Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

L.C. van der Gaag

Different notions of uncertainty in quasi-probabilistic models

Computer Science/Department of Software Technology Report CS-R8847 December

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

Different Notions of Uncertainty in Quasi-Probabilistic Models

L.C. van der Gaag
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

ABSTRACT

In the early years of the research into plausible reasoning several quasi-probabilistic models for handling uncertainty in rule-based expert systems have been proposed. These models were computationally feasible but could not be justified mathematically. Although current research in this subarea of artificial intelligence concentrates on the development of mathematically sound models, the early quasi-probabilistic models are still employed frequently in present-day rule-based expert systems. In this paper we show that two of these models, the certainty factor model developed by E.H. Shortliffe and B.G. Buchanan and the subjective Bayesian method developed by R.O. Duda, P.E. Hart and N.J. Nilsson, model different notions of uncertainty. We furthermore discuss the difference in the interpretation and application of production rules in the respective models.

1980 Mathematics Subject Classification (1985): 60Bxx, 68Txx.
1987 CR Categories: I.2.3. **[Deduction and Theorem Proving]**: uncertainty, probabilistic reasoning; I.2.5 **[Programming Languages and Software]**: expert system tools and techniques.

Key Words & Phrases: rule-based expert systems, plausible reasoning, certainty factor model, subjective Bayesian method.

1. INTRODUCTION

Probability theory is one of the oldest mathematical theories concerning uncertainty. Therefore, it is no wonder that in the seventies this theory was chosen as the first point of departure for the development of models for handling uncertain information in knowledge-based systems. In these early years, research concentrated more in specific on the application of Bayesian probability theory in *rule-based expert systems*. It was soon discovered that this mathematical theory could not be applied in such a context in a straightforward manner. The researchers were, and still are, confronted with several problems, such as:

- It often is not possible to obtain a fully specified probability function on the field of concern: only a few probabilities are known or can be estimated by an expert in the field, i.e. only a *partially specified probability function* is available.
- In the case where an expert has assessed many of the probabilities, the thus partially specified probability function is likely to be *inconsistent* in the sense that there is not an underlying probability function.
- Bayesian probability theory does not provide *explicit computation rules* for computing probabilities from a partially specified probability function for all (intermediary) results derived during an actual consultation of the system.

From then, research for a short period of time centered around the development of modifications of Bayesian probability theory that should overcome these problems and that could be applied efficiently in a rule-based environment. The two most well-known models developed in this period are the *certainty factor model* especially designed for the expert system (E)MYCIN by E.H. Shortliffe and B.G. Buchanan [1], and the *subjective Bayesian method* developed by R.O. Duda, P.E. Hart and

Report CS-R8847
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

N.J. Nilsson for the PROSPECTOR system [2]. Neither of these models presents a theoretically well-founded solution to the problems encountered in handling uncertainty in a rule-based system: the models have an ad hoc character, thus justifying our phrase *quasi-probabilistic models*.

Current research on plausible reasoning concentrates on the development of mathematically sound models for dealing with uncertain information. Since this research has not yet yielded a model both computationally feasible and semantically clear, the early quasi-probabilistic models are still employed frequently in present-day expert systems. We therefore feel that it is worth comparing the models mentioned in the foregoing. Since both models were developed for handling uncertainty in a rule-based expert system, we informally introduce the notions of production rules and top-down inference independent of the two models, in Section 2. Only the propagation of uncertain evidence will be discussed in some detail. In the Sections 3 and 4 the subjective Bayesian method and the certainty factor model are discussed only with respect to the propagation of uncertain evidence; an investigation of the respective schemes for propagating uncertainty suffices to elucidate a conceptual difference between the two models. In Section 5 we argue that in the certainty factor model and the subjective Bayesian method different notions of uncertainty are employed. From this difference in their respective concepts of uncertainty we furthermore find that the models' respective interpretations of production rules differ as well.

2. Production Rules and the Propagation of Uncertainty

In a rule-based expert system the expert knowledge concerning the domain in which the system is to be employed, is modelled in a set of *production rules* of the form $e \to h$. The left hand side e of a production rule in general is a Boolean combination of atomic *conditions*. The right hand side h of a production rule in general is a conjunction of atomic *conclusions*. For the aforementioned investigation of the schemes for propagating uncertain evidence in quasi-probabilistic models however, it suffices to assume the left hand sides of the rules to be atomic. Notice that this assumption is not realistic in a more general case. In the sequel, the atomic left-hand side of a production rule is called a (*piece of*) *evidence*. We furthermore assume that the right hand side h of a rule is atomic as well; such an atomic conclusion will be called a *hypothesis*. A production rule $e \to h$ has the following meaning: if evidence e has been observed, then hypothesis h is true. Production rules can be applied backwards, i.e. from conclusion to condition, to prove one or more *goal hypotheses*: this goal-directed reasoning technique is called *top-down reasoning*. A production rule is said to *succeed* if the evidence mentioned in the rule has been found to be true; otherwise the rule is said to *fail*. Notice that during such an inference process *reasoning chains* arise from applying the production rules; these chains constitute a so-called *inference network*, [3].

In practice a hypothesis seldom is confirmed to absolute certainty by the observation of a certain piece of evidence. Therefore, the notion of a production rule is extended to include a measure of uncertainty: with the hypothesis h of the rule $e \to h$ a measure of uncertainty is associated indicating the degree to which h is confirmed by the observation of evidence e. So, a model for handling uncertain information provides an expert with a means for representing the uncertainties that go with the pieces of information he has specified. In the Sections 3 and 4 we will see that in both the certainty factor model and the subjective Bayesian method the measures of uncertainty to be associated with the hypotheses of the production rules are probability-based.

During a top-down inference process, several intermediary hypotheses are confirmed to some degree. These uncertain hypotheses can in turn be used as evidence for the confirmation of other hypotheses, i.e. as evidence in other production rules. This situation is shown below; hypothesis e has been confirmed to the degree x on account of prior evidence e' and is subsequently used as evidence in the production rule $e \to h$.

$$e' \xrightarrow{\;x\;} e \xrightarrow{\;y\;} h$$

We recall that the measure of uncertainty y associated with the hypothesis h in the rule $e \rightarrow h$ indicates the degree to which h is confirmed by the actual observation of e. The measure of uncertainty to be associated with h after applying the production rule in the situation shown above is dependent upon this measure of uncertainty y and the measure of uncertainty that has been associated with the evidence e used in confirming h. Each quasi-probabilistic model provides a function f for computing the actual measure of uncertainty to be associated with the hypothesis h after applying the rule in the situation shown above:

$$e' \xrightarrow{\;f(x,y)\;} h$$

e' now denotes all prior evidence. f will be called the *function for propagating uncertain evidence.*

Beside a propagation function each quasi-probabilistic model provides another three functions for combining measures of uncertainty, [1,2,3]. Here, we only deal with the respective functions for propagating uncertain evidence in the certainty factor model and the subjective Bayesian method; examination of the remaining combination functions however will show that these functions do not influence the results stated in this paper.

3. The Subjective Bayesian Method

The *subjective Bayesian method* was developed as a model for handling uncertainty by R.O. Duda, P.E. Hart and N.J. Nilsson in the seventies for PROSPECTOR, an expert system for assisting non-expert field geologists in exploring sites, [2,4]. Hereto PROSPECTOR, among other information, contains representations of ore deposite models in a collection of production rules.

We recall that a model for handling uncertainty provides a means of expressing the uncertainties that go with the hypotheses of the production rules. In the subjective Bayesian method, an expert has to express his uncertainty concerning the hypothesis h of a production rule $e \rightarrow h$ in two *conditional probabilities*: $P(h \mid e)$, the probability of h given the actual occurrence of evidence e, and $P(h \mid \bar{e})$, the probability of the hypothesis h given that e has definitely not occurred[1]. Furthermore it is assumed that the prior probabilities $P(h)$ as well as $P(e)$ are known to the system. So, for modelling uncertainty probability theory is used. Notice that in general the probability function P has not been specified for the entire problem domain; P has only been specified partially. For reasoning with uncertainty, the model provides solutions to the problems mentioned in the introduction, that arise in applying probability theory in a straightforward manner.

In Section 2 we have discussed that during a top-down inference process several intermediary hypotheses are confirmed or disconfirmed to some degree. Now suppose that for the intermediary hypothesis e the posterior probability $P(e \mid e')$ has been computed given some prior observations e', and that e subsequently is used as evidence for the confirmation of the hypothesis h in the production rule $e \rightarrow h$, as shown in the following inference network:

1. Actually, an expert has to provide the two likelihood ratios $\lambda = \frac{P(e \mid h)}{P(e \mid \bar{h})}$ and $\bar{\lambda} = \frac{P(\bar{e} \mid h)}{P(\bar{e} \mid \bar{h})}$. From these likelihood ratios however, the two conditional probabilities mentioned in the text can easily be computed.

$$e' \xrightarrow{P(e\,|\,e')} e \xrightarrow{P(h\,|\,e),P(h\,|\,\overline{e})} h$$

We now are interested in the probability $P(h\,|\,e')$ such that

$$e' \xrightarrow{P(h\,|\,e')} \mathrm{h}$$

Notice that in general, the probability $P(h\,|\,e')$ has not been given by the expert and cannot be computed from the probability function P since P has been specified only partially; $P(h\,|\,e')$ therefore has to be approximated from the probabilities that actually are known. In general, we have

$$\begin{aligned} P(h\,|\,e') &= P(h \cap e\,|\,e') + P(h \cap \overline{e}\,|\,e') = \\ &= P(h\,|\,e \cap e')P(e\,|\,e') + P(h\,|\,\overline{e} \cap e')P(\overline{e}\,|\,e'). \end{aligned}$$

We assume that if we know e to be absolutely true (or false), then the observations e' relevant to e do not provide any further information on h, i.e. we assume the Markov property. Under this assumption we obtain

$$\begin{aligned} P(h\,|\,e') &= P(h\,|\,e)P(e\,|\,e') + P(h\,|\,\overline{e})P(\overline{e}\,|\,e') = \\ &= (P(h\,|\,e) - P(h\,|\,\overline{e})) \cdot P(e\,|\,e') + P(h\,|\,\overline{e}). \end{aligned}$$

We have that $P(h\,|\,e')$ is a linear interpolation function in $P(e\,|\,e')$. In Figure 3.1 an interpolation function for the situation of the production rule $e \rightarrow h$ shown above, is depicted.

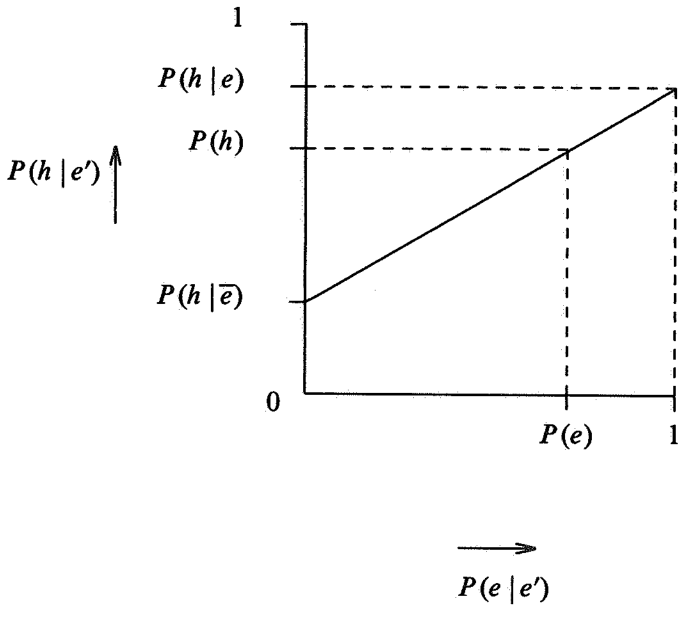


FIGURE 3.1. $P(h\,|\,e')$ as a linear interpolation function in $P(e\,|\,e')$.

For $P(e\,|\,e') = 0$ we have $P(h\,|\,e') = P(h\,|\,\overline{e})$; for $P(e\,|\,e') = 1$ we have $P(h\,|\,e') = P(h\,|\,e)$. For any $P(e\,|\,e')$ between 0 and 1 the corresponding value for $P(h\,|\,e')$ can be read from the figure. For instance, if evidence e' has been observed confirming e, i.e. if $P(e \mid e') > P(e)$, we find that the probability of h increases from applying the production rule $e \rightarrow h$: $P(h \mid e') > P(h)$. Notice that this effect is exactly what is meant by the rule. In the special case where $P(e \mid e') = P(e)$, we have $P(h \mid e') = P(h \mid e)P(e) + P(h \mid \bar{e})P(\bar{e}) = P(h)$.

In principle, this interpolation function offers an explicit computation rule for propagating uncertain evidence. Duda, Hart and Nilsson however have observed that when an expert is asked to assess for each rule $e \rightarrow h$ the four probabilities $P(h)$, $P(e)$, $P(h \mid e)$ and $P(h \mid \bar{e})$, the specified values are likely to be *inconsistent*, in the sense that there is not an underlying actual probability function. More in specific, the relation between $P(h)$ and $P(e)$ as shown in Figure 3.1 will be violated. We show to which problems such an inconsistency may lead.

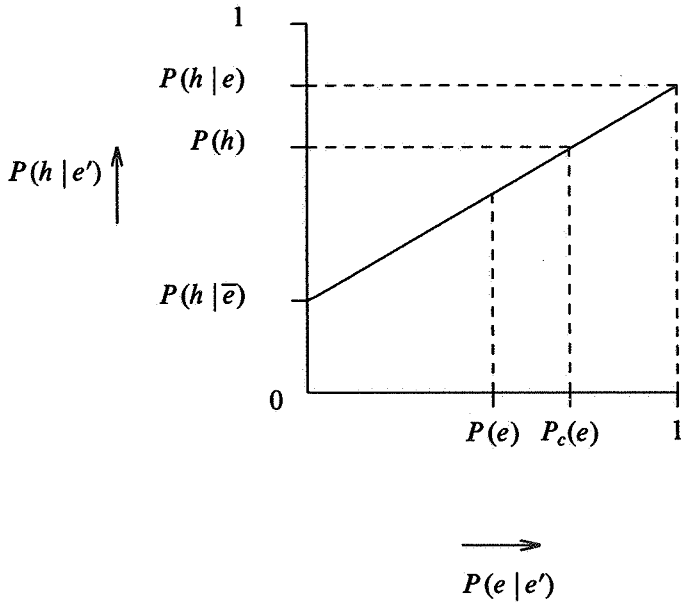


FIGURE 3.2. Inconsistently specified prior probabilities.

Consider Figure 3.2: the expert has assessed the probabilities $P(h)$, $P(e)$, $P(h \mid e)$ and $P(h \mid \bar{e})$. The specified values are inconsistent; the consistent value for $P(e \mid e')$ corresponding with $P(h)$ is shown as $P_c(e)$. Now suppose that evidence e' has been observed confirming e to a degree $P(e \mid e')$ such that $P(e) < P(e \mid e') < P_c(e)$. From Figure 3.2 we have that $P(h \mid e') < P(h)$. The production rule $e \rightarrow h$ however was meant to express that confirmation of e leads to confirmation of h: due to the inconsistency, the reverse has been achieved. A natural solution to this problem would be to reassess $P(e)$ by choosing $P(e) = P_c(e)$ (or, in case the assessment of $P(h)$ is less certain than the assessment of $P(e)$, to reassess $P(h)$ by choosing a consistent value for $P(h)$). The hypotheses h and e however may occur in several places in a given set of production rules and each reassessment affects all these occurrences. Reassessing prior probabilities therefore is not a feasible solution to the problem we have discussed.

Duda, Hart and Nilsson have developed several methods for employing inconsistently specified probabilities, one of which has been implemented as the function for propagating uncertain evidence in PROSPECTOR. The main idea of the method that has been chosen for implementation is shown in Figure 3.3; the original interpolation function is splitted in two separate interpolation functions on the intervals $[0,P(e)]$ and $(P(e),1]$ respectively, such that the property $P(h \mid e') = P(h)$ if $P(e \mid e') = P(e)$ is enforced. Notice that the closer the function value for $P(e)$ is to the value for $P(e)$ from the original interpolation function, the better the initial assessments of $P(e)$ and $P(h)$ are.

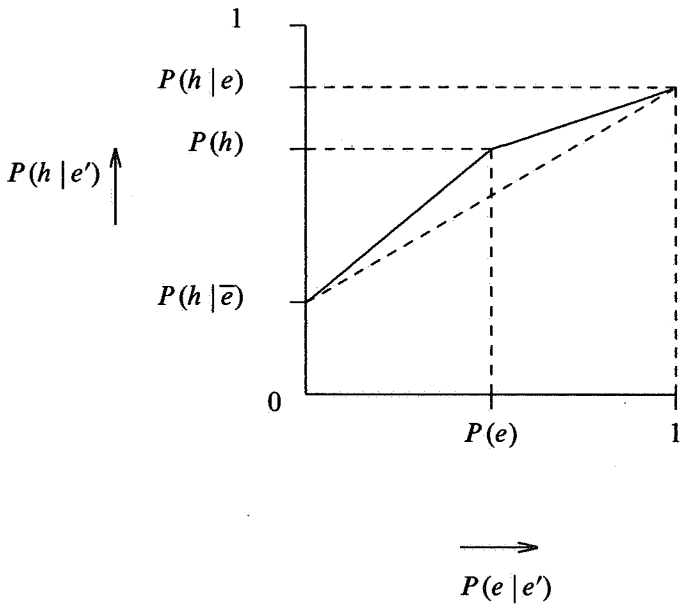


FIGURE 3.3. The 'consistent' propagation of uncertain evidence.

The function

$$P(h \mid e') = \begin{cases} P(h \mid \overline{e}) + \dfrac{P(h) - P(h \mid \overline{e})}{P(e)} \cdot P(e \mid e') & \textit{if } 0 \leqslant P(e \mid e') \leqslant P(e) \\ P(h) + \dfrac{P(h \mid e) - P(h)}{1 - P(e)} \cdot (P(e \mid e') - P(e)) & \textit{if } P(e) < P(e \mid e') \leqslant 1 \end{cases}$$

is the resulting interpolation function.

4. THE CERTAINTY FACTOR MODEL

In the seventies another quasi-probabilistic model for handling uncertainty in rule-based systems was proposed: the *certainty factor model*, [1]. This model has been developed by E.H. Shortliffe and B.G. Buchanan for the purpose of introducing the notion of uncertainty in MYCIN, the well-known rule-based expert system for diagnosing infectious diseases.

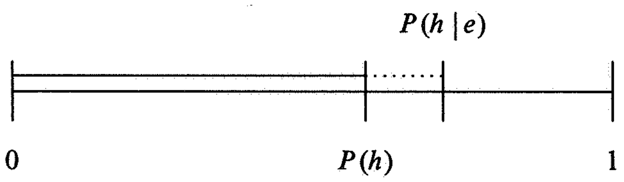
We recall that a model for handling uncertainty provides a means for expressing the uncertainties to be associated with the hypotheses in the production rules as well as a means for combining several uncertainties into new ones. In the certainty factor model, an expert has to express his uncertainty concerning the hypothesis h in a production rule $e \rightarrow h$, in a *certainty factor* $CF(h,e)$. Such a certainty factor is a real number between -1 and $+1$. A positive certainty factor is associated with the hypothesis h given the evidence e if h is confirmed in some degree by the observation of e; the certainty factor $CF(h,e) = 1$ indicates that the occurrence of e completely proves the hypothesis h. A negative certainty factor is suggested by the expert if the observation of evidence e disconfirms the hypothesis h. A certainty factor equal to zero is proposed if the observation of e does not influence the confidence in the hypothesis h. The initially given certainty factors are considered to be function values of a two-argument function CF; notice that in general, this function CF has been specified only partially.

In present-day implementations of the certainty factor model, the experts and the users are confronted only with the certainty factor function and the associated functions for combining certainty factors, of which the function for propagating uncertain evidence will be discussed shortly. In [1] however, Shortliffe and Buchanan have introduced the certainty factor function as a function composed of two different notions of uncertainty, both probability-based: the measure of belief and the measure of disbelief. These new notions of uncertainty were devised to capture the intuitive concepts of confirmation and disconfirmation, and have been inspired to a large extent by confirmation theory, [1,5].

The *measure of (increased) belief MB* is a function having two arguments, a hypothesis and a piece of evidence, and yielding function values in the interval [0,1]. The function is defined by

$$MB(h,e) = \begin{cases} 1 & \textit{if } P(h) = 1 \\ max \left\{ 0, \dfrac{P(h \mid e) - P(h)}{1 - P(h)} \right\} & \textit{else} \end{cases}$$

The measure of belief can be accounted for intuitively as follows: we depict the prior probability of the hypothesis h, i.e. $P(h)$, on a scale from 0 to 1:



The maximum amount of belief that can still be added to the prior belief in h, equals $1 - P(h)$. If a piece of evidence e is observed confirming h, i.e. such that $P(h \mid e) > P(h)$, then this observation results in adding the amount of belief $P(h \mid e) - P(h)$ to the prior belief in h. The belief in h therefore has been increased to the degree $\dfrac{P(h \mid e) - P(h)}{1 - P(h)}$.

The other basic uncertainty function used in the certainty factor model is the *measure of (increased) disbelief MD*, equally having two arguments and yielding function values in the interval [0,1]:

$$MD(h,e) = \begin{cases} 1 & \textit{if } P(h) = 0 \\ max \left\{ 0, \dfrac{P(h) - P(h \mid e)}{P(h)} \right\} & \textit{else} \end{cases}$$

This function can be accounted for similarly.

According to Shortliffe and Buchanan the need for new notions of uncertainty arose from their observation that an expert often was unwilling to accept the logical implications of his probabilistic statements, such as: if $P(h \mid e) = x$ then $P(\bar{h} \mid e) = 1 - x$. They state that in the mentioned case an expert would claim that evidence e in favor of hypothesis h should not be construed as evidence against the hypothesis as well. The reason that the logical implication concerning $P(\bar{h} \mid e)$ may seem counterintuitive is explained by J. Pearl as follows, [6]. The phrase 'evidence e in favor of hypothesis h' is interpreted as stating an *increase* in the probability of the hypothesis from $P(h)$ to $P(h \mid e)$, with $P(h \mid e) > P(h)$: $P(h \mid e)$ is viewed in relation with $P(h)$. On the other hand, in the argument of Shortliffe and Buchanan $P(\bar{h} \mid e)$ seems to be taken as an absolute probability irrespective of the prior $P(\bar{h})$. This somehow conveys the false idea that $P(\bar{h})$ increases by some positive factor. However if for example $P(\bar{h}) = 0.9$ and $P(\bar{h} \mid e) = 0.5$, then no expert will construe this considerable decrease in the probability of $\bar{h}$ as supporting the negation of h! Anyhow, the measures of belief and disbelief explicitly capture the notion that one piece of evidence cannot both favor and disfavor a single hypothesis: it can easily be shown that $MB(h,e) = 0$ if $MD(h,e) > 0$ and vice versa.

As we will see shortly, the certainty factor function CF is defined in terms of the measures of belief and disbelief only. So, we may assume for the moment that with the hypothesis h in a production rule $e \rightarrow h$ the two function values $MB(h,e)$ and $MD(h,e)$ are associated. Again it should be noted that in general these functions have only been specified partially, i.e. only for some arguments h and e. We have seen before that during a top-down inference process uncertain hypotheses that have been confirmed or disconfirmed to some degree, are used as evidence in confirming other hypotheses. Now suppose that for the intermediary hypothesis e the actual measures of belief and disbelief $MB(e,e')$ and $MD(e,e')$ have been computed given some prior observations e', and that e subsequently is used as evidence in the production rule $e \rightarrow h$, as shown in the following inference network:

$$e' \xrightarrow{MB(e,e'),\ MD(e,e')} e \xrightarrow{MB(h,e),\ MD(h,e)} h$$

We are interested in obtaining measures of belief and disbelief $MB(h,e')$ and $MD(h,e')$ such that

$$e' \xrightarrow{MB(h,e'),\ MD(h,e')} \mathrm{h}$$

In the certainty factor model the following functions for combining measures of belief and disbelief are employed for the propagation of uncertain evidence:

$$MB(h,e') = MB(h,e) \cdot MB(e,e'), \text{ and}$$
$$MD(h,e') = MD(h,e) \cdot MB(e,e').$$

Notice the asymmetry in these formulas. The two functions for propagating uncertain evidence cannot be justified from the probabilistic definitions of MB and MD, [7]. Unfortunately, Shortliffe and Buchanan were not able to formulate correct functions for computing measures of belief and disbelief from the initially given function values only.

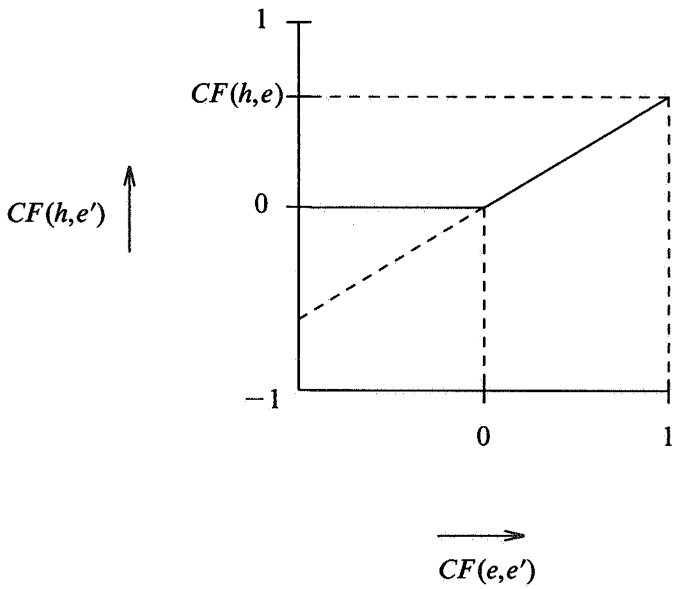


FIGURE 4.1. The propagation of uncertain evidence.

The certainty factor function as already has been mentioned before, is defined in terms of the measures of belief and disbelief:

$$CF(h,e) = MB(h,e) - MD(h,e)^1.$$

Using this definition and the functions for propagating the measures of belief and disbelief of a piece of uncertain evidence, Shortliffe and Buchanan have derived a function for propagating uncertain evidence in terms of certainty factors only; this function for propagating uncertain evidence is the following one:

$$CF(h,e') = CF(h,e) \cdot max\{0, CF(e,e')\}.$$

The function is depicted in Figure 4.1 for a given $CF(h,e)$.

5. Different Notions of Uncertainty

In the preceding two sections we have discussed several components of the certainty factor model and the subjective Bayesian method as models for handling uncertainty in a rule-based expert system. We have seen that both models provide a means for expressing uncertainty. In the subjective Bayesian method uncertainty is modelled using (subjective) probability theory; the developers of the certainty factor model have introduced another notion of uncertainty. In this section we compare the behaviour of the two models on a given set of production rules: we take the behaviour of the models to be equivalent if top-down inference with a given set of production rules employing one model yields the same conclusions concerning the goal hypotheses in the same order when sorted according to their associated uncertainties, as top-down inference with the same set of rules employing the other model. In this section we show that the behaviour of the certainty factor model is different from that of the subjective Bayesian method although the notions of uncertainty used in both models are probability-based. The difference between the two models becomes apparent when closely examining the propagation of uncertain evidence.

Consider the production rule $e \rightarrow h$. In the subjective Bayesian method the two conditional probabilities $P(h \mid e)$ and $P(h \mid \bar{e})$ are associated with the hypothesis h of the rule; in the certainty factor model the certainty factor $CF(h,e)$ is associated with h. Now suppose that the intermediary hypothesis e is confirmed to some degree by prior observations e'.

- In the subjective Bayesian method confirmation of e corresponds with the situation $P(e \mid e') > P(e)$. From Figure 3.3 we read that applying the production rule $e \rightarrow h$ results in an increase of the probability of the hypothesis h: $P(h \mid e') > P(h)$. The increase of the probability of h is linear in the increase of the probability of the evidence e. The maximum value for $P(h \mid e')$ is attained when e is absolutely certain, and equals the probability $P(h \mid e)$ that has been associated with the rule.
- In the certainty factor model confirmation of evidence e corresponds with $CF(e,e') > 0$. From Figure 4.1 we read that applying the rule $e \rightarrow h$ equally results in confirmation of the hypothesis h to a certain degree: $CF(h,e') > 0$. The degree to which h is confirmed is linear in the degree to which the evidence e is confirmed. The maximum value for $CF(h,e')$ is attained when e is certain, and equals the certainty factor $CF(h,e)$ that has been associated with the rule.

It is obvious that in both models no change in the belief in the evidence e results in no change in the

1. Actually, the certainty factor function is defined as $CF(h,e) = \dfrac{MB(h,e) - MD(h,e)}{1 - min\{MB(h,e), MD(h,e)\}}$.

The scaling factor $1 - min\{MB(h,e), MD(h,e)\}$ has been added to the notion of the certainty factor function for pragmatical reasons. In our case, this factor always equals 1. When taking the other functions for combining measures of belief and disbelief in consideration as well, this scaling factor may differ from 1 due to the errors introduced in combining these uncertainty measures.

belief in the hypothesis h. So far, the models behave equivalently.

Now consider the case where prior observations e' disconfirm e.

- In the subjective Bayesian method we have $P(e \mid e') < P(e)$. From Figure 3.3 we read that applying the production rule $e \rightarrow h$ in this case results in a decrease of the probability of the hypothesis h: $P(h \mid e') < P(h)$. The minimum is attained when $\bar{e}$ is absolutely certain, and equals $P(h \mid \bar{e})$, the other conditional probability associated with the rule.
- In the certainty factor model disconfirmation of e corresponds with $CF(e,e') < 0$. From Figure 4.1 we read that applying the rule $e \rightarrow h$ does not result in any change in the belief in the hypothesis h: h is not disconfirmed by the disconfirmation of the evidence e.

From these observations we have that the two models do not interpret the production rule $e \rightarrow h$ the same way. In the subjective Bayesian method the rule is interpreted as expressing that 'confirmation of e results in confirmation of h *and* disconfirmation of e results in disconfirmation of h', whereas in case of the certainty factor model the rule is interpreted as expressing that 'confirmation of e results in confirmation of h' only. So, when the certainty factor model is employed a production rule can contribute either positively or negatively to the belief in a hypothesis, but not both ways; Shortliffe and Buchanan have succeeded in modelling a notion of uncertainty that is conceptually different from the notion of uncertainty used in probability theory and in the subjective Bayesian method, more in specific. Notice that the interpretation of the rule $e \rightarrow h$ in the sense of the subjective Bayesian method is similar to the interpretation of the set of rules $\{e \rightarrow h, \bar{e} \rightarrow \bar{h}\}$ viewed according to the certainty factor model[1].

From this difference in their respective interpretation of production rules, it will be evident that given a specific set of production rules the models in general will behave differently: when employing one model the system may rank the confirmed hypotheses in an order differing from the one obtained from employing the other model, and what is still worse, it may even come up with entirely different answers. So, the two models are not interchangeable.

6. Conclusion

In this paper we have discussed two of the early quasi-probabilistic models designed for handling uncertainty in knowledge-based systems: the subjective Bayesian method and the certainty factor model. Although these models cannot be justified mathematically, they are still employed frequently in present-day rule-based expert systems. We have closely examined the means the respective models provide for expressing the uncertainty to be associated with the hypothesis of a production rule, and we have compared their respective schemes for propagating uncertain evidence. From these observations it became apparent that the two models employ different notions of uncertainty, resulting in a conceptual difference in the interpretation of production rules and eventually in a different behaviour given a specific set of production rules. We have argued that for a given set of production rules in which certainty factors are employed, we cannot simply substitute conditional probabilities for certainty factors: the system may come up with different answers. Before choosing a model for handling uncertainty in a rule-based expert system therefore, the conceptual notion of uncertainty an expert has in mind should be elicited; this task should not be taken too lightly.

1. The case is somewhat more complicated than is stated here due to the way negations are handled by the certainty factor model.

REFERENCES

[1] E.H. SHORTLIFFE, B.G. BUCHANAN (1984). A model of inexact reasoning in medicine, in: B.G. BUCHANAN, E.H. SHORTLIFFE (eds.) *Rule-Based Expert Systems. The MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley, Reading, Massachusetts.

[2] R.O. DUDA, P.E. HART, N.J. NILSSON (1976). *Subjective Bayesian Methods for Rule-Based Inference Systems*, Technical Note 124, Artificial Intelligence Center, SRI International, Menlo Park.

[3] L.C. VAN DER GAAG (1987). *A Network Approach to the Certainty Factor Model*, Report CS-R8757, CWI, Amsterdam.

[4] R. REBOH (1981). *Knowledge Engineering Techniques and Tools in the PROSPECTOR Environment*, Technical Note 243, SRI International, Menlo Park.

[5] R. CARNAP (1950). *Logical Foundations of Probability*, University of Chicago Press, Chicago.

[6] J. PEARL (1985). *How To Do With Probabilities What People Say You Can't*, Report CSD-850031, University of California, Los Angeles.

[7] L.C. VAN DER GAAG (1988). *The Certainty Factor Model and its Basis in Probability Theory*, Report CS-R8816, CWI, Amsterdam.