its uncertainty. Fig. 12 shows how an item of information that originated with D is passes in parallel through two intermediate sources B and C – both of whom deliver it to the analyst, A. Even though the relatively poor reliability of B and C decrease the certainty of the information that each deliver, their fusion results in increased certainty when compared to the information individually provided by B and C.

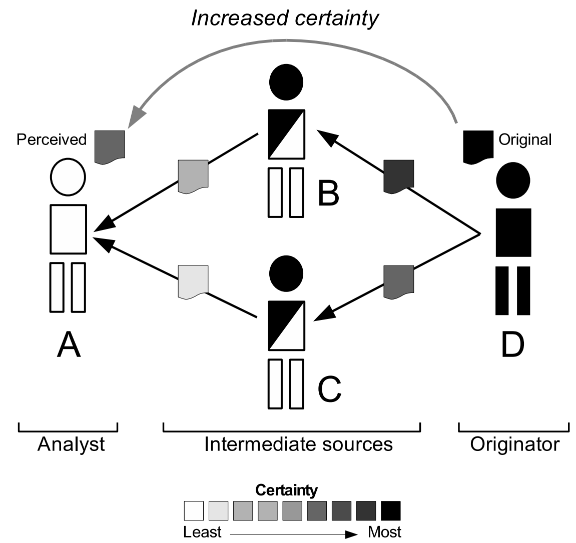


Fig. 12. Parallel paths increase the certainty of the received information

Subjective Logic can be used to address the issue of source reliability and multiple sources using the *discounting* [26], [27] and *consensus* [23] operators. Fig. 13 illustrates an example problem of reconciling information about an item of evidence e that is provided by three different sources with varying degrees of reliability. The figure shows three sources X, Y and Z providing information about the same topic. The task for the analyst is to arrive at an overall assessment as to the credibility and reliability of the information, so the information can be used as input into an existing ACH-SL analytical model. In Fig. 13, this overall assessment is represented by a 'virtual' source that provides a single piece of evidence e for use in the model.

The information provided by source X, e^X, is rated as "B-3" on the example 'Admiralty Scale' [13] in Table I, meaning that the source is considered

[13]For a brief discussion of the *Admiralty Scale*, see [28]

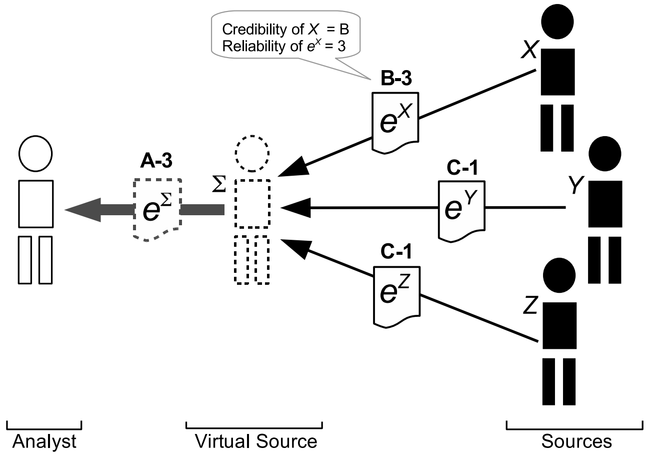


Fig. 13. Reconciling information of different value from sources of differing credibility results in a consensus value of the information and an combined source credibility rating.

to be "usually reliable", and the information they have provided *prima facie* implies that e is "likely". Similarly, the information provided by Y and Z (e^Y, e^Z) are both rated as "C-1". This means that the sources are considered to be "fairly reliable" and that each piece of information suggests that e is "almost certainly true".

With Subjective Logic, each of the categories in Table I can be mapped to a Subjective Opinion, and using the discount operator $\otimes$ [26], [27], an overall opinion can be obtained about the value and uncertainty of the information provided by each source, where the credibility of each piece of information is discounted by the reliability of its source, i.e.:

$$\begin{aligned}\mathrm{Cred}(X) \otimes \mathrm{Rel}(e^X) \rightarrow \omega_e^X &= (0.48, 0.32, 0.2, 0.5)\\ \mathrm{Cred}(Y) \otimes \mathrm{Rel}(e^Y) \rightarrow \omega_e^Y &= (0.57, 0.03, 0.4, 0.5)\\ \mathrm{Cred}(Z) \otimes \mathrm{Rel}(e^Z) \rightarrow \omega_e^Z &= (0.57, 0.03, 0.4, 0.5)\end{aligned}$$

The individual opinions about the information provided by each source can then be combined using the consensus operator $\oplus$ [23] to produce an overall opinion about the value and uncertainty of the evidence, given the opinion of the value and uncertainty of each piece of information provided, i.e.:

$$\omega_{\mathrm{e}}^X \oplus \omega_{\mathrm{e}}^Y \oplus \omega_{\mathrm{e}}^Z \rightarrow \omega_{\mathrm{e}}^{X\Diamond Y\Diamond Z} = (0.66, 0.22, 0.12, 0.5)$$

This single subjective opinion represents both the value of the evidence and the uncertainty associated with it. It can be input directly into the ACH-SL calculations, and it can be ascribed an overall semantic interpretation that represents a 'virtual' source with a single piece of information (Fig. 13). Using the method described in *Representations of subjective opinions* [2], this subjective opinion is given an interpretation of "A-3" indicating that the evidence is "likely" and that this assessment should be considered to be "almost always reliable", i.e. the certainty of this assessment is very high.

B. *Effect of belief decay*

Belief decay is important to model in situation-based decision making, where there is a temporal delay between observations and any decisions that are predicated on the probability expectation and certainty of those observations [29]. A belief about some continuing state of the world, based on an observation, will be less certain in the future if no additional

TABLE I

EXAMPLE SUBJECTIVE OPINION MAPPINGS FOR A 6×6 'ADMIRALTY SCALE'.

Credibility of Source *Collector's assessment of Source's reliability*			Reliability of Information *Source's assessment of the likelihood of information*		
Score	**Interpretation**	**Belief**[a]	**Score**	**Interpretation**	**Belief**[a]
A	Almost Always Reliable	$(0.95, 0.05, 0.0, a)$	**1**	Almost Certainly True	$(0.95, 0.05, 0.0, a)$
B	Usually Reliable	$(0.8, 0.2, 0.0, a)$	**2**	Very Likely	$(0.8, 0.2, 0.0, a)$
C	Fairly Reliable	$(0.6, 0.4, 0.0, a)$	**3**	Likely	$(0.6, 0.4, 0.0, a)$
D	Fairly Unreliable	$(0.4, 0.6, 0.0, a)$	**4**	Unlikely	$(0.4, 0.6, 0.0, a)$
E	Unreliable	$(0.1, 0.9, 0.0, a)$	**5**	Very Unlikely	$(0.1, 0.9, 0.0, a)$
F	Untested	$(0.0, 0.0, 1.0, a)$	**6**	Unknown	$(0.0, 0.0, 1.0, a)$

[a] The numbers in the parentheses constitute subjective logic beliefs in the form (b, d, u, a) – denoting *belief*, *disbelief*, *uncertainty* and *base rate*, respectively. For more information, please consult Appendix A.

observation is possible. Beliefs are therefore subject to *exponential decay*, where the rate of decay of a belief's certainty is proportional to its current certainty; and where the probability expectation of a decaying belief will approach its base rate, as its remaining certainty approaches zero.

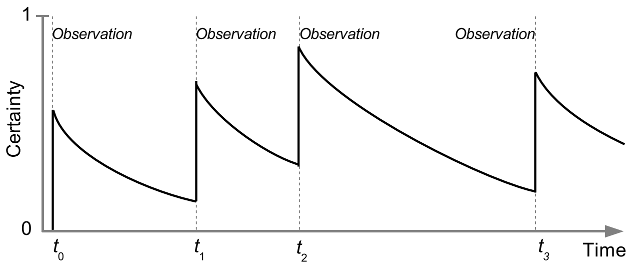


Fig. 14. Successive observations over time increase the certainty of decaying beliefs.

For example, if we observe that "John is at home" at 9:00am, we will in general be less certain that this is still true at 12:00pm, and far less certain at 3:00pm. In this way, we say that the certainty of our belief has *decayed* since the initial observation on which the belief was predicated. Of course, if we periodically observed John at his home over the course of the day, then the belief that "John is at home" over the entire course of the day will vary in certainty. Fig. 14 provides such an example where the resultant belief from an initial observation at time t_0 erodes with time, until succesively reinforced by subsequent observations at times t_1, t_2, and t_3.

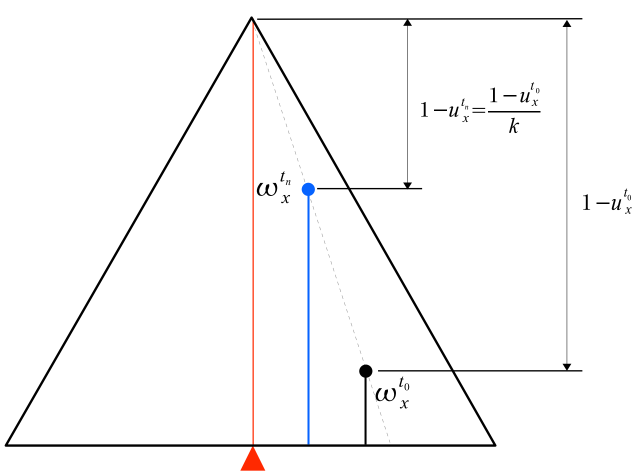


Fig. 15. A decaying belief ω_x at some time t_0, and later at t_n. The remaining certainty at $\omega_x^{t_n}$ is proportional to $k = e^{-\lambda(t_n - t_0)}$ of $\omega_x^{t_0}$.

In Fig. 15, an initial belief ω_x at time t_0 can be written as $\omega_x^{t_0}$, and the same belief with decayed certainty (i.e. increased uncertainty) at some time t_n can be written as $\omega_x^{t_n}$. The proportion of remaining certainty k after the decay of some interval time $t = t_n - t_0$ represents the tendency of the belief to become less certain with time where

$$k = e^{-\lambda t} = \frac{1 - u_x^{t_n}}{1 - u_x^{t_0}}$$

The temporal distance $t = t_n - t_0$, and the decay constant λ, acts to *discount* the original belief in the same way as a dogmatic opinion [22] with an expectation $\mathrm{E}(\omega_y) = e^{-\lambda t}$. We therefore can the express the decay of beliefs in a similar manner to the discounting of beliefs [26], [27]. The mathematical definintion of belief decay is provided in Appendix B.

C. *Source reliability revision*

The belief in the reliability of a source, whether it be a person, organization, sensor, or system does not remain static. It will usually change with the perceived reliability of the source, and decay according to the temporal distance between the time at which some judgement about the reliability of the source was possible, and the time at which an observation from that source is used for some decision.

Sources that provide reliable information now may not always do so in the future, and sources that were once unreliable may improve. Various factors can influence this, such as the upgrade of sensor or analytical systems; financial problems that motivate a source to provide poor quality information for quick financial gain; and better or worse access to valuable information as the result of a promotion or reorganization within an organization – to name just a few. Continuous belief revision is therefore necessary to ensure that the *perceived* reliability of a source reflects its *actual* reliability. Without it, the estimated reliability of a source will deviate more and more from its actual reliability, and therefore the information provided by the source cannot be safely judged.

Reputation systems [30] are ideally suited for representing and managing source reliability. A reputation system gathers and aggregates feedback about the behaviour of entities in order to derive measures of reputation [31]. While reputation systems typically collect ratings about members of a community from within the same community, the approach can also be used to manage reputation about any set of entities, using ratings provided by a closed community, such as an intelligence or law enforcement agency. Using a closed, trusted community also avoids most of the typical complaints about reputation systems, such as unfair ratings, retaliation, 'ballot box stuffing', and inadequate authentication [32].

In the case of a source's reliability, the feedback provided to the reputation system is the completeness and accuracy of the information provided, evaluated at some later time when there is sufficient certainty concerning its accuracy. This evaluation typically occurs independently of its use in any analysis and is an 'after the event' assessment. Complete and accurate information that is provided by the source increases the perceived reliability of the source, while incomplete or inaccurate information acts to decrease the perception of reliability.

It is an intuitive principle that recent observations are generally a better indicator of performance than are distant past observations [32] – all other things being equal. Therefore, the quality of recently provided information is generally a better indication of source reliability than that of information provided in the distant past. Simulations of markets that implement this principle in their reputation systems have been shown to facilitate honest trading, while those that do not remember past performance – or never forget past performance – have been shown to facilitate dishonest trading [32]. This implies that the perceived reliability of a source should naturally decay over time, increasing the uncertainty, and moving its reliability measure closer to the base rate - usually a neutral position. This ensures that more recently provided information will influence the derived measure of reliability more than old information. Further details are available in [31], [32].

VII. Conclusion

ACH-SL allows analysts to include weak diagnostic information that could be used to expose possible deceptions and to provide more balanced conclusions as to the expected outcomes.

Its use of Subjective Logic to represent belief in evidence and conclusions allows partial or complete uncertainty about evidence to be expressed both in the analytical model and the conclusions that follow. In addition, the Subjective Logic calculus allows information to be aggregated from different sources that have varying reliability, while the use of reputation systems can be used to capture and maintain source reliability and provides a formal way to model the reliability of sources – and therefore monitor and moderate the credibility of the information provided.

This paper is not intended as a complete solution to the problem of deception and misperception in analysis. While the metrics provided by ACH-SL can be very useful for detecting anomalies that may be indicative of of deception or misperception, further research needs to be performed to provide a high-level analysis of the likelihood of various types and locations of deception or misperception; and also account for the influence of an adversary's strategic goals in deciding how likely the anomalies are to be indicators of deception or misperception.

In addition, some of the other areas of importance have only been roughly sketched in this paper. The design of a reputation system for use in intelligence source management is an interesting and important research task. At first glance, such a task may seem simple, but a more studied viewing will reveal subtleties and complexities that make it far more difficult to actually design one for an intelligence collection or assessment organisation.

The same is also true for much of the concepts outlined in the paper. In order to make them useful, they need to be situated within an intelligence systems framework that allows sufficiently rich and flexible constructs – and yet that is simple and efficient enough for a human analyst to use without negative effects on cognitive load.

REFERENCES

[1] 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

[2] S. Pope and A. Jøsang, "Analsysis of competing hypotheses using subjective logic," in *Proceedings of the 10th International Command and Control Research and Technology Symposium*. United States Department of Defense Command and Control Research Program (DoDCCRP), 2005.

[3] B. Whaley, *Military Deception and Strategic Surprise*. Frank Cass, 1982, ch. Toward a General Theory of Deception.

[4] J. B. Bowyer, *Cheating*. St. Martins Press, 1982.

[5] J. B. Bell and B. Whaley, *Cheating and Deception*. Transaction Publishers, 1991.

[6] P. Johnson, S. Grazioli, K. Jamal, and R. G. Berryman, "Detecting deception: adversarial problem solving in a low base-rate world," *Cognitive Science*, vol. 3, pp. 355–392, May-June 2001.

[7] R. J. Heuer, "Strategic deception and counterdeception: A cognitive process approach," *International Studies Quarterly*, vol. 25, no. 2, pp. 294–327, June 1981.

[8] 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

[9] ——, "Deception detection by analysis of competing hypothesis," in *Proceedings of the 2005 International Conference on Intelligence Analysis*, 2005. [Online]. Available: http://analysis.mitre.org/proceedings/Final_Papers_Files/94_Camera_Ready_Paper.pdf

[10] H. Simon, *Models of Man: Social and National*. John Wiley, 1957.

[11] A. Tversky and D. Kahneman, "Judgement under uncertainty: Heuristics and biases," *Science*, vol. 125, pp. 1124–1131, 1974.

[12] ——, "The framing of decisions and the psychology of choice," *Science*, vol. 211, no. 4481, pp. 453–458, 1981.

[13] 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.

[14] K. J. Gilhooly, *Thinking: Directed, Undirected and Creative*. Academic Press, 1996.

[15] 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

[16] 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

[17] 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.

[18] 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.

[19] 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

[20] 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.

[21] 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

[22] 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.

[23] ——, "The consensus operator for combining beliefs," *Artificial Intelligence Journal*, vol. 142, no. 1–2, pp. 157–170, October 2002.

[24] K. Popper, *Objective Knowledge; An Evolutionary Approach*. Oxford University Press, 1972.

[25] A. Jøsang, S. Pope, and M. Daniel, "Conditional deduction under uncertainty," in *Proceedings of the 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2005)*, 2005.

[26] A. Jøsang and S. Pope, "Semantic constraints for trust tansitivity," in *Proceedings of the Asia-Pacific Conference of Conceptual Modelling (APCCM) (Volume 43 of Conferences in Research and Practice in Information Technology)*, S. Hartmann and M. Stumptner, Eds., Newcastle, Australia, February 2005.

[27] A. Jøsang, R. Hayward, and S. Pope, "Trust network analysis with subjective logic," in *Proceedings of the* 29^{th} *Australasian Computer Science Conference (ACSC2006), CRPIT Volume 48*, Hobart, Australia, January 2006.

[28] "Intelligence – human sources, the strategic fit," *AUSPOL*, no. 3, pp. 13–21, 2003.

[29] J. Yang, C. K. Mohan, K. G. Mehrotra, and P. K. Varshney, "A tool for belief updating over time in bayesian networks," in *14th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'02)*, November 2002, pp. 284–289. [Online]. Available: http://ieeexplore.ieee.org/iel5/8408/26514/01180816.pdf?tp=&arnumber=1180816&isnumber=26514

[30] A. Jøsang and R. Ismail, "The beta reputation system," in *Proceedings of the 15th Bled Electronic Commerce Conference*, Bled, Slovenia, June 2002.

[31] A. Jøsang, S. Hird, and E. Faccer, "Simulating the effect of reputation systems on e-markets," in *Proceedings of the First International Conference on Trust Management*, N. C., Ed., Crete, May 2003.

[32] A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision (to appear)," *Decision Support Systems*, vol. 00, no. 00, pp. 00–00, 2006.

[33] P. Smets, "The transferable belief model for quantified belief representation," in *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol.1*, D. Gabbay and P. Smets, Eds. Kluwer, Doordrecht, 1998, pp. 267–301.

APPENDIX
SUBJECTIVE LOGIC THEORY AND DEFINITIONS

A. 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 [22]. Belief calculus is suitable for approximate reasoning in situations where the probability of events can not be expressed with certainty, and is ideally suited for both human and machine representations of belief.

Belief theory applies to state spaces of possible exclusive events, and uses the notion of belief masses that are distributed over subsets of a state space, where the sum of belief masses so assigned must add up to one. A subset with belief mass is called a focal element of the state space. When all the belief mass is assigned to the whole state space, the belief distribution function is called *vacuous*, because it expresses no certainty about any event in particular. When all the focal elements are atomic events, the belief function is called Bayesian, which corresponds to a classical probability model. When no belief mass is assigned to the whole state space, the belief function is *dogmatic* [33]. Let us note, that trivially, every Bayesian belief function is dogmatic, and that in case of binary state spaces, every dogmatic belief function is Bayesian.

Subjective Logic [22] 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 the truth of statement x. Here b, d, and u represent belief, disbelief and uncertainty, and base rate [14] respectively where $b_x^A, d_x^A, u_x^A, a_x^A \in [0,1]$ and the following equation holds:

$$b_x^A + d_x^A + u_x^A = 1\ . \tag{A-1}$$

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 base rate is used for computing an opinion's probability expectation value expressed by:

$$\mathrm{E}(\omega_x^A) = b_x^A + a_x^A u_x^A\ , \tag{A-2}$$

meaning that a determines how uncertainty shall contribute to $\mathrm{E}(\omega_x^A)$. 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. 16 illustrates an example where the opinion about a proposition x from a binary frame of discernment has the value $\omega_x = (0.7,\ 0.1,\ 0.2,\ 0.5)$.

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

[14] Also referred to as *relative atomicity* in [22] and in other later papers

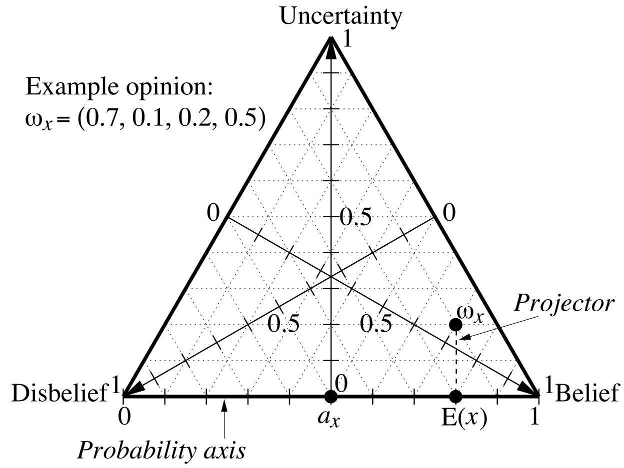


Fig. 16. Opinion triangle with example opinion

disbelief and belief vertices is called the probability axis. The base rate 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 base rate 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 [22]:

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 $\mathrm{beta}(\alpha, \beta)$ can be expressed using the gamma function Γ as:

$$\mathrm{beta}(\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}\, p^{\alpha-1}(1-p)^{\beta-1} \tag{A-3}$$

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:

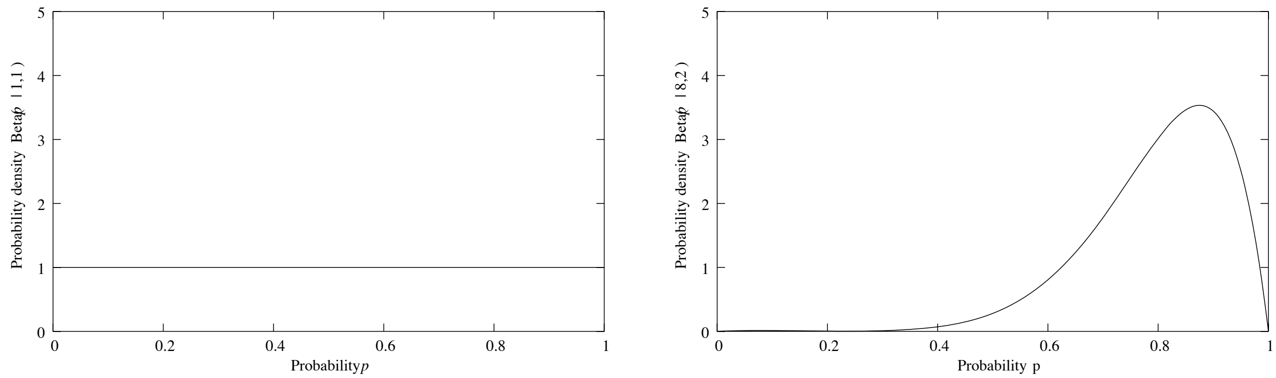
$$\mathrm{E}(p) = \alpha/(\alpha + \beta). \tag{A-4}$$

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

$$(b_x, d_x, u_x, a_x) \longmapsto \mathrm{beta}\left(\frac{2b_x}{u_x} + 2a_x,\ \ \frac{2d_x}{u_x} + 2(1 - a_x)\right) \tag{A-5}$$

This means for example that an opinion with $u_x = 1$ and $a_x = 0.5$ which maps to $\mathrm{beta}\,(1, 1)$ is equivalent to a uniform PDF. It also means that a dogmatic opinion with $u_x = 0$ which maps to $\mathrm{beta}\,(b_x\eta,\ d_x\eta)$ where $\eta \to \infty$ 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. 17a. 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. 17b below. This corresponds to the opinion of Fig. 16 through the mapping of (A-5).



(a) Uniform beta PDF: beta(1,1)

(b) Uniform beta PDF: beta(8,2)

Fig. 17. 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. 17b. is $\mathrm{E}(p) = 0.8$. 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 (A-3) is a probability variable, so that for a given p the probability density $\mathrm{beta}(\alpha, \beta)$ represents second order probability. The first-order variable p represents the probability of an event, whereas the density $\mathrm{beta}(\alpha, \beta)$ 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.

Subjective Logic includes the standard set of operators that are used for probability calculus and binary logic; specialized operators for belief fusion and discounting; and additional functions for analyzing directed graphs of subjective opinions.

Within ACH-SL, the value and certainty of evidence and hypotheses are represented as subjective opinions which can be transformed to and from other representations of belief. These representations include 'fuzzy' terms used by humans, as well as various machine representations, including Bayesian representations.

B. Decay

Definition 1.1 (Decay Constant): Let some initial belief ω_x at time t_0 can be written as $\omega_x^{t_0}$. Let the same belief with decayed certainty (i.e. increased uncertainty) at some time t_n be written as $\omega_x^{t_n}$. The proportion of remaining certainty k after the decay of some interval time $t = t_n - t_0$ can be calculated as

$$k = \frac{1 - u_x^{t_n}}{1 - u_x^{t_0}} \tag{B-1}$$

and therefore the decay constant $\lambda \geq 0$ for exponential decay is

$$k = e^{-\lambda t} \tag{B-2}$$

which can also be written as

$$\lambda = \frac{-\ln k}{t} \tag{B-3}$$

Proof: Since the certainty is subject to exponential decay, then the rate of decay of a belief ω_x must be proportional to its certainty $(1 - u_x^{t_0})$, and is expressed by the following differential equation, where $\lambda \geq 0$ is the decay constant

$$\frac{du}{dt} = \lambda(1 - u_x^{t_0}) \tag{B-4}$$

which has the solution

$$\begin{aligned} 1 - u_x^{t_n} &= (1 - u_x^{t_0})\, e^{-\lambda t} \\ \frac{1 - u_x^{t_n}}{1 - u_x^{t_0}} &= e^{-\lambda t} \end{aligned}$$

Substituting k from B-1 gives

$$k = e^{-\lambda t}$$

■

Definition 1.2 (Instantaneous Decay): The instantaneous decay of a belief $\omega_x = (b_x, d_x, u_x, a_x)$ to a relative certainty of k (B-1) produces the belief $\omega_{x'} = (b_{x'}, d_{x'}, u_{x'}, a_{x'})$ where

$$\text{Decay}(\omega_x, k) \rightarrow \omega_{x'} = \begin{cases} b_{x'} = & kb_x \\ d_{x'} = & kd_x \\ u_{x'} = & 1 + ku_x - k \\ a_{x'} = & a_x \end{cases} \tag{B-5}$$

Remark 1.3: Given a decaying belief ω_x with a decay constant of λ, there must be some time $t_{1/2}$, when the remaining certainty $(1 - u_x^{t^{1/2}})$ is half of the certainty of the original belief $(1 - u_x^{t_0})$, i.e.

$$1 - u_x^{t_{1/2}} = \frac{1}{2}\,(1 - u_x^{t_0}) \tag{B-6}$$

Definition 1.4 (Half-life): The half-life is the length of time $t_{1/2}$ that the certainty of a belief decays (with the decay constant λ) before there remains one half of the certainty of the original belief, and is given by

$$t_{1/2} = \frac{\ln 2}{\lambda} \tag{B-7}$$

Corollary 1.5 (Decay Constant from Half-life): The decay constant λ can be calculated from the half-life, $t_{1/2}$ and is given by

$$\lambda = \frac{\ln 2}{t_{1/2}} \tag{B-8}$$

C. Projection

Decay always produces a belief with the same or decreased certainty at some future point in time. The complement of decay is *projection*, which is agnostic with respect to time, and projects the existing opinion ω_x to a position of either increased or decreased certainty $\omega_{x'} \to u_{x'} = u_\gamma$ along the 'path of least resistance' defined by the point at the top vertex of the triangle, $(0, 0, 1, a_x)$, and the opinion (b_x, d_x, u_x, a_x) (see Fig. 18).

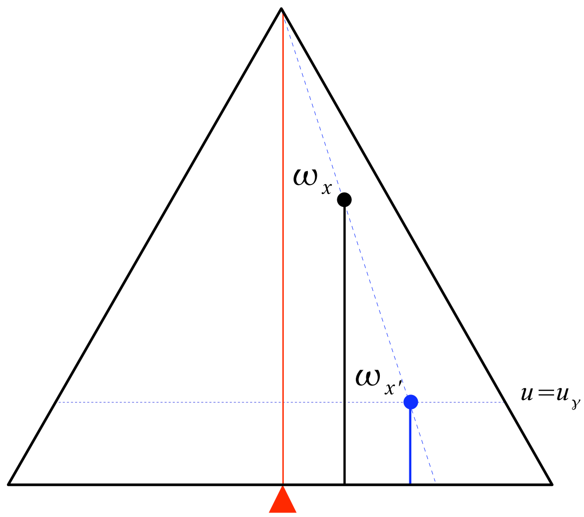


Fig. 18. An opinion ω_x with uncertainty u_x projected to a new opinion $\omega_{x'}$ with a target uncertainty of $u_{x'} = u_\gamma$

Projection is needed to approximate the past value of a decayed belief, to a specified level of certainty. Like decay, projection treats both b_x and d_x fairly, and produces an opinion that maintains their relative sizes. In the special case of the opinion being completely uncertain $u_x = 1$, both $b_x = d_x = 0$, and the projected opinion will lie along the line $\mathrm{E}(\omega_{x'}) = a_x$.

Definition 1.6 (Projection): The projection of a subjective opinion $\omega_x = (b_x, d_x, u_x, a_x)$ to some target uncertainty $u_\gamma \in [0, 1]$ produces a new opinion $\omega_y = (b_{x'}, d_{x'}, u_{x'}, a_{x'})$ as defined by

$$\mathrm{Proj}(\omega_x, u_\gamma) \to \omega_{x'} = \begin{cases} u_x = 1, & \begin{cases} b_{x'} = & a_x(1 - u_\gamma) \\ d_{x'} = & (1 - a_x)(1 - u_\gamma) \\ u_{x'} = & u_\gamma \\ a_{x'} = & a_x \end{cases} \\ & \\ u_x < 1, & \begin{cases} b_{x'} = & b_x + \frac{b_x}{b_x + d_x}(u_x - u_\gamma) \\ d_{x'} = & b_x + \frac{d_x}{b_x + d_x}(u_x - u_\gamma) \\ u_{x'} = & u_\gamma \\ a_{x'} = & a_x \end{cases} \end{cases} \qquad \text{(C-1)}$$

D. Mean of a set of opinions

This mean, μ, of a set of a set of k probabilities is, in effect, an unweighted compromise of all probabilities in the input set and its formula is simply

$$\mu = \frac{\sum_{i=1}^{k} \mathrm{p}(x_i)}{k} \qquad \text{(D-1)}$$

Similarly, the mean of a set of Subjective Opinions is the opinion that represents a compromise of all opinions in the input set, equally weighted with respect to the certainty, expectation, and base rate of each opinion. It is unsuited for belief fusion since it is an unweighted measure, and should not be confused with the *consensus* operator [23] that is used for belief fusion.

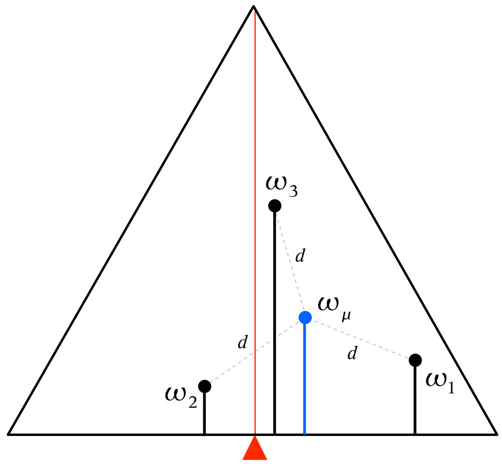


Fig. 19. The mean of two or three opinions with the same base rate is point with the smallest equidistant distance from each

The effects of the operator are portrayed in Fig. 19, where the resultant mean opinion ω_μ is an opinion whose expectation, certainty, and base rate values are respectively the mean of the expectation, certainty and base rate values of the input opinions $(\omega_1, \omega_2, \omega_3)$. Visually, for two or three opinions that share a common base rate, the mean opinion will lie equidistant from each of them in the opinion triangle. In the example in Fig. 19, the mean opinion lies at a distance d from each of the input opinions. For more than three opinions – or where the base rates are different – no such visual symmetry is guaranteed.

Definition 1.7 (Mean): Let $\Theta = \{\omega_1, \omega_2, \ldots \omega_k,\}$ represent a set of $k > 1$ opinions. Then the mean opinion of Θ, $\omega_\mu = (b_\mu, d_\mu, u_\mu, a_\mu)$ where

$$\mathrm{E}(\omega_\mu) = \frac{\sum_{i=1}^{k} \mathrm{E}(\omega_i)}{k},\ u_\mu = \frac{\sum_{i=1}^{k} u_i}{k},\ a_\mu = \frac{\sum_{i=1}^{k} a_i}{k} \tag{D-2}$$

and therefore is defined as

$$\omega_\mu = \begin{cases} b_\mu & = \frac{\sum_{i=1}^{k}(b_i + a_i u_i) - \sum_{i=1}^{k} a_i \cdot \sum_{i=1}^{k} u_i}{k} \\ d_\mu & = 1 - \frac{\sum_{i=1}^{k}(b_i + a_i u_i) - \sum_{i=1}^{k} a_i \cdot \sum_{i=1}^{k} u_i + \sum_{i=1}^{k} u_i}{k} \\ u_\mu & = \frac{\sum_{i=1}^{k} u_i}{k} \\ a_\mu & = \frac{\sum_{i=1}^{k} a_i}{k} \end{cases} \tag{D-3}$$

E. Diagnosticity

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

Definition 1.8 (Diagnosticity of evidence): Let $\Phi = \{h_1, h_2, \ldots h_k\}$ be a state space for a set k hypotheses where one and only one $h_i \in \Phi$ is true. Let $\Omega^\Phi = \{\Theta_1, \Theta_2, \ldots \Theta_m\}$ be the corresponding set of m state spaces for a single item of evidence, e (where $\Theta_i = \{e_i, \overline{e}_i\}, \Theta_i \in \Omega^\Phi$) that represent the conditionals $\omega_{h_i|e}, \omega_{h_i|\overline{e}}$ for each hypothesis $h_i \in \Phi$. Then we define the *diagnosticity* of the evidence e with respect to an arbitrary subset of hypotheses $H \subseteq \Phi$ with $k > 0$ elements to be

$$\mathrm{D}(e,H) = \begin{cases} \mathrm{E}_{\text{total}}(e,H) = 0, & 0; \\ \mathrm{E}_{\text{total}}(e,H) > 0, & \frac{\sum_{n=1}^{k} \left| \mathrm{E}(\omega_{h_n|e}) - \mathrm{E}(\omega_{h_n|\bar{e}}) - \mathrm{D}_{\text{mean}}(e,H) \right|}{\mathrm{E}_{\text{total}}(e,H)} \end{cases} \tag{E-1}$$

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

$$\mathrm{D}_{\text{mean}}(e,H) = \frac{\sum_{n=1}^{k} \left[\mathrm{E}(\omega_{h_n|e}) - \mathrm{E}(\omega_{h_n|\bar{e}}) \right]}{k} \tag{E-2}$$

$$\mathrm{E}_{\text{total}}(e,H) = \sum_{n=1}^{k} \left[\mathrm{E}(\omega_{h_n|e}) + \mathrm{E}(\omega_{h_n|\bar{e}}) \right] \tag{E-3}$$

Then the *diagnosticity* of the evidence e with respect to an arbitrary subset of hypotheses H can be rewritten as (substituting E-2 and E-3 into E-1):

$$\mathrm{D}(e,H) = \frac{\sum_{n=1}^{k} \left| \mathrm{E}(\omega_{h_n|e}) - \mathrm{E}(\omega_{h_n|\bar{e}}) - \sum_{n=1}^{k} \left[\frac{\mathrm{E}(\omega_{h_n|e}) - \mathrm{E}(\omega_{h_n|\bar{e}})}{k} \right] \right|}{\sum_{n=1}^{k} \left[\mathrm{E}(\omega_{h_n|e}) + \mathrm{E}(\omega_{h_n|\bar{e}}) \right]} \tag{E-4}$$

Remark 1.9: It can be seen that $\mathrm{D}(e,H) \in [0..1]$ 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. $\mathrm{E}(\omega_e) = 0 \vee \mathrm{E}(\omega_e) = 1$), 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 1.10 (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

$$\mathrm{D}(e,\Phi) = \frac{\sum_{n=1}^{m} \left| \mathrm{E}(\omega_{h_n|e}) - \mathrm{E}(\omega_{h_n|\bar{e}}) \right|}{2} \tag{E-5}$$

F. Sensitivity

The relative degree to which the the value of a hypothesis (or a subset of hypotheses) will be determined by the value of an item of evidence is known as its *sensitivity*, and is useful in sensitivity analysis for quickly identifying the items of evidence that will likely have the greatest impact. It is calculated for a subset of hypotheses and an item of evidence as its relative *potential energy* – which is defined as the product of the difference in the expectation of the conditionals, and the mean of the squares of their certainty (see Fig. 20).

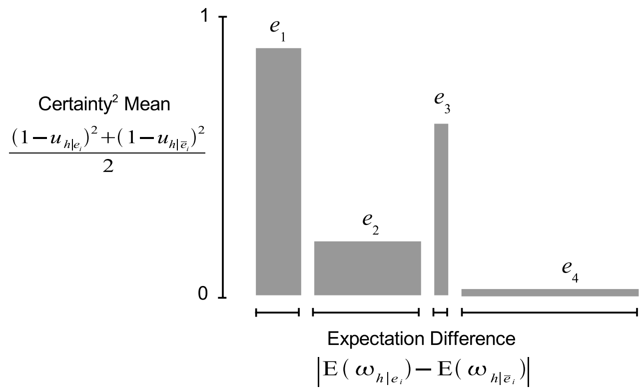


Fig. 20. The sensitivity of a hypothesis with respect to an item of evidence is the relative product of the mean of the squares of their certainties and expectation difference of their causal logical conditionals

Definition 1.11 (Potential energy of a pair of logical conditionals): Let $\Phi = \{h_1, h_2, \ldots h_m\}$ be a state space for a set m hypotheses where one and only one $h_i \in \Phi$ is true. Let e be an item of evidence and $\Omega^{\Phi} = \{\Theta_1, \Theta_2, \ldots \Theta_n\}$ be the corresponding set of n state spaces (where $\Theta_j = \{e_j, \bar{e}_j\}, \Theta_j i \in \Omega^{\Phi}$) that represent the conditionals $\omega_{h_i|e}, \omega_{h_i|\bar{e}}$ for each hypothesis $h_i \in \Phi$. Then we define the *potential energy* of an arbitrary subset of hypotheses $H \subseteq \Phi$ with $k > 0$ elements with respect to an item of evidence e to be

$$\mathrm{U_p}(H, e) = \sum_{i=1}^{k} \left[\frac{\left|\mathrm{E}(\omega_{h_i|e}) - \mathrm{E}(\omega_{h_i|\bar{e}})\right| \cdot \left[(1 - u_{h_i|e})^2 + (1 - u_{h_i|\bar{e}})^2\right]}{2} \right] \tag{F-1}$$

Definition 1.12 (Sensitivity of a result): Let $\Phi = \{h_1, h_2, \ldots h_m\}$ be a state space for a set m hypotheses where one and only one $h_i \in \Phi$ is true. Let $\Psi = \{e_1, e_2, \ldots e_n\}$ be a set of n items of evidence. Let $\mathrm{U_p}(H, e_i)$ be the potential energy (F-1) of an arbitrary subset of hypotheses $H \subseteq \Phi$ with $k > 0$ elements with respect to an item of evidence $e_i \in \Phi$. Then we define the *sensitivity* of the subset of hypotheses H with respect to a single item of evidence $e_j \in \Psi$, relative to Ψ to be

$$\mathrm{S_{rel}}(H, e_j, \Psi) = \frac{\mathrm{U_p}(H, e_j)}{\sum_{i=1}^{n} \mathrm{U_p}(H, e_i)} \tag{F-2}$$

G. Support

Definition 1.13 (Support): Let x be some proposition and $\omega_x = \{b_x, d_x, u_x, a_x\}$ be an opinion about x. Then the degree to which the opinion supports the proposition x is

$$\mathcal{S}_{\omega_x} = \mathrm{E}(\omega_x) - a_x \tag{G-1}$$

H. Concordance

Definition 1.14 (Concordance): Let $\Theta = \{\omega_1, \omega_2, \ldots \omega_k,\}$ represent a set of $k > 1$ opinions. Let ω_μ represent the mean opinion of Θ. Let $\Theta' = \{\omega'_1, \omega'_2, \ldots \omega'_k,\}$ be the projections (C-1) of the elements of Θ to uncertainty u_μ, i.e. $\forall i, \omega_{i'} = \mathrm{Proj}(\omega_i, u_\mu)$. Then the concordance of Θ, $\mathcal{K}_\Theta$, is

$$\mathcal{K}_\Theta = 1 - \frac{\sum_{i=1}^{k} |\mathrm{E}(\omega_i) - \mathrm{E}(\omega_\mu)| + \sum_{i=1}^{k} |\mathrm{E}(\omega'_i) - \mathrm{E}(\omega_\mu)|}{k} \tag{H-1}$$

I. Consistency

Definition 1.15 (Consistency of Analytical Result): Let h be some hypothesis. Let $\Psi = \{e_1, e_2, \ldots e_m\}$ be the set of m items of evidence. Let the intermediate analytic results for the hypothesis be $h||\mathcal{E} = \{\omega_{h||e_1}, \omega_{h||e_2}, \ldots \omega_{h||e_m}\}$, and the corresponding set of *support* values (G-1) for each $\omega_{h||e_i} \in h||\mathcal{E} \mapsto \mathcal{S}_{\omega_{h||e_i}} \in \Upsilon = \{\mathcal{S}_{\omega_{h||e_1}}, \mathcal{S}_{\omega_{h||e_2}}, \ldots \mathcal{S}_{\omega_{h||e_m}}\}$. Let S_μ be the mean of the support (G-1) of Υ, as defined by

$$\mathrm{S}_\mu = \frac{\sum_{i=1}^{m} \mathcal{S}_{\omega_{h||e_i}}}{m} \tag{I-1}$$

Then the consistency of h with respect to $\mathcal{E}$ is

$$\mathcal{C}_{h||\mathcal{E}} = \frac{\left|\sum_{i=1}^{m} \mathcal{S}_{\omega_{h||e_i}}\right| + 1 - \mathrm{S}_\mu}{\sum_{i=1}^{m} \left|\mathcal{S}_{\omega_{h||e_i}}\right| + 1 - \mathrm{S}_\mu} \tag{I-2}$$

which in full form is written as

$$\mathcal{C}_{h||\mathcal{E}} = \frac{m\left|\sum_{i=1}^{m} \mathcal{S}_{\omega_{h||e_i}}\right| + m - \sum_{i=1}^{m} \mathcal{S}_{\omega_{h||e_i}}}{m\sum_{i=1}^{m} \left|\mathcal{S}_{\omega_{h||e_i}}\right| + m - \sum_{i=1}^{m} \mathcal{S}_{\omega_{h||e_i}}} \tag{I-3}$$