Hypothesis Management Framework: a flexible design pattern for belief networks in decision support systems

Sicco Pier van Gosliga, Imelda van de Voorde
TNO Defence, Security and Safety
The Hague, The Netherlands

Abstract

This article discusses a design pattern for building belief networks for application domains in which causal models are hard to construct. In this approach we pursue a modular belief network structure that is easily extended by the users themselves, while remaining reliable for decision support. The Hypothesis Management Framework proposed here is a pragmatic attempt to enable analysts and domain experts to construct and maintain a belief network that can be used to support decision making, without requiring advanced knowledge engineering skills.

1 INTRODUCTION

Since their introduction by Kim and Pearl [10] belief networks have become a popular framework for decision support and automated reasoning. Also at TNO, the Netherlands Organisation for Applied Scientific Research, Bayesian reasoning is used in an increasing number of projects and application domains. One of these application domains is decision support for criminal investigations. The typical application in this field is to perform a quick scan on available evidence to select the most likely hypothesis, and to prioritize unavailable evidence to aid further investigations. The need for sound probabilistic reasoning is quite large in this area, and belief networks are becoming an accepted tool for modeling reasoning.

Well-known examples of belief networks such as the Alarm [2] and Hailfinder [1] networks are quite complex and their development requires the co-operation between both Bayesian specialists and domain experts. Also, currently available software packages (e.g. HUGIN, Netica and GeNie)[1] for modeling and analysing belief networks require expertise and skill in belief networks. Whereas in the field of criminal investigations, the typical user of such decision support software is usually not a Bayesian specialist but either an analyst or an expert on the area being analyzed, a so-called domain expert. To get belief networks accepted as a standard tool in criminal investigations, we should improve the usability to such a degree that a domain expert is able to produce useful models without the assistance of a Bayesian specialist. Obviously, analysts should find it beneficial for performing their analyses as well.

Besides offering criminal investigators a method to use belief networks, also some effort should be focused on preventing bias arising in analyses. Where much attention goes into getting unbiased and accurate prior probabilities, in this paper we are more concerned with any bias within the topology; the choice of variables included in the model. When an analyst looks for support for a certain hypothesis, it is easy to get into a so-called tunnel view in which contradicting evidence and alternate hypotheses are neglected. When a plausible alternative perspective is missing in the model, a potential bias is present yet invisible. It seems impossible to always exclude such a bias, but applying certain strategies in the design of a belief network may lead to more balanced and less biased models. Among others, the following strategies might be considered. Firstly, different domain experts can add an alternative point of view to the same model. Secondly, each domain expert can work independently on a different hypothesis or counter-hypothesis. And finally, domain experts can design reusable templates that are not tailored for a specific case, but for generic classes of cases. Whatever combination of strategies may work best to avoid a bias, the case for a flexible and modular way

[1] The software packages HUGIN Expert, Netica and GeNie are respectively found at: `http://www.hugin.com`, `http://www.norsys.com/netica.html` and `http://genie.sis.pitt.edu/`

to design belief networks to aid better decision making should be apparent.

Various systematic techniques are available to guide the modeling of a belief network in a systematic manner. Many of these generate a belief network by translation of another type of model, e.g. ontologies [19], rule-based systems [11], causal maps [16], or by merging quantitative and qualitative statements in a canonical form [5]. However, all these techniques rely on a sound understanding of the application domain to establish the qualitative aspect of a belief network: the topology of the graph. When a domain is modeled that is dynamic in nature and of which causality is not fully known, the technique used to construct a belief network must above all be modular and easily extendible as new insights constantly change the perspective of what variables matter to the hypotheses of interest.

This led to the development of the hypothesis management framework (HMF) at TNO. This design pattern enables a domain expert to independently create and maintain a belief network, and an analyst to evaluate evidence in a criminal investigation. The HMF is a modular belief network structure that is easily expandable by the users themselves, while remaining reliable for decision support. The HMF adds a layer of abstraction to the belief network, so the belief network can be kept hidden from the user. Multiple users can independently modify or extend the model based on his or her domain knowledge. The HMF ensures that all parts of the model remain a coherent whole, suitable for consistent reasoning.

2 THE PURPOSE OF HMF

While devising the HMF design pattern we had one particular goal in mind: to enable the design of modular and extendible Bayesian models for users that are no Bayesian specialist. Once a first version of a model has been developed, it should be easily extended and maintained later-on. It is likely that the set of variables as well as the subjective priors for conditional probability tables require regular revisions as the field of investigation changes over time. Therefore it should be possible to reconsider the set of variables, without having to elicit all of the priors on each change of the model. The need for multiple revisions of a developing model was addressed by the AI group at the University of Kentucky in [14]. A design pattern should preferably be such that it enables the use of templates, generalized submodels within the belief network, that can be maintained independently by a group of domain experts. Such templates should be applicable within multiple belief networks.

To maximize its applicability in real world applications the following two requirements should be met:

1 *Reliability (or consistency)* The belief network should capture the knowledge of domain experts. Given the same set of evidence, the domain experts should agree on the same most likely hypotheses and the results of the model should intuitively make sense.

2 *Usability* The number of priors to be elicited should be kept to a practical minimum. We prefer to have a limited set of well founded priors, rather than a larger set of priors of which the domain expert is less confident. Conditional probability tables with a small set of priors are easier to maintain and validate, especially when the number of conditioning parent variables is limited. Furthermore, it should be unambiguous to domain experts (as well as the analysts) what the variables and their priors stand for.

These requirements are indeed very common, and generally accepted as basic requirements in the context of system development. We think, however, that they are hard to comply with without the use of a generalized framework.

3 AN OVERVIEW OF HMF

The HMF places each variable of interest within a predefined structure, as visualized in Figure 4(c). Furthermore it prescribes which variables may be instantiated with evidence, and for some variables the content of conditional probability tables. All variables must be categorized by the user in hypotheses, indicators or information sources. Each type has its own place and role within the topology of the belief network:

1 *Hypotheses* are statements of which we would like to get a posterior probability distribution. In general, hypotheses are unobserved. The user can specify unconditional priors for each hypothesis, or use a uniform nondiscriminative distribution instead. As an option, one can add alternative hypotheses to represent known facts that explain observed indicators in an other way than existing hypotheses.

2 *Indicators* are statements related to hypotheses. Knowledge of an indicator helps to reveal the states of related hypotheses. Indicators describe events that are dependent on the occurrence of one or more hypotheses. Causal relations between hypotheses and indicators are not always obvious, or present at all. Indicators are assumed to be

'caused' by hypotheses, not the other way around. For each relation between an indicator and a hypothesis, a domain expert should specify conditional probabilities for that specific relation.

3 *Information Sources* are used to express the reliability of sources related to an indicator, when the user does not want to enter 'hard evidence'. For instance, an information source may be a report, a sensor or a person. An indicator can be associated to multiple information sources.

Although common, it is not necessary for an arc in a belief network to imply causality. The HMF makes use of this freedom by taking a more abstract perspective on the relations between variables of interest. The structure is based on the relatively simple notion of hypotheses and indicators. Indicators may all support or contradict any of the hypotheses, but the indicators themselves are assumed independent of one another. Hypotheses are independent (root nodes) and typically have many children. Quite similar, so-called 'naive Bayes' structures [6], have been effective in other areas where causality is unknown or too dynamic in nature (e.g. e-mail spam filtering [15]).

If more structure is desired, this modeling style may be applied in a recursive fashion in which a hypothesis may have sub-hypotheses, who are modeled in an similar way. This is not demonstrated in this article.

It is good practice to use a causal model whenever possible [17], and it should be stressed that HMF does not aim to substitute such models. The HMF design pattern is specifically designed for domains in which causal dependencies are debated or not fully known. As pointed out by Biedermann and Taroni [3], in forensic science the availability of hard numerical data is not a necessary requirement for quantifying belief networks and Bayesian inference could therefore be used nonetheless. By using HMF, a Bayesian model can be constructed even when the qualitative aspects of a belief network are hard to obtain.

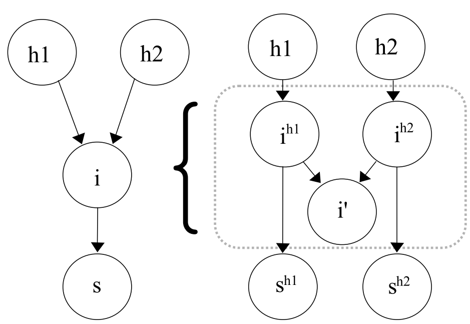


Figure 1: Indicators are substituted by multiple intermediate variables and one combining variable.

There are various options to elicit priors for such large CPTs. One could apply linear interpolation over a subset of elicited priors [20], but this requires more elicited priors and is less flexible than the solution found for HMF. Rather than connecting indicators directly to hypotheses, the HMF uses intermediate variables. In this article all variables are booleans. This is not a strict requirement, but a general recommendation when it simplifies the elicitation of prior probabilities. Elicited priors will be stored in the intermediate variable between the indicator and the hypothesis. This reduces the number of prior probabilities to elicit, and conditions to consider for each prior. In fact, the HMF splits up each indicator in multiple variables (Figure 1): one or more intermediate variables (i^{h1}, i^{h2}) and a variable that combines them (i'). For three hypotheses i would require 16 priors, instead of 12 priors for the three intermediate variables together.

When evidence is available for an indicator, we instantiate all associated intermediate variables. Alternatively, one can use information sources. An information source for an indicator (s in Figure 1) may exist as multiple variables with identical priors in the HMF belief network (s^{h1}, s^{h2}). The priors of an information source variable represent the reliability of the source in regard the associated indicator. Information source variables are children of intermediate variables, and have only one parent and no children. Either all information sources of an indicator are instantiated for evidence, or all associated information source variables. Instantiating intermediate variables of an indicator d-separates information sources from hypotheses, rendering all information sources for that indicator obsolete.

When there is no evidence for an indicator, the combining indicator variable (i') will resemble the posterior probability of the original indicator (i) by taking the average probability of all intermediate variables. This information is useful to predict the likelihood of unobserved indicators or for selecting the most influential unobserved indicator. Equation 1 is used to construct the conditional probability table of the combining indicator variable. Note, that the HMF does not use a logical function (e.g. OR/MAX, AND/MIN or XOR). Logical functions that assume independence of causal influence, in a discrete or noisy variant, have been long in use [8] as a solution for variables with many parents. An extensive overview of such methods are described by Diez and Drudzel in [4]. Although many alternatives may be considered, our preference goes to an averaging method to avoid scalability problems. The scalability problem will be further discussed in Section 5, while the results of using the averaging method in Equation 1 are discussed in Section 6.

$$P(X|parents(X)) = 1.0 - \frac{maxValueOf(parents(X))}{valueOf(parents(X))} \quad (1)$$

This article focuses on how HMF can aid the construction of belief networks. It does not elaborate on how a software tool might facilitate this process. Nonetheless, we would like to discuss briefly how we envision such a tool and how the HMF might be presented to the user. We differentiate two types of roles for users: domain experts and analysts. A user may have both roles in practice. By using the HMF, the GUI can effectively hide the underlying belief network from the user. Both types of users need a different user interface.

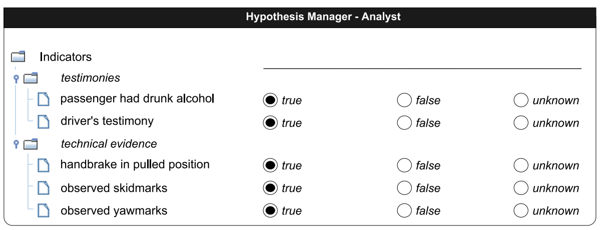


Figure 2: The GUI for an analyst.

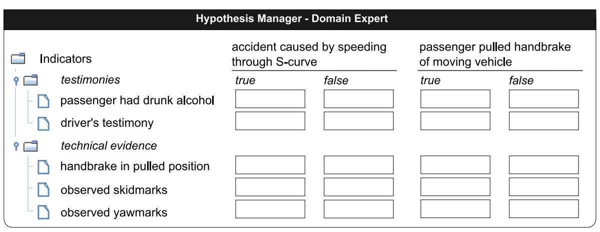


Figure 3: The GUI for a domain expert.

An analyst processes information sources and selects evidence for indicators to support or contradict hypotheses. For analysts the GUI (Figure 2) shows indicators in a foldable tree-like structure. The indicators are organized in categories and sub-categories. For each indicator the analyst can choose a state (e.g. true or false) based on observed evidence. If the analyst is uncertain about an observation, the analyst is given the ability the express the reliability of each information source for that specific indicator. This requires a prior probability for both positive observations and false positives, given that the indicator is a boolean.

Domain experts evaluate the conditional probabilities of an indicator given an hypothesis, and choose prior probabilities for hypotheses. The GUI should enable a domain expert to construct and maintain a list of indicators and hypotheses. A domain expert is responsible for relating indicators to hypotheses in a sensible manner, and assign conditional probabilities to each relation. Figure 3 shows how this may be presented to the domain expert. There is a column for each hypothesis. Assuming only booleans are used, the respective column requires only two elicited priors: one prior for the likelihood of observing the indicator given the hypothesis is true, and another for when the hypothesis is false. Qualitative descriptions or frequencies can be more effective than probabilities [7]. Such notations can be used instead of probabilities, as long these descriptions are consistently translated into conditional probability tables.

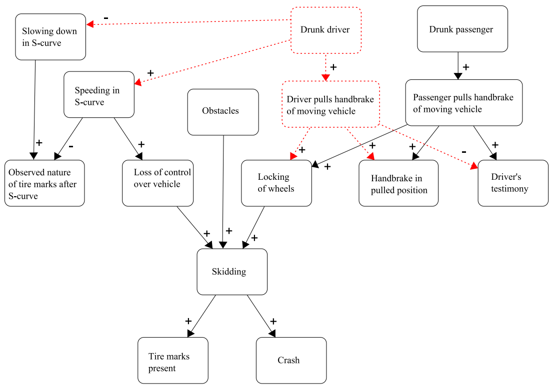
4 HMF WALKTHROUGH BY AN EXAMPLE

To explain how the HMF may be used and why we have chosen this specific topology, we will now discuss three different models based on a civil case concerning a car accident. The first is a logical causal model by Prakken and Renooij [18]. The second is a Bayesian belief network by Huygen [9], directly based on Prakken's logical model. Third and finally, a Bayesian belief network that follows the HMF is constructed for the same case. [2]

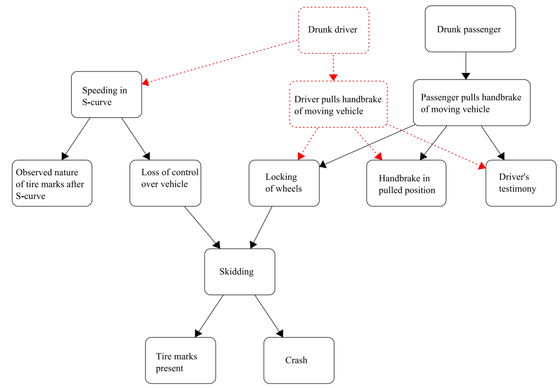
The legal case concerns a nightly car accident involving a driver and a passenger, after a party which both persons attended. The police that arrived at the scene after the accident observed that the car crashed just beyond an S-curve and the handbrake was in a pulled position. The police did observe tire marks (skid marks and jaw marks), but did not observe any obstacles. The driver claims that the passenger was drunk and pulled the handbrake. The passenger claims that the driver speeded through the S-curve. The judge had to decide whether it is plausible that the passenger caused the accident, rather than the driver.

The logical model about this case by Prakken and Renooij is aimed at reconstructing the reasoning behind the court decision on this case. Figure 4(a) shows the causal structure for the case. Nodes within the structure visualize causal concepts (propositions), and arcs represent causal rules between them. Each arc is annotated to show whether the proposition at the head supports (+) or contradicts (-) the proposition at the tail of the arc. By using abductive-logical reasoning on the structure given evidence for some concepts, one can determine whether other concepts are plausible. Although such a model, a causal map, like the one in Figure 4a may resemble a belief network, it lacks the quantitative information required for Bayesian inference. Nadkarni and Shenoy [16] discussed how a causal

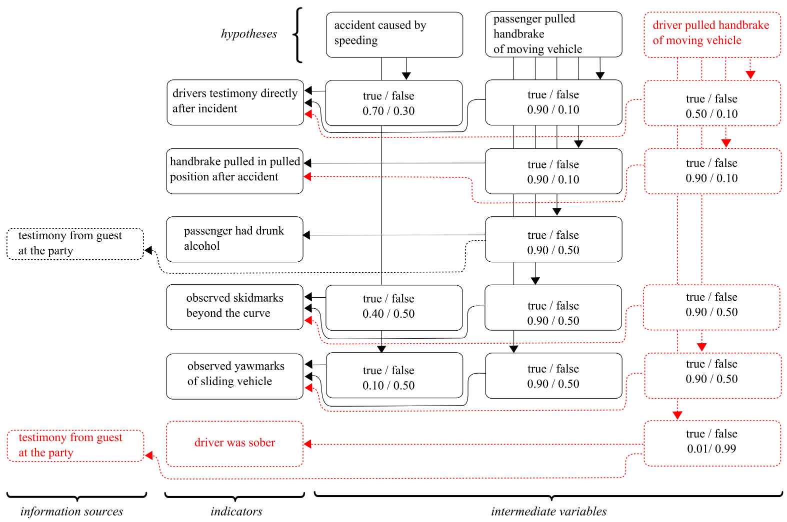
[2] the belief networks discussed in this article are available for download at: `http://www.science.uva.nl/~spg`



(a) A logical causal model by Prakken and Renooij.



(b) The belief network by Huygen.



(c) A variant that uses the HMF design pattern.

Figure 4: Three different models of the same case. The colour red is used to highlight the proposed extensions.

map, can be used as a foundation for constructing belief networks when supplemented with casual values that express the strength of a causal connection.

There is evidence for the following facts: $\neg$*obstacles*, *tire marks present*, *observed nature of tire marks after S-curve*, *handbrake in pulled position*, *driver's testimony* and *drunk passenger*. The hypotheses *speeding in S-curve* and *loss of control over vehicle* explain two facts but contradicts three others. Whereas the hypothesis *passenger pulled handbrake of moving vehicle* explains three rules and contradicts nothing. This makes the drivers point of view more convincing.

Huygen used the causal model of Prakken to construct a belief network for the same case (Figure 4b). The topology was slightly changed: the node for *obstacles* has been removed and the propositions for *speeding* and *slowing down in S-curve* have been replaced by a single boolean that represents both. Furthermore, each node is accompanied with a conditional probability table or prior probability distribution (not visible in Figure 4(b)). This effectively replaces the annotations along arcs in the causal map. Huygen decided not to use evidence for variables on tire marks, because in the sentence of the court it was not explicitly stated that the nature of the tire marks were proof for not speeding, but gave insufficient support for the suggestion that the driver had speeded. Huygens suggests to change the priors, when one would like to use this evidence.

Given evidence for: *pulled position*, *driver's testimony*, *passenger drunk* and *crash*, it is highly likely that the passenger pulled the handbrake ($\approx$ 100%). Since the evidence against the passenger explains away the car crash, it is unlikely that the crash was caused by lost control of the vehicle after speeding through the S-curve (0.1%). The bayesian belief network comes to the same conclusion as the causal map of Prakken and Renooij.

When we model the same case using the HMF, we get a radically different topology (Figure 4(c)) that does not resemble the causal map of Prakken and the belief network of Huygen. Both claims are modeled as hypotheses in the HMF model: *accident caused by speeding* and *passenger pulled handbrake of moving vehicle*. These hypotheses correspond to similarly named predicates in Figure 4a and probability variables in Figure 4b. Uniform probability distributions were used as priors for these hypotheses. We use indicators to support our beliefs in the hypotheses, these are: *driver's testimony directly after incident*, *handbrake in pulled position after incident*, *passenger had drunk alcohol*, *observed yawmarks of sliding vehicle* and *observed skidmarks beyond the curve*.

By choosing different priors, the evidence for tire marks is now usable. Some intermediate variables that relate facts with the two hypotheses are no longer in use. These are *locking of wheels* and *loss of control over vehicle*. The information source of *passenger had drunk alcohol* is undisclosed. Suppose the source was a guest at the party, than the reliability of this testimony is represented by an information source variable (Figure 4(c)).

Given the available evidence, we get a high likelihood for the passenger pulling the handbrake of the moving vehicle ($\approx 100\%$). The propability for speeding is much lower ($\approx 27\%$), and therefore far less convincing.

All three approaches can adequately model the case and derive equally sensible conclusions. Abductive-logical reasoning over a causal map explains the logical correctness and contradictions of propositions. The advantage of a Bayesian approach is that by quantifying influence, it is able to give insight in what hypothesis is most credible as well as the relevance of evidence. The models of Prakken, Renooij and Huygen are based on a causal map. Although HMF follows a different approach to the construction of belief networks, and therefore uses a rather different topology, it does derive the same conclusions.

5 ISSUES REGARDING EXTENDIBILITY

Extendibility as well as modularity are important requirements. The models by Prakken and Huygen are 'static' models in the sense that they were designed to model one single case with a fixed set of evidence and hypotheses. This is feasible when consensus has been developed on all aspects of the case. However, supporting decision making at an earlier stage requires a high level of flexibility. The HMF was developed to facilitate decision making when the set of evidence (or indicators) and hypotheses is still evolving and a constant topic of discussion. Models designed with the HMF are flexible, meaning that a model is decomposable into independent modules. So that each module can be maintained or extended by a different domain expert. This section will discuss issues that concern the extendibility of models developed with the HMF. These issues will be illustrated by extending the existing models from the previous section.

We have pursued extendibility by modular independence of the elicited priors. When an indicator is added to the model, the only priors to elicit are those for the intermediate nodes of that specific indicator. Priors that were elicited before do not have to be reconsidered. The same holds for adding hypotheses. We will illustrate this by considering an additional hypothesis for the car accident case. Suppose the *driver* pulled the handbrake of the moving vehicle. If the driver was under influence of alcohol, that would have also influenced the driving behavior and therefore the likelihood of speeding as well as the possibility of pulling the handbrake of the moving vehicle. In all three models we would have to add and update existing prior knowledge.

To add the alternative hypothesis to the logical model of Prakken and Renooij a proposition is needed for the new hypothesis, and another to represent the possibility that the driver was under the influence of alcohol. These additional causal relations are highlighted in red in Figure 4(a). Together, these additions extend the existing set of 12 rules with 6 more.

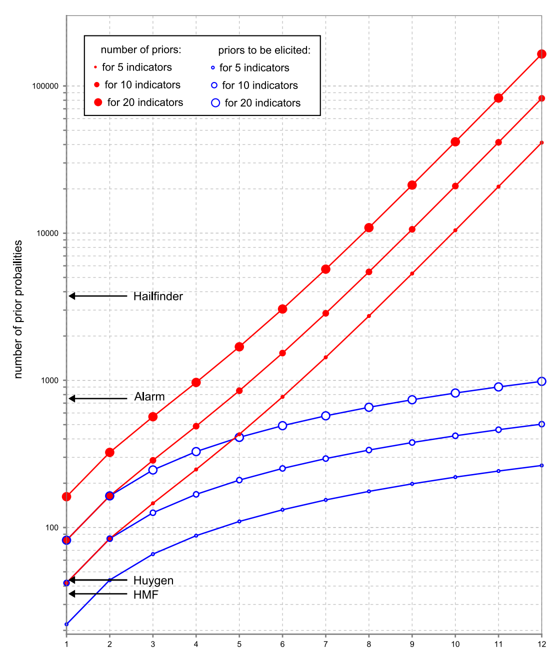


Figure 5: How extending the model affects the number of priors to elicit.

Table 1: Extending the models.

priors	Prakken	Huygen	elicited HMF
in original model	12	44	36
after extension	18	64	58
unchanged	12	30	36
updated and added	6	34	22
relative workload	50%	77%	61%

When we add similar variables and relations to the belief network of Huygen, we need to specify new conditional probability tables for *locking of wheels*, *handbrake in pulled position* and *driver's testimony*. Furthermore, we would have to replace the prior probability distributions of *speeding through S-curve* with a new conditional probability table. These changes comprise the elicitation of 34 new priors that substitute 14 previously elicited priors.

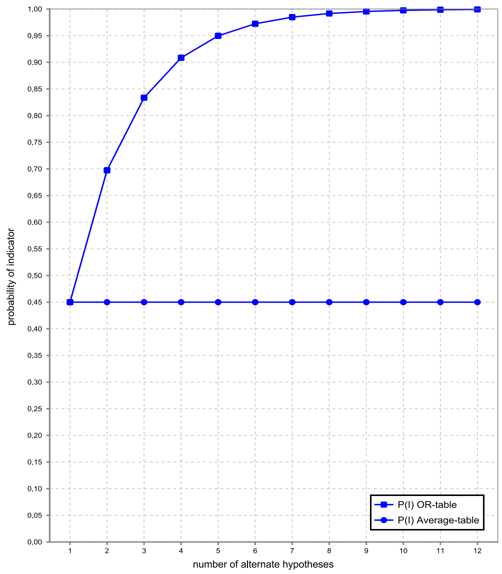
To add to the HMF model the hypothesis *driver pulled the handbrake of the moving vehicle*, requires a new column in the model in Figure 4. The possibility of the driver being under the influence of alcohol is modeled as an indicator, which adds a new row to the model. Table I shows how many elicited priors are required for extending the models. The extensions of the HMF model comprise only 22 elicited priors, all 36 existing priors remain unchanged. This makes HMF considerably cheaper to extend than the belief network of Huygen. The original causal model of Prakken is even simpler to extend. That model, however, lacks quantitative support for probabilistic inference.

As the car accident case shows, the HMF is tolerant to extensions. Figure 5 shows the general effect of adding hypotheses and indicators to a model by outlining the maximum number of elicited priors. While the total number of parameters grows exponentially when more hypotheses are added, the amount of elicited priors grows in a linear fashion. The figure assumes the worst case in which each indicator is associated to all hypotheses. Although the model assumes boolean variables and two priors for each intermediate variable would suffice, it is assumed that *all* priors for intermediate variables are elicited as well as a prior probability distribution for each hypothesis. Note that we have excluded all other parameters that require elicitation such as variable names and state definitions. As a reference Figure 5 includes the number of priors of Hailfinder (3741), Alarm (752), the original belief network of Huygen (44) and the HMF model from Section 4 (36). The extensions proposed in this Section were excluded from the HMF model.

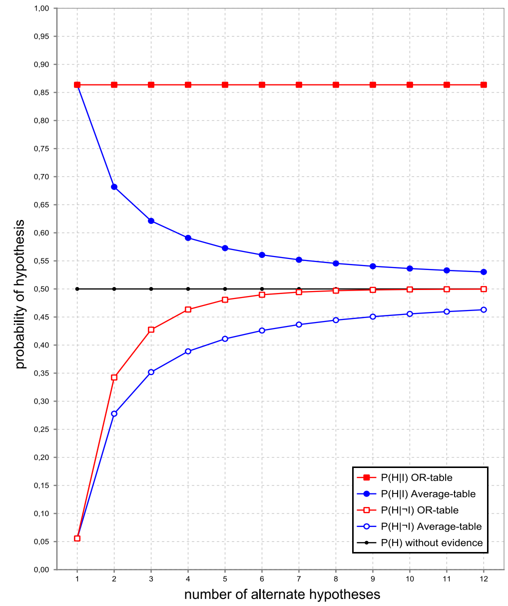
As mentioned in Section 3 indicators are modeled by intermediate variables and one combining variable. The more hypotheses are associated to an indicator, the more probabilities of intermediate variables will have to be combined. On each extension the combining variable gets an extra parent, and as a consequence its conditional probability table (CPT) doubles in size. In the HMF an averaging function has been chosen as the preferred option for these CPTs. By default, the CPT of a combining variable effectively takes the average posterior distribution of all intermediate variables (Equation 1).

Arguably, one might find a logical OR-function [8] more intuitive. However, we have chosen not to use an OR or AND function for these CPTs since a methodical bias may arise in the model if it is extended. A practical drawback of using OR-tables in this situation arises when more than (approximately) five alternative hypotheses are connected to an indicator. By adding more parents to a deterministic OR-table the probability for the child variable quickly converges to unity, or alternatively a pre-defined upper bound. This is shown in Figure 6(a). It is likely that this will lead to unintentional overestimation of the occurrence of unobserved indicators. This can be illustrated by extending the belief network of Huygen, where the variable *locking of wheels* is modeled as an OR-table with an upper bound of 0.80. Suppose the case would be extended to include one or two additional drunk backseat passengers who may have pulled the handbrake of the moving vehicle. The extra backseat passengers are modeled in the same way as the passenger in front, using the original priors $P(locking|pulled) = 80\%$ and $P(locking|\neg pulled) = 0\%$ (where pulled is *true* when any of the persons in the vehicle pulled the handbrake). Given that the driver is sober and all passengers are drunk, the probability of *locking the wheels* increases rapidly (one drunk passenger: 2.4%, two drunk passengers: 4.7%, three drunk passengers: 7.0%). Even when we have not instantiated any other variables (e.g. *crash* or *driver's testimony*). After these extensions, one might like to reconsider the original priors of $P(pull|drunk)$ to prevent overestimating the probability of locked wheels. This potential problem is avoided when the method in Equation 1 is used.

Another potential problem that is associated with OR-tables is the asymmetric influence of an indicator: positive observations have less impact than a negative observation. This is shown in Figure 6(b)). Where observed indicators will only have marginal impact on hypotheses when observed true, the impact on intermediate variables of an indicator observed as *false* is deterministic and therefore usually stronger. It is likely that the user will be unaware of these effects. This makes the model relatively vulnerable to errors in the priors. Therefore, we advice to use Equation 1 as the default method. Other methods for constructing CPTs of combining variables may hinder extending



(a) The likelihood of an indicator.



(b) The impact of evidence for an indicator.

Figure 6: Extending the model affects the probabilities.

the model.

6 ISSUES REGARDING RELIABILITY

To evaluate the outcomes of HMF belief networks we have translated the Asia belief network, as introduced by Lauritzen and Spiegelhalter in [12], into the HMF format.

We will use abbreviations that correspond to the first character of each variable. The original model is shown in Figure 7 (left), the HMF version of Asia is shown on the right. In the HMF model of Asia we distinguish hypotheses: $\{b, l, t\}$, indicators: $\{s, v, x, d\}$ and intermediate nodes: $\{s_b, s_l, v_t, x_b, x_l, x_t, d_b, d_l, d_t\}$. The variable $TbOrCa$ is missing from the HMF model, which in the original belief network combines the probabilities of tuberculosis and lung cancer with a logical OR function has become obsolete.

In the HMF model of Asia, the prior information for the indicators is specified separately for each associated hypothesis. This assumes that the influence of e.g. lung cancer on dyspnea is unaffected by bronchitis. The following probabilities will have to be elicited from a domain expert, when using HMF on Asia. Unconditional priors for each hypothesis: $P(b)$, $P(l)$, $P(t)$ and conditional priors for all intermediate nodes: $P(s_b|b)$, $P(s_l|l)$, $P(v_t|t)$, $P(x_b|b)$, $P(x_l|l)$, $P(x_t|t)$, $P(d_b|b)$, $P(d_l|l)$, $P(d_t|t)$.

The Asia model uses only boolean variables and therefore only one probability for each hypothesis has to be elicited and two for each association of an indicator with a hypothesis. For Asia this gives a total of 21 probabilities. In this case the priors for the hypotheses and intermediate nodes were derived from the joint probability table of the original Asia belief network.

We computed the posteriors of the hypotheses for all possible scenario's of evidence for the indicators. In each of these scenarios each indicator was either observed or not. Note that we instantiate the intermediate nodes for evidence, rather than the combining variables. As mentioned in Section 3 an indicator is represented by both intermediate variables and a combining variable. The conditional probability table of the combining variable is implemented by Equation 1, whereas the elicited priors are stored in the intermediate variables. Instantiating only the combining variable would undervalue those elicited priors.

The results are shown in Table II. For each indicator and hypothesis, the table shows the average and maximum absolute difference in posteriors, as well as the Jensen-Shannon divergence [13]. The bottom row shows the percentage of scenario's in which the outcomes (i.e. the most likely state) for the variables were equal. Especially this last criterion is important for decision making, as the 'real' priors and posteriors will always be open to debate when a causal model is hard to obtain. The table shows that while posterior distributions may vary between both versions, on average the difference is relatively small (< 4 percentage points). For almost all scenario's the outcomes

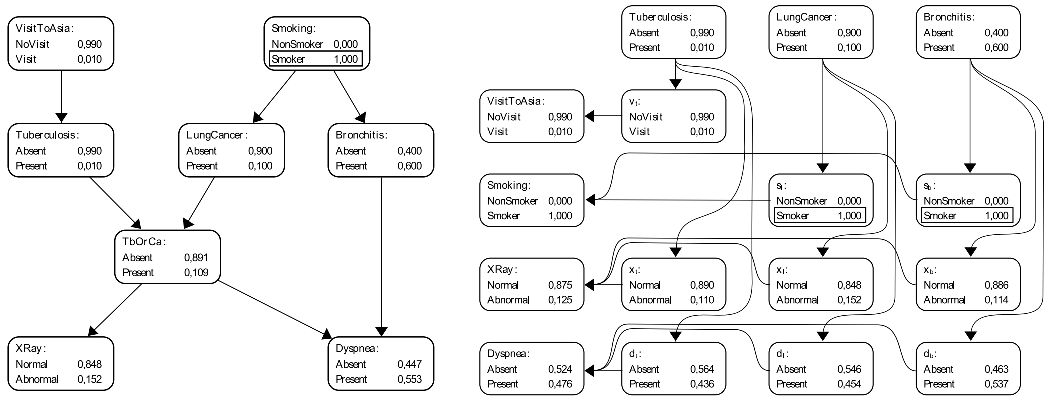


Figure 7: Left: the original Asia belief network. Right: HMF version of Asia. Both with evidence for *Smoking*.

Table 2: Divergence between HMF verion of Asia and the original.

vertice	d	v	x	b	t	c	s
max dif	0,162	0,004	0,071	0,308	0,193	0,209	0,095
av. dif	0,023	0,000	0,009	0,036	0,020	0,017	0,014
max J-S	0,021	0,000	0,006	0,074	0,029	0,035	0,008
av. J-S	0,002	0,000	0,001	0,006	0,002	0,002	0,001
match(%)	91,4	100,0	100,0	97,5	98,8	98,8	97,5

are identical. The few exceptions are caused by the synergistic effect between an abnormal X-Ray and the presence of dyspnea. This synergistic affect is absent in the HMF version, and in those situations we get the relatively large differences in the posterior distributions of bronchitis and long cancer.

7 CONCLUSIONS

The current HMF design pattern is extendible and modular. In our opinion the HMF succeeds in its purpose. We have confidence that HMF comes as a relief to those application domains that so far have been relatively underequipped with practical decision support tools, due to the lack of 'hard and solid' domain knowledge that can be used as a basis for probabilistic models.

The arrangement of the HMF supports a working method which deals with tunnel-view in a well considered manner. The HMF will not explicitly reduce or prevent bias occurring within the topology of a model. However, it offers the possibility to use certain strategies during the design of a model which lead to more balanced and thus less biased models. Using such strategies will enlarge the awareness about tunnel-view (and bias) and as such may partly prevent it.

Although the requirements of reliability and usability are not validated by domain experts and analysts, several issues concerning these requirements have been discussed in this paper. The Asia example shows that posteriors via a HMF model can be quite similar to those derived via a belief network based on causality. The issues that we have encountered so far in applying belief networks for criminal investigations have been addressed in this paper. However, it is a continuous effort to further improve the HMF.

8 FUTURE RESEARCH

One of the complementary wishes of the authors involves a bias measurement combined with automated commentary that highlights useful missing evidence. By calculating how discriminative the indicators and the evidence is to each hypothesis and counterhypothesis, we can evaluate whether tunnel vision may be present. It can also be used to investigate the added value of collecting evidence for unobserved indicators. One way of getting this information is by simulating evidence and evaluate the posteriors of all hypotheses. Since the maximum potential impact of an indicator may only occur at a certain combination of evidence for other indicators, the simulation should consider all possible combinations of evidence for all unobserved indicators. This may be a costly operation. Alternatively one may derive the maximum impact directly from the conditional probability tables of the variables, and use message passing to investigate the maximum potential impact of each indicator.

The naive structure of a HMF belief network may in some occasions not capture the targeted effects. In those cases we would like to extend the HMF model with constraining variables that model the synergistic effect between indicators (or in between hypotheses). We have not been able to test such mechanisms in realistic cases so far. Therefore these need further investigation to test the feasibility of adding constraints, and whether the implications of such mechanisms violate the extendibility and modularity.

The HMF has been applied on several study cases based on real data by the authors. In the foreseeable future it is expected that domain experts will work with this framework. Their experience will be very useful for validating the usability and reliability of this method, and for finding ways to further improve it.

References

[1] B. Abramson, J. Brown, W. Edwards, A. Murphy, and R.L. Winkler. Hailfinder: A bayesian system for forecasting severe weather. *International Journal of Forecasting*, 12(1):57–72, 1996.

[2] Ingo Beinlich, Jaap Suermondt, Martin Chavez, and Gregory Cooper. The alarm monitoring system: A case study with two probabilistic inference techniques for belief networks. In *Second European Conference on Artificial Intelligence in Medicine*, 1988.

[3] A. Biedermann and F. Taroni. Bayesian networks and probabilistic reasoning about scientific evidence when there is a lack of data. *Forensic Science International*, (157):163–167, 2006.

[4] F. J. Díez and M. J. Druzdzel. Canonical probabilistic models for knowledge engineering. Technical Report CISIAD-06-01, UNED, Madrid, Spain, 2006.

[5] Marek J. Druzdzel and Linda C. van der Gaag. Elicitation of probabilities for belief networks: Combining qualitative and quantitative information. In *Proceedings of the Eleventh Annual Conference on Uncertainty in Artificial Intelligence*, pages 141–148. Morgan Kaufmann, 1995.

[6] N Friedman, D Geiger, and M Goldszmidt. Bayesian network classifiers. In *Machine Learning*, volume 29, pages 131–163, 1997.

[7] Gerd Gigerenzer and Ulrich Hoffrage. How to improve bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102:684–704, October 1995.

[8] Max Henrion. Some practical issues in constructing belief networks. In *Proceedings of the 3rd Annual Conference on Uncertainty in Artificial Intelligence (UAI-87)*, pages 161–174, New York, NY, 1987. Elsevier Science.

[9] P.E.M. Huygen. Use of bayesian belief networks in legal reasoning. In *17th BILETA Annual Conference*, 2002.

[10] J. H. Kim and J. Pearl. A computation model for causal and diagnostic reasoning in inference systems. *Proceedings of the 8th International Joint Conference on AI*, pages 190–193, 1983.

[11] M. Korver and P. Lucas. Converting a rule-based expert system into a belief network. *Medical Informatics*, 18(3):219–241, 1993.

[12] S.L. Lauritzen and D.J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Readings in uncertain reasoning table of contents*, pages 415–448, 1990.

[13] J. Lin. Divergence measures based on the shannon entropy. *Information Theory, IEEE Transactions on*, 37(1):145–151, 1991.

[14] Krol Kevin Mathias, Cynthia Isenhour, Alex Dekhtyar, Judy Goldsmith, and Beth Goldstein. When domains require modeling adaptations. In *4th Bayesian Modelling Applications Workshop at UAI*, 2006.

[15] V Metsis, I Androutsopoulos, and G Paliouras. Spam filtering with naive bayes: which naive bayes. In *In 3rd Conference on Email and Anti-Spam, Mountain View, ca*, 2006.

[16] Sucheta Nadkarni and Prakash P. Shenoy. A causal mapping approach to constructing bayesian networks. *Decision Support Systems*, 38(2):259–281, November 2004.

[17] J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2000.

[18] H. Prakken and S. Renooij. Reconstructing causal reasoning about evidence: a case study. In *Legal Knowledge and Information Systems. JURIX 2001: The Fourteenth Annual Conference*, pages 131–142. IOS Press, Amsterdam, The Netherlands, 2001.

[19] Rita Sharma, David Poole, and Clinton Smyth. A system for ontologically-grounded probabilistic matching. In *Proceedings of the Fifth UAI Bayesian Modeling Applications Workshop*, 2007.

[20] B.W. Wisse, S.P. van Gosliga, N.P. van Elst, and A.I. Barros. Relieving the elicitation burden of bayesian belief networks. *Sixth Bayesian Modelling Applications Workshop on UAI 2008.*