UvA-DARE (Digital Academic Repository)

An analysis of five strategies for reasoning in uncertainties and their suitability for pathology

van Ginneken, A.M.; Smeulders, A.W.M.

Publication date
1988

Published in
Machine intelligence and pattern recognition

[Link to publication](#)

Citation for published version (APA):

van Ginneken, A. M., & Smeulders, A. W. M. (1988). An analysis of five strategies for reasoning in uncertainties and their suitability for pathology. *Machine intelligence and pattern recognition*, 367-380.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

A078355648
NCC/IBL AANVRAAGBON

KOPIE MONOGRAFIE EGB

(13)
18-04-2005

Datum indienen : 17-04-2005 14:30 5493-1 Clearing House
Datum plaatsen : 17-04-2005 14:30
Aanvrager : 0004/9998
Aanvraagident :
Aanvragerident : 0004/9999
Eindgebruiker : 041631433

Telefoonnummer : 050-3635057
Coöperatiecode : R

Leverwijze : Elektronisch
Fax :
Ftp :
E-Mail : m.s.van.delden@rug.nl
Ariel :

Plaatscode : 233356657 ; IN 516.35 p022 ; ;

(1) ☐ Origineel gestuurd	(6) ☐ Niet beschikbaar
(2) ☐ Copie gestuurd	(7) ☐ Uitgeleend
(3) ☐ Overige	(8) ☐ Wordt niet uitgeleend
(4) ☐ Nog niet aanwezig	(9) ☐ Bibliografisch onjuist
(5) ☐ Niet aanwezig	(0) ☐ Bij de binder

A078355648
NCC/IBL AANVRAAGBON

Verzamelnota volgt.
KOPIE MONOGRAFIE EGB

18-04-2005

Datum indienen : 17-04-2005 14:30 9560-1 UvA Keur
Datum plaatsen : 17-04-2005 14:30 UB Groningen
Aanvrager : 0004/9998 Broerstraat 4
Aanvraagident : 9700 AN Groningen
Aanvragerident : 0004/9999
Eindgebruiker : 041631433 tav

Aantal

PPN Titel : 233356657
Titel : Pattern recognition and artificial intelligence : towards an
: integration : proceedings of an international workshop held
Auteur : Gelsema, Edzard S. (ed.), 1988
Deel/Supplem. :
Corporatie : Externe database:
Jaar/Editie : 1988 Extern nummer :
Uitgave : Amsterdam North-Holland
Serie/Sectie : Machine-intelligence and pattern recognition. Machine-intellige
Pag-ISSN/ISBN : recognition ; vol.0-444-87137-3

Plaatscode : 233356657 ; IN 516.35 p022 ; ;

Jaar :
Volume :
Aflevering :
Eindgebruiker : 041631433 Aanvraagident. :
Auteur : Aanvragerident. : UVA KEUR (UB GRONINGEN)
Artikel :
Bladzijden : 367-380
Bron :
Opmerking : arno ID: 136702

AN ANALYSIS OF FIVE STRATEGIES FOR REASONING IN
UNCERTAINTIES AND THEIR SUITABILITY FOR PATHOLOGY

Astrid M. van Ginneken, Arnold W.M. Smeulders

Department of Medical Informatics
Erasmus University
Rotterdam

In reasoning systems, uncertainty plays a crucial part, especially for those fields where judgements are essential, as in pathology. Uncertainty has several aspects, such as prevalence of diseases, occurrence of findings and the predictive value of findings.

For the functioning of a reasoning system two aspects are crucial: first, the internal representation of the uncertainty and second, the way in which the uncertainty is propagated in the reasoning process when combining formal statements.

Five well known reasoning strategies are compared: probability theory, MYCIN's certainty factor model, fuzzy set theory, the theory of Dempster-Shafer and the scoring scheme of Internist. The comparison addresses, among others, the following questions:

- Can the different aspects of uncertainty be dealt with as separate entities?
- How are unknown uncertainties dealt with?
- How is evidence in favor of a hypothesis combined with evidence against it?
- How does the model treat the simultaneous occurrence of more than one disorder, that is, how does the model support reasoning with compound hypotheses?

It is preliminarily concluded, that only in Internist and probability theory, the different aspects of uncertainty are expressed as separate entities. Hence, the other models do not accurately represent uncertain knowledge. Also, theoretically attractive models such as Bayes, MYCIN and the theory of Dempster-Shafer can only function properly under the tight condition of mutual exclusiveness of hypotheses, not always suited for broader parts of pathology. They may, however, be suited for smaller parts with a limited number of defined diseases and a limited number of features. All models but Bayes lack a predictable performance as there is no or only a partial underlying theory to guarantee minimization of the overall error.

1. INTRODUCTION

In pathology as in many other medical fields, the increase of the body of medical knowledge makes it increasingly difficult for a general pathologist to master the whole field of the discipline at the present level of available knowledge. The general scope of our project is to make available in an efficient way reference knowledge, now less accessible in books and journals in the form of a computerized system.

Such a reference system consists of two major parts: a knowledge base and a reasoning system (inference engine). The knowledge base holds a collection of interpreted facts, i.e. facts with a meaning in the context of a specific subfield. The reasoning system provides directives as to which facts of knowledge to combine and how to combine them to arrive at an answer.

In this process uncertainties play a crucial part in several ways. First, there is uncertainty with respect to the presence or absence of a finding either alone or in combination or, in general, the reliability thereof. For example, nucleoli may be considered large by one expert and moderately enlarged by another expert. Second, there is uncertainty as how to interpret facts, i.e. uncertainty in the knowledge. Certainty about the meaning of thyroid tissue strongly depends on its localization. If normal appearing thyroid follicles are found in the thyroid, pathologists will usually believe, that this section is obtained from a normal thyroid, and certainty of the diagnosis benignancy is close to 1. However, if the same morphological appearance is found in the lung or in other places certainty about the diagnosis "normal ectopic thyroid tissue" now is less than 1. Third, there is the problem of consensus about the knowledge: the same criterion may be assigned different meanings by different experts.

Though the reliability of histopathological diagnoses is well accepted in general, the certainty of a concluded diagnosis is not always equal to one, due to uncertainty with respect to one or more of the aspects mentioned. Often, the surgical pathologist will be aware of the uncertainty in a particular case and will seek additional information to confirm or exclude a certain diagnosis. Each new observation may add to the certainty of a particular conclusion, provided the new observations are not too strongly correlated with the previous findings.

A reasoning system can only deal with observational uncertainty if such uncertainty is explicitly provided with the data on which it operates. The problem of consensus about the knowledge is outside the scope of a reasoning system. Here, the term uncertainty refers to uncertainty in the knowledge, i.e. the interpretation of findings.

This paper will focus on techniques dealing with uncertainty and the combination thereof in the reasoning part of an expert system. When knowledge is combined, in what way are uncertainties combined (the combinatorics) to reach the conclusion and in what way do they affect the conclusion?

Strategies for the combinatorics of uncertainties have been studied in a number of papers. We will discuss the following five most well-known combination techniques to gain insight in their suitability for pathology:

- Bayes theorem (probability reasoning)
- Certainty factor model of MYCIN (rule-based reasoning)
- Theory of Dempster - Shafer (reasoning with hypothesis sets)
- Fuzzy logic (reasoning with hierarchical hypothesis groups)
- Internist (heuristic, hypothesis directed reasoning)

Prior to the analysis of these five models, a short description will be given of the definition and combinatorics of these models. In the analysis, the following comparisons are addressed:

- Which aspects of uncertainty are expressed?
- How are unknown uncertainties dealt with?
- How is evidence in favor of a hypothesis combined with evidence against that hypothesis?
- Is it possible to reason with uncertainties assigned to hypothesis sets? I.e. is it possible to zoom in from a large hypothesis set to a smaller one as more information becomes available?
- What part do uncertainties play in the selection of questions?
- What are the conditions for application?

2. DEFINITIONS AND COMBINATORICS

Bayes

In probability theory three parameters play an important part:

$P(D|e)$ = probability of diagnosis D given the finding e , also predictive value

$P(e|D)$ = probability of the finding e given diagnosis D , also sensitivity

$P(D)$ = a priori probability of diagnosis D

$P(e)$ can be calculated and is defined by $P(e|D) * P(D)$.

The rule of Bayes [1] for combination of evidence is given by:

$$P(D_i|e) = \frac{P(e|D_i) * P(D_i)}{\sum_j P(e|D_j) * P(D_j)}$$

CF - model of MYCIN

MYCIN [2] is a rule-based system and Certainty Factors are assigned to the rules themselves and their premises. A CF does not reflect the predictive value of a finding with respect to a diagnosis, but acts as a belief update in the computation process of the final certainty of that diagnosis. A CF is, however, partly based on the predictive value. Its definition [3] is given by:

$$CF(D|e) = \begin{cases} \frac{P(D|e) - P(D)}{P(D|e) * (1 - P(D))} & \text{if } P(D|e) \geq P(D) \\ \frac{P(D|e) - P(D)}{P(D) * (1 - P(D|e))} & \text{if } P(D|e) < P(D) \end{cases}$$

When two rules have an identical "then - part", their CF's have to be combined according to the rule for parallel combination:

$$CF(D|e1 \wedge e2) = \begin{cases} CF(D|e1) + CF(D|e2) - CF(D|e1) * CF(D|e2) & \text{if both CF's} > 0 \\ \frac{CF(D|e1) + CF(D|e2)}{1 - \min(|CF(D|e1)|, |CF(D|e2)|)} & \text{if one of both CF's} < 0 \\ CF(D|e1) + CF(D|e2) + CF(D|e1) * CF(D|e2) & \text{if both CF's} < 0 \end{cases}$$

When the "then - part" of a rule serves as the "if - part" of another rule then the formula for sequential combination should be applied, where e_1 acts as evidence for e_2 and e_2 in its turn for D :

$$CF(D|e_1) = \begin{cases} CF(e_2|e_1) * CF(D|e_2) & \text{if both CF's } \geq 0 \\ -CF(e_2|e_1) * CF(D|\neg e_2) & \text{if } CF(e_2, e_1) < 0 \end{cases}$$

Theory of Dempster-Shafer

In the theory of Dempster-Shafer [4], a basic probability assignment, BPA, assigns belief to a subset of all possible diagnoses, based on evidence. A BPA lies in the interval $[0,1]$ and all BPA values together sum to 1. The interpretation of a BPA comes closest to the predictive value of a finding with respect to a diagnosis (set). BPA values are assigned by a function m , which fulfills the following conditions:

$$\begin{aligned} m_e(\{\emptyset\}) &= 0 \\ m_e(D_i) &= 0 && \text{when } e \text{ does not support } D_i \\ m_e(D_j) &\in [0,1] && \text{when } e \text{ supports } D_j \\ m_e(D_{tot}) &= 1 - \sum_j m(D_j) && \text{for all } D_j, \text{ supported by } e \end{aligned}$$

Two m functions based on evidence e_1 and e_2 respectively, are combined as follows:

$$\begin{aligned} m_{e_1 e_2}(\{\emptyset\}) &= 0 \\ m_{e_1 e_2}(D) &= \frac{\sum (m_{e_1}(D_i) * m_{e_2}(D_j))}{\sum (m_{e_1}(D_k) * m_{e_2}(D_l))} \end{aligned} \quad \begin{aligned} &\text{for } i, j \text{ where } D_i \cap D_j = D \\ &\text{for } k, l \text{ where } D_k \cap D_l \neq \{\emptyset\} \end{aligned}$$

When all evidence has been gathered and combined, the total Belief for a subset D_i of D_{tot} , based on the combined evidence E is given by:

$$B_E(D_i) = \sum m_E(D_j) \quad \text{for all } j \text{ such, that } D_j \subseteq D_i$$

Fuzzy set theory

Fuzzy set theory [5] can be used to classify a patient to a single diagnosis or a set of diagnoses, based on clinico-pathological findings. A degree of membership value (DOM) expresses the belief in the assignment of a patient to a diagnosis (set). A DOM value lies in the interval $[0,1]$, where 0 represents total belief in non-membership and 1 total belief in membership. The values do not have to sum to 1. Intuitively, DOM values express a concept close to predictive value. The combinatoric for DOM values with respect to diagnosis D and evidence $e_1 \dots e_k$ is simple and given by:

$$DOM(D|e_1 \dots e_k) = \min/\max [DOM(D|e_1), \dots, DOM(D|e_k)]$$

Internist

During the reasoning process of Internist [6,7], diagnoses receive a numerical score, based on clinico-pathological findings. Two parameters are assigned to express uncertainties with respect to each relevant diagnosis-finding pair:


```
evoking strength: likelihood of a diagnosis given the finding,  
                  expressed on a scale of integers of 1 to 5  
frequency        : likelihood of a finding given the diagnosis,  
                  expressed on a scale of integers of 0 to 5
```

In addition, each finding is assigned an importance value on a scale of integers ranging from 1 to 5. This value is high, when the clinical consequences of a finding are regarded high.

A diagnosis receives a score based on the available findings. A finding, explained by the diagnosis will add positively to the score of that diagnosis with bonuspoints calculated on the basis of the evoking strength value. A finding, which should not be present in a particular diagnosis reduces the score of that diagnosis with a number of points, computed on the basis of the frequency value. Findings, not explained by a diagnosis decrease the score to a degree, which is calculated from the importance value. It is important to mention, that the addition to the diagnosis score based on evoking strengths is relatively larger, than the reduction of the score due to frequency values. In other words: evidence supporting a diagnosis is favored over evidence against that diagnosis.

Depending on the numerical distances between the scores of several possible diagnoses, the system will ask additional information to confirm the diagnosis with the highest score or to achieve more spreading in the scores when they lie close together [8].

3. ANALYSIS

3.1 What aspects of uncertainty are expressed?

We will consider two important aspects of uncertainty:

- The predictive value of a finding: the probability of a diagnosis given the finding
- The sensitivity of a finding: the probability of the finding given the diagnosis

In probability theory, $P(D|e)$ expresses the predictive value, whereas $P(e|D)$ expresses the sensitivity of a finding. Evidence in favor of a hypothesis is expressed by the fact that $P(e|D) > P(e)$, where $P(e)$ reflects the a priori probability of the finding e .

In Bayes' rule, the ratio $P(e|D)/P(e)$ represents the degree to which $P(D)$ is updated, when evidence e becomes available. This ratio is larger than 1, when evidence in favor of D is involved and smaller than 1 when the evidence disfavors D . Data to apply Bayes' rule can be obtained directly from statistical archives, computed from statistical data or may be estimated.

In MYCIN, CF values act as belief updates, where the interpretation of belief comes closest to predictive value of one or a combination of findings. A positive CF value in the interval $[0,1]$ corresponds to a value of $P(e|D)$ in the interval $[P(e),1]$. A negative CF value in the interval $[-1,0]$ corresponds to $P(e|D) \subseteq [0,P(e)]$. CF values can be computed from statistical data, but this results in loss of information. Usually, values are given by the expert.

A BPA value in Dempster-Shafer's theory closely resembles the expression of the predictive value of a finding with respect to one diagnosis or a set of diagnoses. A fundamental difference between $P(D|e)$ in Bayes' theory and a BPA value is the fact that a BPA value cannot only be assigned to singleton hypotheses, but to hypothesis sets as well. Since all BPA values with respect to the same evidence sum to one, an m function resembles a probability density function.

However, unassigned belief is treated differently as is described in the next paragraph. As in the CF - model BPA values are estimated, probably based on available probabilistic data.

DOM values express the likeness of one or more diagnoses given the finding and as such, the aspect of uncertainty they express corresponds best to the predictive value of the finding. Since DOM values do not have to sum to 1, they cannot be interpreted as real probabilities. Intuitively, they have the strongest similarity with BPA values from the theory of Dempster-Shafer and $P(D|e)$ from Bayes' theory with D being a member of a set of hypotheses, which are not necessarily mutually exclusive. DOM values have to be estimated.

Internist has a heuristic reasoning system with two different parameters for expressing uncertainties. Intuitively frequencies correspond to the sensitivity of a finding and evoking strengths to the predictive value, be it that they are expressed on a very small scale of 1 and 0 to 5 respectively.

Only in the models of Bayes and Internist is the sensitivity of findings expressed explicitly. Though the explicit expression of these two aspects of uncertainty requires more values, these values are easier to approximate. When more than one aspect of uncertainty is incorporated in one parameter, it is difficult to determine how much each aspect should add to the final value.

3.2 How are unknown uncertainties dealt with?

The five strategies will now be compared with respect to the way they deal with unknown uncertainties. When it is unknown how a certain piece of evidence should update the certainty of a particular diagnosis, the presence if this evidence should have no influence on the existing certainty of that diagnosis.

When using Bayes' rule, hypotheses with unknown a priori probabilities receive equal a priori values, such that all these values sum to 1. The remaining four models do not explicitly express a priori probabilities or equivalents thereof. If a piece of evidence becomes available for which $P(e|D_i)$ is not known it receives the value $P(e)$. As then $P(D_i|e) = P(D_i)$, evidence e does not affect the probability of D_i , i.e. it has no influence. The same holds for MYCIN, where $CF(D_i, e)$ receives the value 0 if it is unknown how evidence e updates hypothesis D_i . According to the combining functions for CF's, evidence e will not update the belief in D_i . Similarly, in the theory of Dempster-Shafer, the BPA value of D_i will be 0 when the significance of a piece of evidence with respect to D_i is unknown. Belief which cannot explicitly be assigned to one or more hypotheses is assigned to the complete set D_{tot} . Consequently, unknown uncertainties do not influence the value of $BEL(D_i)$. In fuzzy set theory and Internist, unknown uncertainties receive no default values as in the previously discussed models. As DOM values do not have to sum to a particular value, it is not necessary to assign default values. Therefore, existing DOM values are not influenced. In Internist, unexplained manifestations add negatively to the score of that hypothesis, depending on their importance. Though importance cannot be regarded as an expression of uncertainty, it influences the effect of evoking strengths and frequencies by decreasing the hypothesis score.

It can be concluded, that unknown uncertainties, whether or not assigned default values, do not update existing belief in a hypothesis except for Internist. In Internist, unexplained findings have a negative effect on the hypothesis score, i.e. reduce the chance for a hypothesis to become a concluded diagnosis.

3.3 How is evidence in favor of a hypothesis combined with evidence against that hypothesis?

With respect to this topic two aspects are important:

- Does the presence of pieces of evidence with an opposite effect on the certainty of a diagnosis remain visible throughout the combination process?
- Is the net effect of the combination of pieces of evidence with an opposite effect on the certainty of a diagnosis proportionate to their weight?

The first aspect is important, since it makes a difference whether a diagnosis is concluded on positive evidence alone or the combination of both positive and negative evidence. The second aspect reflects the desire that all pieces of evidence contribute to the certainty of a diagnosis proportionally. When e_1 is evidence in favor of hypothesis D and e_2 pleads against it (with e_1 and e_2 independent), then Bayes rule gives:

$$P(D|e_1 \wedge e_2) = \frac{P(e_1|D) * P(e_2|D)}{P(e_1) * P(e_2)} * P(D)$$

This quotient determines whether $P(D|e_1 \wedge e_2)$ will increase or decrease. The new probability of D does not reveal separately the positive and negative influence of evidence, leading to the change of $P(D)$. When the quotient exceeds the value 1, the combined evidence of e_1 and e_2 supports D , though e_1 and e_2 do not both support D . An equal amount of evidence in favor of a hypothesis D and against it does not differ from the absence of evidence since in both cases the existing belief in D remains unchanged.

In the CF-model evidence in favor of a hypothesis is represented by a positive CF value as opposed to a negative CF value for evidence against a hypothesis. When CF's with opposite signs are combined, the the CF with the largest absolute value will not only determine the sign of the new combined CF, but is also proportionally slightly favored over the other CF value since the denominator of the combination formula is less than 1. Here again, information with respect to the presence of a combination of positive and negative evidence is lost. As in Bayes, the presence of an equal amount of positive and negative evidence is indistinguishable from the absence of knowledge as the CF value in both cases is 0.

In the strategy of Dempster-Shafer a hypothesis can only receive BPA values based upon evidence in favor of that hypothesis. If evidence against a hypothesis becomes available, only its negation will receive a BPA, but not D . In Dempster-Shafer's strategy positive and negative evidence remain recognizable in the belief values and are very well distinguished from the absence of evidence. In addition it is obvious that each piece of evidence adds to D or the negation of D proportionate to its weight.

In fuzzy logic DOM values assigned to evidence in favor of a hypothesis range from 0 to 0.5, whereas DOM values corresponding to positive evidence range from 0.5 to 1. The combination rule for DOM values prescribes that the combined DOM value is equal to the lowest or highest value of the composite DOM values. Assume, that the lowest DOM value is decisive. The correctness of this combination rule is questionable when positive evidence is absent for a hypothesis as it implicitly acts against that hypothesis. However, this rule has another consequence: a low DOM value, resulting from negative evidence overrules all evidence present with higher DOM

values. Evidence in favor of and against a hypothesis is not at all treated proportionally, nor does it remain recognizable in the resulting DOM value whether the evidence is positive, negative or a combination of both.

In Internist, positive manifestations add positively to the hypothesis score. Negative manifestations, i.e. evidence against the hypothesis add negatively. A decrease of the hypothesis score may also stem from unexplained manifestations, depending on their importance. Thus, the reasoning strategy of Internist not only attempts to find the most likely diagnosis, but takes into account the consequence of a conclusion as well. However, positive manifestations are favored, since the increase in score due to positive manifestations is larger than the decrease resulting from similar negative manifestations. The resulting hypothesis score does not reveal whether positive and negative evidence is present in combination.

It can be concluded, that the models of Bayes, MYCIN and Dempster-Shafer treat positive and negative evidence proportionally, but only the latter separates the influence of positive and negative evidence. The fuzzy set model and Internist do not treat evidence proportionally and do not separate evidence in favor of and evidence against a diagnosis. In Internist hypothesis selection is not only based on evidence, but influenced by clinical consequences as well.

3.4 Is it possible to reason with uncertainties assigned to hypothesis sets?

In pathology the pathologist usually focusses on a decreasing number of diagnoses as more patient data become available. In an initial stage, when little is known about a patient, it is attractive when a reasoning system can establish which disease or combination of diseases is most likely. Each time when new evidence is gathered the hypothesis set should become smaller.

Using Bayes' rule or the CF model, only one disease is assumed to be present. When little information about the patient is available, many probabilities and CF's receive default values respectively. Then, the reliability of a conclusion is poor and as more data become available the concluded disease may change many times before the system concentrates on one disease hypothesis. A pathologist will not easily understand such a jumpy behavior, which also obscures the line of questioning.

When using strategies like Dempster-Shafer, fuzzy logic and to some extent Internist, it is possible to zoom in from the conclusion of a large hypothesis set based upon little data to a smaller and more specific set of hypotheses as more data becomes available. Hypothesis sets, serving as intermediate conclusions, contain diseases which are related on the basis of one or more common properties. These hypothesis sets do not necessarily correspond to disease groups as found in pathological classification trees. As a consequence, additional patient data are asked on a general groups level first and a specific, detailed level later. A better understandable line of questioning is the result.

3.5 What part do uncertainties play in the selection of questions?

In discussing the topic of question selection our present emphasis will be on how uncertainties may play a part in the strategy of question selection. The knowledge base, upon which the reasoning

system operates, contains the relations between findings and diagnoses. These relations enable the system to select questions, appropriate for a particular diagnosis or group of diagnoses. To assess the usefulness of a particular question, the predictive value of the possible answers, the likeliness of them to occur and their importance have to be considered. The more specific a finding is for a disease, the more important it is to acquire that evidence. It is obvious when two pieces of evidence have the same predictive value, that the one most likely to occur is the most attractive one to ask. The crucial issue is: which question is most likely to give maximum information gain?

In probability theory and in Internist a finding is likely to occur when both the values for $P(e)$ and frequency are high. $P(e)$ can be computed from sensitivities and a priori probabilities or approximated by an expert. Sensitivity and predictive value are not necessarily related. A finding occurring in almost every patient with a certain diagnosis and not in patients without that diagnosis, has both a high sensitivity and predictive value for that diagnosis. If, on the other hand, the same finding is also often seen in patients without that diagnosis, then the finding still has a high sensitivity, but a low predictive value. In the CF - model, fuzzy sets and the theory of Dempster-Shafer, sensitivity is not expressed in their parameters for uncertainty. Question selection strategies, using these models depend on other sources for the information on frequency of findings. A well-known strategy for selecting the question from which maximum information gain may be expected is the use of the entropy formula. However, this formula requires Bayesian probabilities to be available.

It is important for a system to decide when it stops to ask questions; when no appropriate questions remain or when the remaining available questions have a too small predictive value. user may expect from the reasoning system to ask questions, which enable this system to make progress in its diagnosis selection. A question which will not bring about any change in the systems conclusion should not be asked. The CF model of MYCIN stops asking questions, when the remaining appropriate rules have CF's of 0.2 or less. Nothing is mentioned about negative CF values. Internist will not ask questions with importance values of 2 or less. When the task of the reasoning system is to purely generate a list of possible diagnoses by descending certainty, importance values should only serve as a threshold when to stop question selection. For the remaining models, threshold values for stopping question selection can be defined.

Predictive value and to a lesser extent the occurrence of findings are important for question selection. All five models express some form of predictive value in their uncertainties, but only the models of Bayes and Internist explicitly express sensitivities. When not present in a model, the subtlety of question selection can be improved by defining the thresholds for question selection and parameters to express the occurrence of findings.

3.6 What are the conditions for application?

Prior to mentioning which condition has to be met by which model, the conditions will be discussed.

The condition of mutually exclusive hypotheses implies, that only one diagnosis is assumed to be present. In reality, several diagnoses may be present in a patient. Therefore, the condition is difficult to realize since it requires that all possible combinations of diagnoses be defined as separate diagnoses. Though the condition might limit the applicability of models with this

condition in many fields of medicine, this is less true for pathology. Multiple diseases play a less important part in pathology: abnormalities in patient tissue are generally spatially separated, each abnormal area posing its own diagnostic problem. As a consequence, application of a model, which has to meet the condition of mutually exclusive hypotheses, may require a limited number of "newly defined" mixture diagnoses, or leads to a minor violation of the condition.

The condition of independence of evidence implies, that the effect of a new piece of evidence on the certainty of any diagnosis is not influenced by the presence of previously acquired evidence. This condition is formulated in some models for the purpose of simplification and modularity. It is however difficult to satisfy. Firstly, there is not enough information or knowledge available to determine for each possible piece of evidence its dependency relation with any other piece of evidence. Secondly, even if it were possible to determine which pieces of evidence are independent of each other, it is still questionable whether enough evidence remains to differentiate between all separately defined diagnoses. It may seem convenient to have a model, which does not require independence of evidence. If, however, interdependence of evidence is completely ignored, some diagnoses may be disproportionately favored, due to the fact that each piece of evidence adds to the certainty as much as it would if it were independent of other evidence.

Bayes' rule, the CF-model of MYCIN and the theory of Dempster-Shafer all require that the hypotheses are mutually exclusive. Only the latter two require independence of evidence as well.

4. CONCLUSION

To gain insight in the suitability for pathology of the five strategies mentioned above, we now concentrate on those topics, which are important for the choice or design of a reasoning system and which differ considerably in the five strategies. The topics are: the condition of mutual exclusive hypotheses and independence of evidence, the explicit expression of aspects of uncertainty, the feasibility of data acquisition, reasoning with hypothesis sets and the suitability of the combinatorics.

The condition of mutual exclusive hypotheses, which has to be met by the models of Bayes, MYCIN and Dempster-Shafer, seems not to be a major disadvantage when a well defined part of the pathology spectrum is considered. The condition of independence of evidence cannot be easily met, but the omission of the condition is not desirable either. It is attractive when a model allows for the expression of interdependence of evidence. The latter is possible in Bayes, but of little practical use as long as the data are not available. It is interesting to study how the models of Bayes, MYCIN and Dempster-Shafer respond quantitatively to changes in the dependence of evidence.

The degree to which knowledge about uncertainties can be made explicit, depends on which aspects of uncertainty are expressed in uncertainty parameters. All five models allow the expression of predictive value, but only Bayes and Internist can explicitly express sensitivity. Explicit knowledge about the predictive value of findings is important for an efficient reasoning and question selection strategy. Sensitivities or a priori occurrences of findings have to be expressed explicitly for the question selection strategy, but this is not necessarily so for the reasoning strategy (MYCIN, fuzzy sets and Dempster-Shafer). For the purpose of question

selection it is a disadvantage for a system when its uncertainties do not explicitly express sensitivity or occurrence.

For the acquisition of uncertainties, statistical data are an important source of information, which can be used directly for the application of Bayes' rule or the computation of CF values, as well as indirectly for approximation of uncertainties in the remaining three models. Detailed statistical information is scarce in pathology especially the occurrence of findings, which implies that uncertainties have to be approximated by experts or be given default values. Approximations of uncertainties are difficult even by experts when rare diseases or combinations of diseases are concerned. It is questionable, whether an expert can reliably approximate the prevalence of diseases in the order of two or three per thousand. This difference may seem small, but in the reasoning system the difference is relative and thus 50 percent.

When experts approximate uncertainties, they have to make a proper mapping from uncertainty in reality to uncertainties as used in a reasoning system. Therefore, the expert must be able to form an idea of these uncertainties. The parameters of these five systems are of a different nature: BPA's and DOM values resemble probabilities, CF's represent belief updates and the parameters of Internist correspond to probabilities, but are expressed on a small scale. Uncertainties are likely to be easier to approximate, when their expressions have more similarities with true probabilities. When no statistical data are available, it may be more difficult to approximate a CF value, than a DOM value. In addition, the more aspects of uncertainty are incorporated in one parameter, the more difficult it is to approximate the value for that parameter. Thus it is to be expected, that the parameters of Bayes and Internist are easier to approximate than the parameters of the other three models, as they express different aspects of uncertainty in different parameters.

The strategies of Dempster-Shafer and fuzzy sets allow for reasoning with hypothesis sets as opposed to Bayes and MYCIN. An initial set of possible diseases may be reduced as more patient data become available. This is an attractive characteristic as it corresponds to the general diagnostic procedure. In the Pathfinder system [6], which uses the scoring scheme of Internist, group discriminate mode allows reasoning along a tree of predefined diagnostic groups, but not with any subset of the complete set of diagnoses.

As to the applicability of the combinatorics in pathology, it seems a drawback of fuzzy set theory, when - for example - the lowest DOM value is decisive. Not only negative specific evidence, but also nonspecific evidence overrules positive specific evidence. This implies, that diseases tend to receive less belief than would be expected. Though less prominent than in fuzzy set theory, the models of Internist and MYCIN do not combine pros and cons proportionate to their weight. In Internist, findings with an unknown relation to a diagnosis are not indifferent to the certainty of that diagnosis. Though the models of Bayes, Dempster-Shafer and MYCIN lack these disadvantages, only application of Bayes' rule guarantees a conclusion with minimal overall error.

It can be concluded, that an important disadvantage of the CF - model and the theory of Dempster-Shafer lies in the restrictive condition of independence of evidence. These two models together with Bayes' rule require diagnoses to be mutually exclusive, a condition which might only be slightly violated when a small part of the pathology spectrum is considered. Only Bayes and Internist explicitly express more than one aspect of uncertainty in their parameters, which serves not only in data acquisition but in question selection as well. The strategies of Dempster-Shafer and fuzzy set theory allow reasoning with hypothesis sets, which is in

conceptual accordance with the general diagnostic procedure. As to the handling of unknown data and pros and cons, the combinatorics of Bayes and Dempster-Shafer are slightly preferable over the model of MYCIN and Internist, but largely over the fuzzy combinatorics. Provided the required information is known, Bayes has a completely predictable performance and guarantees a conclusion with minimal overall error. Table I gives a schematic overview of the topics as discussed in the analysis.

TABLE I

	B A Y E S	M Y C I N	D S	F U Z Z Y	I N T E R N I S T
Expression uncertainties					
sensitivity (or analogous concept)	+	-	-	-	+
predictive value (or analogous concept)	+	+	+	+	+
Unknown uncertainties					
indifference	+	+	+	+	-
Pro's and con's					
visibility	-	-	+	-	-
proportionate weighing	+	-	+	-	-
Sets	-	-	+	+	-
Conditions					
not necessarily mutually exclusive	-	-	-	+	+
evidence not necessarily indendent	+	-	-	+	+

ACKNOWLEDGEMENTS

We thank Johan van der Lei for his help in the acquisition of literature and useful discussions. Joop Kuik is acknowledged for increasing our insight in probability theory. Funding for the project was provided by the "Preaventiefonds".

REFERENCES

- [1] Shortliffe, E.H., Buchanan, B.G., A Model of Inexact Reasoning in Medicine, in: Buchanan, B.G., Shortliffe, E.H., "Rule-Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project", pp. 233-262, Addison Wesley, 1984.
- [2] Buchanan, B.G., Shortliffe, E.H., Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project, Addison-Wesley, 1984.
- [3] Heckerman, D.E., Probabilistic Interpretations For Mycin's Certainty Factors, in: Kanal, L.N., Lemmer, J.F. (editors), "Uncertainty in Artificial Intelligence", pp. 167-195, North Holland, 1986.
- [4] Gordon, J., Shortliffe, E.H., The Dempster-Shafer theory of evidence, in: Buchanan, B.G., Shortliffe, E.H., "Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project", pp. 272-292, Addison Wesley, 1984.
- [5] Zvieli, A., Chen, P.P., Entity-Relationship modeling and Fuzzy databases, Second Conference on data engineering, Los Angeles, Feb. 3-6, 1986.
- [6] Horvitz, E.J., Heckerman, D.E., Nathwani, B.N., Fagan, L.M., Diagnostic Strategies in the Hypothesis-Directed PATHFINDER System, in "Proceedings of: The First Conference on Artificial Intelligence Applications", IEEE Computer Society, pp. 630-636, 1984.
- [7] Miller, R.A., Pople, H.E., Myers, J.D., INTERNIST-1, An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine, New England Journal of Medicine, Vol. 307, No. 8, pp. 468-476, 1982.
- [8] Horvitz, E.J., Heckerman, D.E., Nathwani, B.N., Fagan, L.M., The Use of a Heuristic Problem-Solving Hierarchy to Facilitate the Explanation of Hypothesis-Directed Reasoning, MEDINFO 86, IFIP-IMIA, 1986.