

The Base Rate Fallacy in Belief Reasoning

Audun Jøsang
University of Oslo
josang@unik.no

Stephen O'Hara
21CSI, USA
sohara@21csi.com

Abstract – *The base rate fallacy is an error that occurs when the conditional probability of some hypothesis y given some evidence x is assessed without taking account of the "base rate" of y , often as a result of wrongly assuming equality between the inverse conditionals, expressed as: $p(y|x) = p(x|y)$. Because base rates and conditionals are not normally considered in traditional Dempster-Shafer belief reasoning there is a risk of falling victim to the base rate fallacy in practical applications. This paper describes the concept of the base rate fallacy, how it can emerge in belief reasoning and possible approaches for how it can be avoided.*

Keywords: Belief reasoning, base rates, base rate fallacy, subjective logic

1 Introduction

Conditional reasoning is used extensively in areas where conclusions need to be derived from evidence, such as in medical reasoning, legal reasoning and intelligence analysis. Both binary logic and probability calculus have mechanisms for conditional reasoning. In binary logic, Modus Ponens (MP) and Modus Tollens (MT) are the classical operators which are used in any field of logic that requires conditional deduction. In probability calculus, binomial conditional deduction is expressed as:

$$p(y||x) = p(x)p(y|x) + p(\bar{x})p(y|\bar{x}) \quad (1)$$

where the terms are interpreted as follows:

- $p(y|x)$: conditional probability of y given x is TRUE
- $p(y|\bar{x})$: conditional probability of y given x is FALSE
- $p(x)$: probability of the antecedent x
- $p(\bar{x})$: probability of the antecedent's complement
(= 1 - $p(x)$)
- $p(y||x)$: deduced probability of the consequent y

The notation $y||x$, introduced in [1], denotes that the truth or probability of proposition y is deduced as a function of the probability of the antecedent x together with the conditionals. The expression $p(y||x)$ thus represents a derived value,

whereas the expressions $p(y|x)$ and $p(y|\bar{x})$ represent input values together with $p(x)$. Below, this notational convention will also be used for belief reasoning.

The problem with conditional reasoning and the base rate fallacy can be illustrated by a simple medical example. Assume a pharmaceutical company that develops a test for a particular infectious disease will typically determine the reliability of the test by letting a group of infected and a group of non-infected people undergo the test. The result of these trials will then determine the reliability of the test in terms of its *sensitivity* $p(x|y)$ and *false positive rate* $p(x|\bar{y})$, where x : "Positive Test", y : "Infected" and \bar{y} : "Not infected". The conditionals are interpreted as:

- $p(x|y)$: "Probability of positive test given infection"
- $p(x|\bar{y})$: "Probability of positive test in the absence of infection".

The problem with applying these reliability measures in a practical setting is that the conditionals are expressed in the opposite direction to what the practitioner needs in order to apply the expression of Eq.(1). The conditionals needed for making the diagnosis are:

- $p(y|x)$: "Probability of infection given positive test"
- $p(y|\bar{x})$: "Probability of infection given negative test"

but these are usually not directly available to the medical practitioner. However, they can be obtained if the base rate of the infection is known.

The base rate fallacy [2] consists of making the erroneous assumption that $p(y|x) = p(x|y)$. Practitioners often fall victim to this the reasoning error, e.g. in medicine, legal reasoning (where it is called the prosecutor's fallacy) and in intelligence analysis. While this reasoning error often can produce a relatively good approximation of the correct diagnostic probability value, it can lead to a completely wrong result and wrong diagnosis in case the base rate of the disease in the population is very low and the reliability of the test is not perfect.

This paper describes how the base rate fallacy can occur in belief reasoning, and puts forward possible approaches for how this reasoning error can be avoided.

2 Reasoning with Conditionals

Assume a medical practitioner who, from the description of a medical test, knows the sensitivity $p(x|y)$ and the false positive rate $p(x|\bar{y})$ of the test, where x : “Positive Test”, y : “Infected” and \bar{y} : “Not infected”.

The required conditionals can be correctly derived by inverting the available conditionals using Bayes rule. The inverted conditionals are obtained as follows:

$$\begin{cases} p(x|y) = \frac{p(x \wedge y)}{p(y)} \\ p(y|x) = \frac{p(x \wedge y)}{p(x)} \end{cases} \Rightarrow p(y|x) = \frac{p(y)p(x|y)}{p(x)}. \quad (2)$$

On the right hand side of Eq.(2) the base rate of the disease in the population is expressed by $p(y)$. By applying Eq.(1) with x and y swapped in every term, the expected rate of positive tests $p(x)$ in Eq.(2) can be computed as a function of the base rate $p(y)$. This is expressed in Eq.(3) below.

$$p(x) = p(y)p(x|y) + p(\bar{y})p(x|\bar{y}) \quad (3)$$

In the following, $a(x)$ and $a(y)$ will denote the base rates of x and y respectively. The required positive conditional is:

$$p(y|x) = \frac{a(y)p(x|y)}{a(y)p(x|y) + a(\bar{y})p(x|\bar{y})}. \quad (4)$$

In the general case where the truth of the antecedent is expressed as a probability, and not just binary TRUE and FALSE, the opposite conditional is also needed as specified in Eq.(1). In case the negative conditional is not directly available, it can be derived according to Eq.(4) by swapping x and \bar{x} in every term. This produces:

$$\begin{aligned} p(y|\bar{x}) &= \frac{a(y)p(\bar{x}|y)}{a(y)p(\bar{x}|y) + a(\bar{y})p(\bar{x}|\bar{y})} \\ &= \frac{a(y)(1-p(x|y))}{a(y)(1-p(x|y)) + a(\bar{y})(1-p(x|\bar{y}))}. \end{aligned} \quad (5)$$

The combination of Eq.(4) and Eq.(5) enables conditional reasoning even when the required conditionals are expressed in the reverse direction to what is needed.

A medical test result is typically considered positive or negative, so when applying Eq.(1), it can be assumed that either $p(x) = 1$ (positive) or $p(\bar{x}) = 1$ (negative). In case the patient tests positive, Eq.(1) can be simplified to $p(y|x) = p(y|y)$ so that Eq.(4) will give the correct likelihood that the patient actually has contracted the disease.

2.1 Example 1: Probabilistic Medical Reasoning

Let the sensitivity of a medical test be expressed as $p(x|y) = 0.9999$ (i.e. an infected person will test positive in 99.99% of the cases) and the false positive rate be

$p(x|\bar{y}) = 0.001$ (i.e. a non-infected person will test positive in 0.1% of the cases). Let the base rate of infection in population A be 1% (expressed as $a(y_A)=0.01$) and let the base rate of infection in population B be 0.01% (expressed as $a(y_B)=0.0001$). Assume that a person from population A tests positive, then Eq.(4) and Eq.(1) lead to the conclusion that $p(y_A|x) = p(y_A|y) = 0.9099$ which indicates a 91% likelihood that the person is infected. Assume that a person from population B tests positive, then $p(y_B|x) = p(y_B|y) = 0.0909$ which indicates only a 9% likelihood that the person is infected. By applying the correct method the base rate fallacy is avoided in this example.

2.2 Example 2: Intelligence Analysis

A typical task in intelligence analysis is to assess and compare a set of hypotheses based on collected evidence. Stech and Elässer of the MITRE corporation describe how analysts can err, and even be susceptible to deception, when determining conditionals based on a limited view of evidence [3].

An example of this is how the reasoning about the detection of Krypton gas in a middle-eastern country can lead to the erroneous conclusion that the country in question likely has a nuclear enrichment program. A typical ad-hoc approximation is:

$$\begin{aligned} p(\text{Krypton} | \text{enrichment}) &= \text{high} \\ \Downarrow \\ p(\text{enrichment} | \text{Krypton}) &= \text{high} \end{aligned} \quad (6)$$

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

$$p(\text{Krypton} | \text{not enrichment}) = \text{medium to high} \quad (7)$$

so that a more correct approximation is:

$$\begin{cases} p(\text{Krypton} | \text{enrichment}) = \text{high} \\ p(\text{Krypton} | \text{not enrichment}) = \text{medium to high} \end{cases} \quad (8)$$

$$\Downarrow$$

$$p(\text{enrichment} | \text{Krypton}) = \text{uncertain}$$

This additional information should lead the analyst to the conclusion that there is a fair amount of uncertainty of a nuclear program given the detection of Krypton. The assignment of the ‘high’ value to **p(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.

2.3 Deductive and Abductive Reasoning

The term *frame*¹ will be used with the meaning of a traditional state space of mutually disjoint states. We will use the

¹Usually called *frame of discernment* in traditional belief theory

term “parent frame” and “child frame” to denote the reasoning direction, meaning that the parent frame is what the analyst has evidence about, and probabilities over the child frame is what the analyst needs. Defining parent and child frames is thus equivalent with defining the direction of the reasoning.

Forward conditional inference, called *deduction*, is when the parent frame is the antecedent and the child frame is the consequent of the available conditionals. Reverse conditional inference, called *abduction*, is when the parent frame is the consequent, and the child frame is the antecedent.

Causal and derivative reasoning situations are illustrated in Fig.1 where x denotes the cause state and y denotes the result state. Causal conditionals are expressed as $p(\text{effect} | \text{cause})$.

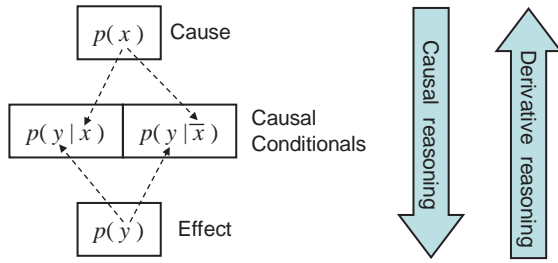


Figure 1: Visualizing causal and derivative reasoning

The concepts of “causal” and “derivative” reasoning can be meaningful for obviously causal conditional relationships between states. By assuming that the antecedent causes the consequent in a conditional, then causal reasoning is equivalent to deductive reasoning, and derivative reasoning is equivalent to abductive reasoning. Derivative reasoning requires derivative conditionals expressed as $p(\text{cause} | \text{effect})$. Causal reasoning is much simpler than derivative reasoning because it is much easier to estimate causal conditionals than derivative conditionals. However, a large part of human intuitive reasoning is derivative.

In medical reasoning for example, the infection causes the test to be positive, not the other way, so derivative reasoning must be applied to estimate the likelihood of being affected by the disease. The reliability of medical tests is expressed as causal conditionals, whereas the practitioner needs to apply the derivative inverted conditionals. Starting from a positive test to conclude that the patient is infected therefore represents derivative reasoning. Most people have a tendency to reason in a causal manner even in situations where derivative reasoning is required. In other words, derivative situations are often confused with causal situations, which provides an explanation for the tendency of the base rate fallacy in medical diagnostics, legal reasoning and intelligence analysis.

3 Review of Belief-Based Conditional Reasoning

In this section, previous approaches to conditional reasoning with beliefs and related frameworks are briefly reviewed.

3.1 Smets’ Disjunctive Rule and Generalized Bayes Theorem

An early attempt at articulating belief-based conditional reasoning was provided by Smets (1993) [4] and by Xu & Smets [5, 6]. This approach is based on using the so-called Generalized Bayes Theorem as well as the Disjunctive Rule of Combination, both of which are defined within the Dempster-Shafer belief theory.

In the binary case, Smets’ approach assumes a conditional connection between a binary parent frame Θ and a binary child frame X defined in terms of belief masses and conditional plausibilities. In Smets’ approach, binomial deduction is defined as:

$$\begin{aligned}
 pl(x) &= m(\theta)pl(x|\theta) + m(\bar{\theta})pl(x|\bar{\theta}) \\
 &\quad + m(\Theta)(1 - (1 - pl(x|\theta))(1 - pl(x|\bar{\theta}))) \\
 pl(\bar{x}) &= m(\theta)pl(\bar{x}|\theta) + m(\bar{\theta})pl(\bar{x}|\bar{\theta}) \\
 &\quad + m(\Theta)(1 - (1 - pl(\bar{x}|\theta))(1 - pl(\bar{x}|\bar{\theta}))) \\
 pl(X) &= m(\theta)pl(X|\theta) + m(\bar{\theta})pl(X|\bar{\theta}) \\
 &\quad + m(\Theta)(1 - (1 - pl(X|\theta))(1 - pl(X|\bar{\theta})))
 \end{aligned} \tag{9}$$

The next example illustrates a case where Smets’ deduction operator produces inconsistent results. Let the conditional plausibilities be expressed as:

$$\Theta \mapsto X : \begin{cases} pl(x|\theta) = \frac{1}{4} & pl(\bar{x}|\theta) = \frac{3}{4} & pl(X|\theta) = 1 \\ pl(x|\bar{\theta}) = \frac{1}{4} & pl(\bar{x}|\bar{\theta}) = \frac{3}{4} & pl(X|\bar{\theta}) = 1 \end{cases} \tag{10}$$

Eq.(10) expresses that the plausibilities of x are totally independent of θ because $pl(x|\theta) = pl(x|\bar{\theta})$ and $pl(\bar{x}|\theta) = pl(\bar{x}|\bar{\theta})$. Let now two basic belief assignments (bbas), m_{Θ}^A and m_{Θ}^B on Θ be expressed as:

$$m_{\Theta}^A : \begin{cases} m_{\Theta}^A(\theta) = 1/2 \\ m_{\Theta}^A(\bar{\theta}) = 1/2 \\ m_{\Theta}^A(\Theta) = 0 \end{cases} \quad m_{\Theta}^B : \begin{cases} m_{\Theta}^B(\theta) = 0 \\ m_{\Theta}^B(\bar{\theta}) = 0 \\ m_{\Theta}^B(\Theta) = 1 \end{cases} \tag{11}$$

This results in the following plausibilities pl , belief masses m_X and pignistic probabilities E , on X in Table 1:

State	Result of m_{Θ}^A on Θ			Result of m_{Θ}^B on Θ		
	pl	m_X	E	pl	m_X	E
x	1/4	1/4	1/4	7/16	1/16	1/4
\bar{x}	3/4	3/4	3/4	1/16	9/16	3/4
X	1	0	n.a.	1	6/16	n.a.

Table 1: Inconsistent results of deductive reasoning with Smets’ method

Because X is totally independent of Θ according to Eq.(10), the bbas on X (m_X) should not be influenced by

the bbas on Θ . It can be seen from Table 1 that the probability expectation values E are equal for both bbas, which seems to indicate consistency. However, the belief masses are different, which shows that Smets' method [4] can produce inconsistent results. It can be mentioned that the framework of subjective logic does not have this problem.

In Smets' approach, binomial abduction is defined as:

$$\begin{aligned} pl(\theta) &= m(x)pl(x|\theta) + m(\bar{x})pl(\bar{x}|\theta) + m(X)(pl(X|\theta)) , \\ pl(\bar{\theta}) &= m(x)pl(x|\bar{\theta}) + m(\bar{x})pl(\bar{x}|\bar{\theta}) + m(X)pl(X|\bar{\theta}) , \\ pl(\Theta) &= m(x)(1 - (1 - pl(x|\theta))(1 - pl(x|\bar{\theta}))) \\ &\quad + m(\bar{x})(1 - (1 - pl(\bar{x}|\theta))(1 - pl(\bar{x}|\bar{\theta}))) \\ &\quad + m(X)(1 - (1 - pl(X|\theta))(1 - pl(X|\bar{\theta}))) . \end{aligned} \quad (12)$$

Eq.(12) fails to take the base rates on Θ into account and would therefore unavoidably be subject to the base rate fallacy, which would also be inconsistent with probabilistic reasoning as e.g. described in Example 1 (Sec.2.1). It can be mentioned that abduction with subjective logic is always consistent with probabilistic abduction.

3.2 Halpern's Approach to Conditional Plausibilities

Halpern (2001) [7] analyses conditional plausibilities from an algebraic point of view, and concludes that conditional probabilities, conditional plausibilities and conditional possibilities share the same algebraic properties. Halpern's analysis does not provide any mathematical methods for practical conditional deduction or abduction.

3.3 Conditional Reasoning with Imprecise Probabilities

Imprecise probabilities are generally interpreted as probability intervals that contain the assumed real probability values. Imprecision is then an increasing function of the interval size [8]. Various conditional reasoning frameworks based on notions of imprecise probabilities have been proposed.

Creedal networks introduced by Cozman [9] are based on credal sets, also called convex probability sets, with which upper and lower probabilities can be expressed. In this theory, a credal set is a set of probabilities with a defined upper and lower bound. There are various methods for deriving credal sets, e.g. [8]. Credal networks allow credal sets to be used as input in Bayesian networks. The analysis of credal networks is, in general, more complex than the analysis of traditional probabilistic Bayesian networks because it requires multiple analyses according to the possible probabilities in each credal set. Various algorithms can be used to make the analysis more efficient.

Weak non-monotonic probabilistic reasoning with conditional constraints proposed by Lukasiewicz [10] is also based on probabilistic conditionals expressed with upper and lower probability values. Various properties for conditional deduction are defined for weak non-monotonic probabilistic reasoning, and algorithms are described for deter-

mining whether conditional deduction properties are satisfied for a set of conditional constraints.

The surveyed literature on credal networks and non-monotonic probabilistic reasoning only describe methods for deductive reasoning, although abductive reasoning under these formalisms would theoretically be possible.

A philosophical concern with imprecise probabilities in general, and with conditional reasoning with imprecise probabilities in particular, is that there can be no real upper and lower bound to probabilities unless these bounds are set to the trivial interval $[0, 1]$. This is because probabilities about real world propositions can never be absolutely certain, thereby leaving the possibility that the actual observed probability is outside the specified interval. For example, Walley's Imprecise Dirichlet Model (IDM) [11] is based on varying the base rate over all possible outcomes in the frame of a Dirichlet distribution. The probability expectation value of an outcome resulting from assigning the total base rate (i.e. equal to one) to that outcome produces the upper probability, and the probability expectation value of an outcome resulting from assigning a zero base rate to that outcome produces the lower probability. The upper and lower probabilities are then interpreted as the upper and lower bounds for the relative frequency of the outcome. While this is an interesting interpretation of the Dirichlet distribution, it can not be taken literally. According to this model, the upper and lower probability values for an outcome x_i are defined as:

$$\text{IDM Upper Probability: } \bar{P}(x_i) = \frac{r(x_i) + W}{W + \sum_{i=1}^k r(x_i)} \quad (13)$$

$$\text{IDM Lower Probability: } \underline{P}(x_i) = \frac{r(x_i)}{W + \sum_{i=1}^k r(x_i)} \quad (14)$$

where $r(x_i)$ is the number of observations of outcome x_i , and W is the weight of the non-informative prior probability distribution.

It can easily be shown that these values can be misleading. For example, assume an urn containing nine red balls and one black ball, meaning that the relative frequencies of red and black balls are $p(\text{red}) = 0.9$ and $p(\text{black}) = 0.1$. The *a priori* weight is set to $W = 2$. Assume further that an observer picks one ball which turns out to be black. According to Eq.(14) the lower probability is then $\underline{P}(\text{black}) = \frac{1}{3}$. It would be incorrect to literally interpret this value as the lower bound for the relative frequency because it obviously is greater than the actual relative frequency of black balls. This example shows that there is no guarantee that the actual probability of an event is inside the interval defined by the upper and lower probabilities as described by the IDM. This result can be generalized to all models based on upper and lower probabilities, and the terms "upper" and "lower" must therefore be interpreted as rough terms for imprecision, and not as absolute bounds.

Opinions used in subjective logic do not define upper and lower probability bounds. As opinions are equivalent to general Dirichlet probability density functions, they always

cover any probability value except in the case of dogmatic opinions which specify discrete probability values.

3.4 Conditional Reasoning in Subjective Logic

Subjective logic [12] is a probabilistic logic that takes *opinions* as input. An opinion denoted by $\omega_x^A = (b, d, u, a)$ expresses the relying party A 's belief in the truth of statement x . Here b, d , and u represent belief, disbelief and uncertainty respectively, where $b, d, u \in [0, 1]$ and $b + d + u = 1$. The parameter $a \in [0, 1]$ is called the base rate, and is used for computing an opinion's probability expectation value that can be determined as $E(\omega_x^A) = b + au$. In the absence of any specific evidence about a given party, the base rate determines the *a priori* trust that would be put in any member of the community.

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.2 illustrates an example where the opinion about a proposition x from a binary state space has the value $\omega_x = (0.7, 0.1, 0.2, 0.5)$.

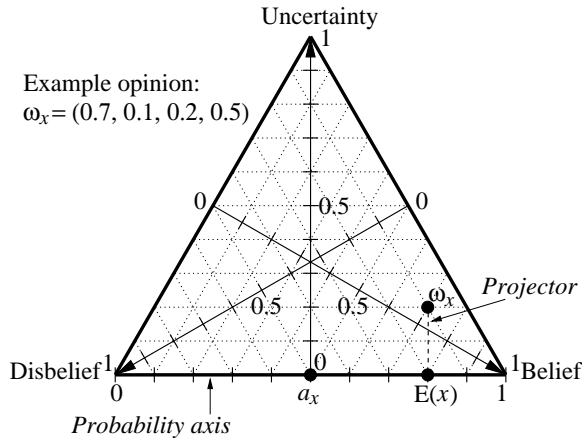


Figure 2: Opinion triangle with example opinion

The top vertex of the triangle represents uncertainty, the bottom left vertex represents disbelief, and the bottom right vertex represents belief. The parameter b_x takes value 0 on the left side edge and takes value 1 at the right side belief vertex. The parameter d_x takes value 0 on the right side edge and takes value 1 at the left side disbelief vertex. The parameter u_x takes value 0 on the base edge and takes value 1 at the top uncertainty vertex. The base of the triangle 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 $E(\omega_x^A)$.

The algebraic expressions for conditional deduction and abduction in subjective logic are relatively long and are therefore omitted here. However, they are mathematically simple and can be computed extremely efficiently. A full presentation of the expressions for binomial conditional deduction is given in [1] and a full description of multinomial deduction is provided in [13]. Only the notation is provided here.

Let $\omega_x, \omega_{y|x}$ and $\omega_{y|\bar{x}}$ be an agent's respective opinions about x being true, about y being true given that x is true, and about y being true given that x is false. Then the opinion $\omega_{y||x}$ is the conditionally derived opinion, expressing the belief in y being true as a function of the beliefs in x and the two sub-conditionals $y|x$ and $y|\bar{x}$. The conditional deduction operator is a ternary operator, and by using the function symbol ' \odot ' to designate this operator, we write:

$$\omega_{y||x} = \omega_x \odot (\omega_{y|x}, \omega_{y|\bar{x}}). \quad (15)$$

Subjective logic abduction is described in [1] and requires the conditionals to be inverted. Let x be the parent node, and let y be the child node. In this situation, the input conditional opinions are $\omega_{x|y}$ and $\omega_{x|\bar{y}}$. That means that the original conditionals are expressed in the opposite direction to what is needed.

The inverted conditional opinions can be derived from knowledge of the supplied conditionals, $\omega_{x|y}$ and $\omega_{x|\bar{y}}$, and knowledge of the base rate of the child, a_y .

Given knowledge of the base rate a_y of the child state where ω_y^{vac} is a vacuous subjective opinion about the base rate of the hypothesis, defined as

$$\omega_y^{\text{vac}} = (b_y, d_y, u_y, a_y) \begin{cases} b_y = 0 \\ d_y = 0 \\ u_y = 1 \\ a_y = \text{base rate of } y \end{cases} \quad (16)$$

and given the logical conditionals $\omega_{x|y}, \omega_{x|\bar{y}}$, then the inverted conditionals $\omega_{y|x}, \omega_{y|\bar{x}}$ can be derived using the following formula

$$\omega_{y|x} = \frac{\omega_y^{\text{vac}} \cdot \omega_{x|y}}{\omega_y^{\text{vac}} \odot (\omega_{x|y}, \omega_{x|\bar{y}})} \quad (17)$$

$$\omega_{y|\bar{x}} = \frac{\omega_y^{\text{vac}} \cdot \neg \omega_{x|y}}{\omega_y^{\text{vac}} \odot (\neg \omega_{x|y}, \neg \omega_{x|\bar{y}})}$$

The abduction operator, $\bar{\odot}$, is written as follows:

$$\omega_{y||x} = \omega_x \bar{\odot} (\omega_{x|y}, \omega_{x|\bar{y}}, a_y) \quad (18)$$

The advantage of subjective logic over probability calculus and binary logic is its ability to explicitly express and take advantage of ignorance and belief ownership. Subjective logic can be applied to all situations where probability calculus can be applied, and to many situations where probability calculus fails precisely because it can not capture degrees of ignorance. Subjective opinions can be interpreted

as probability density functions, making subjective logic a simple and efficient calculus for probability density functions. An online demonstration of subjective logic can be accessed at: <http://persons.unik.no/josang/sl/>.

4 Base Rates in Target Recognition Applications

Evidence for recognizing a target can come from different sources. Typically, some physical characteristics of an object are observed by sensors, and the observations are translated into evidence about a set of hypotheses regarding the nature of the object. The reasoning typically is based on conditionals, so that belief about the nature of objects can be derived from the observations. Let y_j denote the hypothesis that the object is of type y_j , and let x_i denote the observed characteristic x_i . In subjective logic notation the required conditional can be denoted as $\omega_{(y_j|x_i)}$. The opinion about the hypothesis y_j can then be derived as a function of the opinion ω_{x_j} and the conditional $\omega_{(y_j|x_i)}$ according to Eq.15.

It is possible to model the situation where several sensors observe the same characteristics, and to fuse opinions about an hypothesis derived from different characteristics. In order to correctly model such situations the appropriate fusion operators are required.

A common problem when assessing a hypotheses y_j is that it is challenging to determine the required conditionals $\omega_{(y_j|x_i)}$ on a sound basis. Basically there are two ways to determine the required conditionals, one complex but sound, and the other ad-hoc but unsafe. The sound method is based on first determining the conditionals about observed characteristics given specific object types, denoted by $\omega_{(x_i|y_j)}$. These conditionals are determined by letting the sensors observe the characteristics of known objects, similarly to the way the reliability of medical tests are determined by observing their outcome when applied to known medical conditions. In order to correctly apply the sensors for target recognition in an operative setting these conditionals must be inverted by using the base rates of the hypotheses according to Eq.(4) and Eq.(5) in case of probabilistic reasoning, and according to Eq.(17) in case of subjective logic reasoning. In the latter case the derived required conditionals are denoted $\omega_{(y_j|x_i)}$. This method is complex because the reliability of the sensors must be empirically determined, and because the base rate of the hypothesis must be known.

The ad-hoc but unsafe method consists of setting the required conditionals $\omega_{(y_j|x_i)}$ by guessing or by trial-and-error, in principle without any sound basis. Numerous examples in the literature apply belief reasoning for target recognition with ad hoc conditionals or thresholds for the observation measurements to determine whether a specific object has been detected, e.g. citeMACC2005-Fusion (p.4) and [14] (p.450). Such thresholds are in principle conditionals, i.e. in the form of "hypothesis y_j is assumed to be TRUE given that the sensor provides a measure x_i greater than a specific threshold". Examples in the literature typically focus on fusing beliefs about some hypotheses, not on how the

thresholds or conditionals are determined, so it is not necessarily a weakness in the presented models. However, in practice it is crucial that thresholds and conditionals are adequately determined. Simply setting the thresholds in an ad-hoc way is not adequate.

It would be possible to set conditionals $\omega_{(y_j|x_i)}$ in a practical situation by trial and error by assuming that the true hypothesis y_j were known during the learning period. However, should the base rate of the occurrence of the hypothesis change, the settings would no longer be correct. Similarly, it would not be possible to reuse the settings of conditionals in another situation whether the base rates of the hypothesis are different.

By using the sound method of determining the conditionals, their values can more easily be adjusted to new situations by simply determining the base rates of the hypothesis in each situation and deriving the required conditionals $\omega_{(y_j|x_i)}$ accordingly.

5 Base Rates in Global Warming Models

As an example we will test two hypotheses about the conditional relevance between CO₂ emission and global warming. Let x : "Global warming" and y : "Man made CO₂ emission". We will see which of these hypotheses lead to reasonable conclusions about the likelihood of man made CO₂ emission based on observing global warming. There have been approximately equally many periods of global warming as global cooling over the history of the earth, so the base rate of global warming is set to 0.5. Similarly, over the history of the earth, man made CO₂ emission has occurred very rarely, meaning that $a_y = 0.1$ for example. Let us further assume the evidence of global warming, i.e. that an increase in temperature can be observed, expressed as:

$$\omega_x = (0.9, 0.0, 0.1, 0.5) \quad (19)$$

• IPCC's View

According to the IPCC (International Panel on Climate Change) [15] the relevance between CO₂ emission and global warming is expressed as:

$$\omega_{x|y}^{\text{IPCC}} = (1.0, 0.0, 0.0, 0.5) \quad (20)$$

$$\omega_{x|\bar{y}}^{\text{IPCC}} = (0.8, 0.0, 0.2, 0.5) \quad (21)$$

$$(22)$$

According to the IPCC's view, and by applying the formalism of abductive belief reasoning of Eq.(18), it can be concluded that there is man made CO₂ emission with the likelihood $\omega_{y||x}^{\text{IPCC}} = (0.62, 0.00, 0.38, 0.10)$, as illustrated in Fig.3.

This is obviously a questionable conclusion since all but one period of global warming during the history of the earth has taken place without man made CO₂ emission.

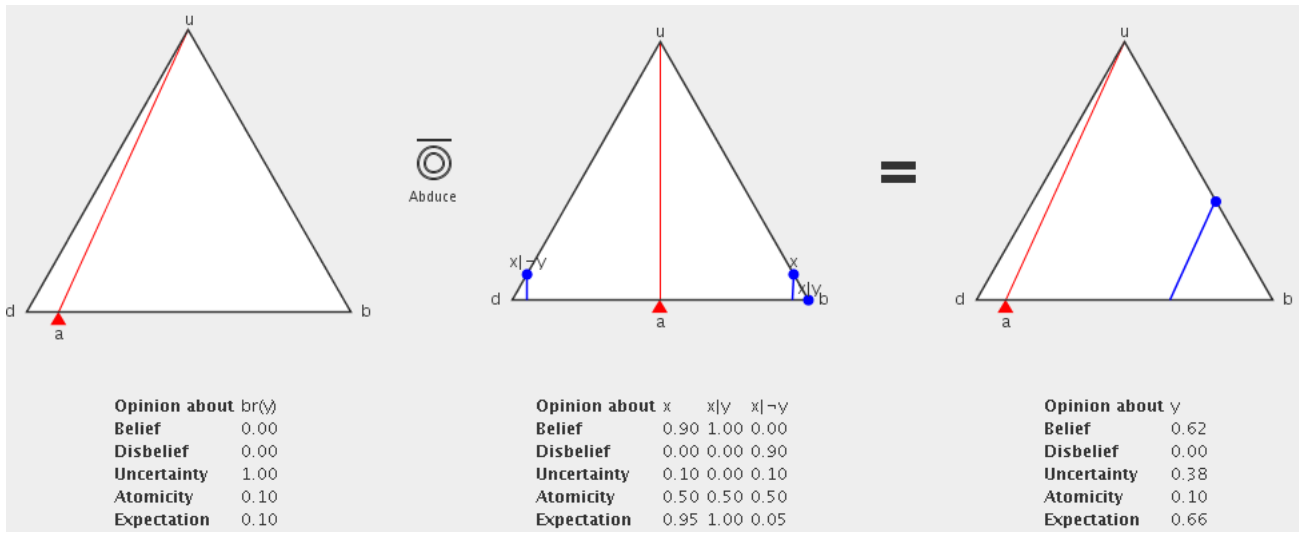


Figure 3: Likelihood of man made CO₂ emission with IPCC's conditionals

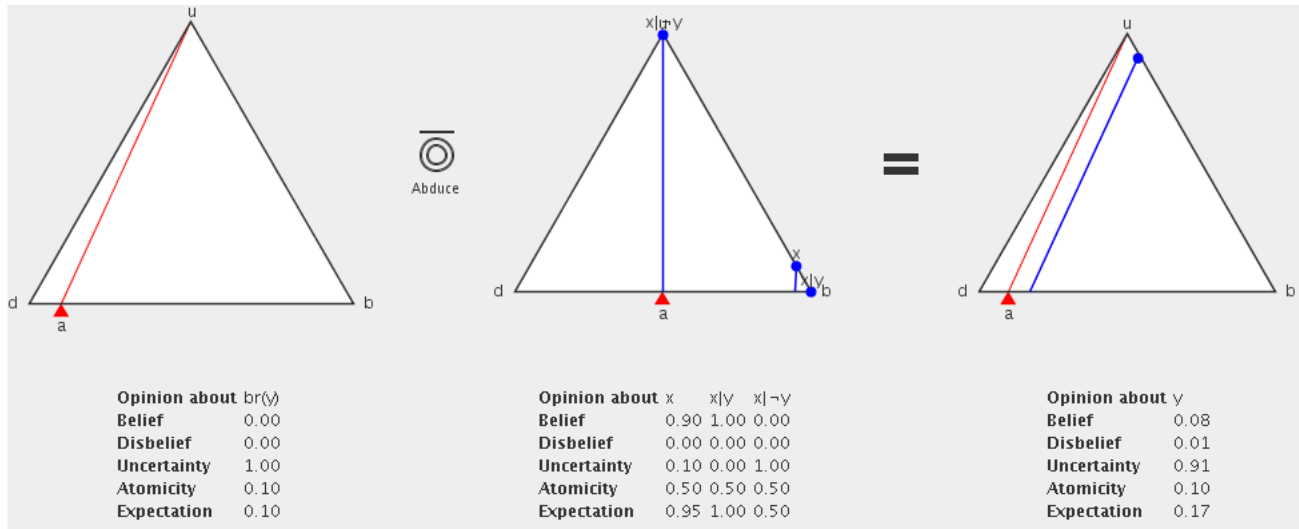


Figure 4: Likelihood of man made CO₂ emission with the skeptic's conditionals

• **The Skeptic's View**

Martin Duke is a journalist who produced the BBC documentary “*The Great Global Warming Swindle*” and who is highly skeptical about the IPCC. Let us skeptic Martin Dukin’s view that we don’t know anything about whether a reduction in man-made CO₂ emission would have had an effect on global warming expressed as:

$$\omega_{x|y}^{\text{Skeptic}} = (1.0, 0.0, 0.0, 0.5) \quad (23)$$

$$\omega_{x|\bar{y}}^{\text{Skeptic}} = (0.0, 0.0, 1.0, 0.5) \quad (24)$$

$$(25)$$

Based on the skeptic’s view, it can be concluded that there is man-made CO₂ emission with the likelihood $\omega_{y||x}^{\text{Skeptic}} = (0.08, 0.01, 0.91, 0.10)$, as illustrated in

Fig.4. This conclusion seems more reasonable in light of the history of the earth.

Based on this analysis the IPCC’s hypothesis seems unreasonable. As global warming is a perfectly normal phenomenon, it is statistically very unlikely that any observed global warming is caused by man-man CO₂ emission.

6 Discussion and Conclusion

Applications of belief reasoning typically assesses the effect of evidence on some hypothesis. We have shown that common frameworks of belief reasoning commonly fail to properly consider base rates and thereby are unable to handle derivative reasoning. The traditional probability projection of belief functions in form of the Pignistic transformation assumes a default subset base rate that is equal to the subset’s relative atomicity. In other words, the default base

rate of a subset is equal to the relative number of singletons in the subset with respect to the total number of singletons in the whole frame. Subsets also have default *relative base rates* with respect to every other fully or partly overlapping subset of the frame. Thus, when projecting a bba to scalar probability values, the Pignistic probability projection dictates that belief masses on subsets contribute to the projected probabilities as a function of the default base rates on those subsets.

However, in practical situations it would be possible and useful to apply base rates that are different from the default base rates. For example, when considering the base rate of a particular infectious disease in a specific population, the frame can be defined as {“infected”, “not infected”}. Assuming that an unknown person enters a medical clinic, the physician would *a priori* be ignorant about whether that person is infected or not before having assessed any evidence. This ignorance should intuitively be expressed as a vacuous belief function, i.e. with the total belief mass assigned to (“infected” \cup “not infected”). The probability projection of a vacuous belief function using the Pignistic probability projection would dictate that the *a priori* probability of having the disease is 0.5. Of course, the base rate of diseases is normally much lower, and can be determined by relevant statistics from a given population. Traditional belief functions are not well-suited for representing this situation. Using only traditional belief functions, the base rate of a disease would have to be expressed through a bba that assigns some belief mass to either “infected” or “not infected” or both. Then after assessing the results of e.g. a medical test, the bba would have to be conditionally updated to reflect the test evidence in order to derive the *a posteriori* bba. Unfortunately, no computational method for conditional updating of traditional bbas according to this principle exists. The methods that have been proposed, e.g. [5], have been shown to be flawed [13] because they are subject to the base rate fallacy [2].

Subjective logic is the only belief reasoning framework which is not susceptible to the base rate fallacy in conditional reasoning. Incorporating base rates for belief functions represents a necessary prerequisite for conditional belief reasoning, in particular for abductive belief reasoning. In order to be able to handle conditional reasoning within general Dempster-Shafer belief theory, it has been proposed to augment traditional bbas with a base rate function [16]. We consider this to be a good idea and propose as a research agenda the definition of conditional belief reasoning in general Dempster-Shafer belief theory.

References

- [1] 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.
- [2] Jonathan Koehler. The Base Rate Fallacy Reconsidered: Descriptive, Normative and Methodological Challenges. *Behavioral and Brain Sciences*, 19, 1996.
- [3] Frank J. Stech and Christopher Elässer. Midway revisited: Deception by analysis of competing hypothesis. Technical report, MITRE Corporation, 2004.
- [4] Ph. Smets. Belief functions: The disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, 9:1–35, 1993.
- [5] H. Xu and Ph Smets. Reasoning in Evidential Networks with Conditional Belief Functions. *International Journal of Approximate Reasoning*, 14(2–3):155–185, 1996.
- [6] H. Xu and Ph Smets. Evidential Reasoning with Conditional Belief Functions. In D. Heckerman et al., editors, *Proceedings of Uncertainty in Artificial Intelligence (UAI94)*, pages 598–606. Morgan Kaufmann, San Mateo, California, 1994.
- [7] J.Y. Halpern. Conditional Plausibility Measures and Bayesian Networks. *Journal of Artificial Intelligence Research*, 14:359–389, 2001.
- [8] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
- [9] F.G. Cozman. Credal networks. *Artif. Intell.*, 120(2):199–233, 2000.
- [10] T. Lukasiewicz. Weak nonmonotonic probabilistic logics. *Artif. Intell.*, 168(1):119–161, 2005.
- [11] P. Walley. Inferences from Multinomial Data: Learning about a Bag of Marbles. *Journal of the Royal Statistical Society*, 58(1):3–57, 1996.
- [12] A. Jøsang. A Logic for Uncertain Probabilities. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9(3):279–311, June 2001.
- [13] A. Jøsang. Conditional Reasoning with Subjective Logic. *Journal of Multiple-Valued Logic and Soft Computing*, 15(1):5–38, 2008.
- [14] Praveen Kumar, Ankush Mittal, and Padam Kumar. Study of Robust and Intelligent Surveillance in Visible and Multimodal Framework. *Informatica*, 31:447–461, 2007.
- [15] Intergovernmental Panel on Climate Change. *Climate Change 2007: The Physical Science Basis*. IPCC Secretariat, Geneva, Switzerland, 2 February 2007. url: <http://www.ipcc.ch/>.
- [16] Audun Jøsang, Stephen O’Hara, and Kristie O’Grady. Base Rates for Belief Functions. In *Proceedings of the Workshop on the Theory on Belief Functions (WTBF 2010)*, Brest, April 2010.