Building Domain Models by Novices in Stochastics: Towards the Probabilistic Semantics of Verbalized Stochastic Relations

C.Möbus[1,2], O. Schröder[2]

[1] *FB Informatik, University, D-26111 Oldenburg, Germany*
[2] *OFFIS Institute, Escherweg 2, D-26121 Oldenburg, Germany*
E-Mail: {moebus, schroeder} @informatik.uni-oldenburg.de

Abstract: In this paper we describe a knowledge acquisition method which makes it possible to teach novices to construct Bayesian network models *of their own domain***.** We and others had to realize that there is a severe knowledge acquisition bottleneck. It is nearly impossible to teach novices how to construct Bayesian net models of their own domain because of the huge number of conditional probabilities that are needed to describe the links of the Bayesian directed acyclic graph (dag). Because of this you have to use "toy" data from textbook examples. This leads to motivational problems because novices are often willing to adapt a new methodology only when it promises an efficiency gain in solving problems without imposing new ones. They expect at least in principle a solution sketch which feasibility can be demonstrated. So we offer the possibility that the students can describe a model of their own domain in verbal terms. The system compiles these statements into a dag. Furthermore, in this paper two methods are proposed that allow the acquisition of quantitative data from the verbally stated qualitative information. The former is needed for the application of the Bayesian network for inference tasks. One of the methods is based on likelihoods, the other one is based on frequency distributions. An important advantage of the latter method is that it substantially reduces the number of necessary knowledge acquisition steps.
Both methods enable the plugin of probability tables which denotate the links of the dag, and which are a necessary part of a Bayesian net. Thus the tables are solely derived from verbal statements about stochastic relations. It is no problem to obtain these verbalizations from domain experts

1. Introduction

Many real-world domains are complex, dynamic, and uncertain. For example, in areas as diverse as business and medicine, the problem solver often has to make decisions when knowledge about the actual situation is incomplete, when relations about important variables are stochastic (or not exactly known), and when unexpected events are possible. This situation is further complicated by the fact that information about stochastic facts and relations is often only stated in verbal form, like for example: "It is well likely that a dislocation of intracranial blood vessels may lead to permanent headache." Even domain experts often hesitate to state numerical relationships, as for example in medicine [9].

On the other hand, domain models should be stated as precisely as possible if they are to be used for further problem solving and inferencing tasks. For example, in medicine a precisely stated domain model is a prerequisite for efficient and effective diagnosis, consultation, and therapy planning. We are currently developing MEDICUS (modelling, explanation, and diagnosis support for complex, uncertain subject matters [3]), a system that supports model building as well as the application of models for diagnosis, consultation, and therapy planning. But MEDICUS can also be applied to stochastic domains other than medicine. An actual example is decision support in the domain of business (project SHAFT). In MEDICUS, uncertainty is handled by the Bayesian network approach. In SHAFT this is extended by influence diagrams [10, 11].

The acquisition and communication of knowledge is a major bottleneck for the construction of complex models in uncertain domains. This problem is exacerbated for novices in probability and uncertainty theory, and for persons (novices or experts) who want to construct uncertain domain models where objective probability data are missing. We think that this problem often even prevents people from building adequate models in uncertain domains. In this paper we want to look at the problem how to support novices in building stochastic models: Is it possible to acquire the quantitative stochastic knowledge necessary for these models from verbally stated "fuzzy" concepts and relations? This means that the quantitative knowledge needed for the specification of a Bayesian network has to be acquired from the verbally stated qualitative knowledge alone. This feature is especially necessary if the user wants to create a Bayesian network model for his *own* domain with some objective probabilities missing. This task can only be efficiently supported if the necessary quantitative knowledge can be acquired in a nonreactive way. There are techniques for acquiring quantitative knowledge from subjects [12] but they are too time-consuming for large domain models. No user will be willing to answer questions about hundreds of conditional probabilities. A more economic method has been developed by Heckerman [4] but it also requires lots of judgements. Approaches trying to assess the (probability) semantics of adverb phrases like "probably", "perhaps", "maybe" etc., and modal verb forms like "should", "will", "may" etc. [7, 15, 16, 18, 19] do not address multivariate distributions which is necessary when we are interested in the semantics of relational terms such as assertions about influences, covariances, or conditional probabilities.

If the semantics of "fuzzy" descriptions of relations can be acquired in the way sketched, then the creation of domain models can be supported by the following steps:

- Letting the user (novice or expert) state verbal assertions about the domain of interest
- Converting these assertions into numerical relationships between the domain variables
- Validating these numerical relationships, and allowing the user criticize and change them
- Letting the user apply the model for diagnosis tasks, consultation tasks, and other inferential applications.

We think that this approach is valuable especially for novices because in this way they can communicate their own model assumptions in a qualitative way and yet make use of the quantitative model for interesting tasks of diagnosis, forecasting, etc. But the approach should also be valuable for experts, at least in those situations where quantitative information is not available.

In this paper, we firstly will give a short overview of MEDICUS. In the main part of the paper, we will focus on the problem of acquiring quantitative information from qualitative relational statements. First we will present an extension of an existing approach based on "fuzzy" membership functions to multivariate situations (*"likelihood approach"*). We will show that this runs into some problems. Then we will present an alternative methodology (*"distribution approach"*) that stays within the Bayesian network approach from the beginning. It allows to acquire the semantics of verbal relational terms with a limited number of judgements. It can be shown though that the likelihood approach can also be seen from a Bayesian view. (Figure 9 summarises both approaches.) We will end with a discussion of the open points and directions of further work.

2. An Overview of MEDICUS

MEDICUS is an environment for supporting model building and inference tasks like diagnosis, forecasting, action planning, and consultation in complex, uncertain domains. Uncertainty of knowledge is handled by the Bayesian network approach. A Bayesian network [10, 11] represents knowledge as a set of propositional variables and probabilistic interrelationships between them by a directed acyclic graph (dag). The variables are represented by the nodes of the graph, and the relations by directed arcs. The relations are quantitatively described by conditional probability matrices (each variable conditioned on its parents in the network) that define a joint probability distribution of the variables. Independencies between variables are represented by omitting arcs, which simplifies the corresponding conditional distributions.

An important reason for choosing the Bayesian network approach is that it supports qualitative reasoning. For example, a physician engaged in medical diagnosis proceeds in a highly selective manner [2]. This selectivity can be described by the independencies also present in Bayesian networks. Qualitative reasoning as supported by Bayesian networks seems to correspond closely to human reasoning patterns [5, 6, 17].

One of the main goals of MEDICUS is to assist the learner in developing a model of perceived causes, effects, and other relationships in a domain of interest with Bayesian networks. Bayesian networks provide a precise base for reasoning and communication, and for deriving consequences (in-/dependencies, aposteriori distributions) useful for applications. But especially for novices, it is necessary that the learner is able to state ideas in an informal way which he is used to. Therefore, we developed a simplified-natural-language *linguistic model editor*. After stating his model in this editor, the system can generate an initial graph automatically.

Figure 1 shows a small example with four sentences. The learner creates sentences with the help of a menu. Relational terms are classified based on i) probabilistic concepts of causality [13] organised according to "kind of influence" (positive / negative) and "direction of influence" (forward, backward, or undirected), and ii) has-part / is-a hierarchies. For example, the verb "causes" (second sentence in Figure 1) expresses a forward, positive influence between two variables A and B: $t_A \leq t_B$, $p(B \mid A) > p(B)$. If the learner uses a verb not yet available in the linguistic model editor, he may classify it according to these dimensions. The sentences created by the learner are checked by a definite clause grammar that gives feedback if it detects errors.

The learner may ask the system to create a graph representation for the model specified (Figure 2). In creating the graph, relations are represented by links whose directions depend on the features of the verb being used in the linguistic model editor. For example, if the relation between variables A and B is expressed by a verb designating a forward, positive influence between A and B, then a link is created that points from node A to node B. For relations describing undirected relations (like "corresponds to"), a dialog is evoked where the learner is asked to specify the direction, or to specify another variable as the common cause or effect of the variables in question.

The graph is an initial heuristic proposal. After the initial formulation of the model, it has to be analysed and revised on a qualitative level. It has to be verified that the dependencies and independencis implied by the graph correspond to the assertions stated by the modeller. In Bayesian networks independencies are expressed by missing links. For example, the graph in Figure 2 states that space requirement and haemodynamic irritation are independent, given injury (that is, p(space requirement | injury) = p(space requirement | injury, haemodynamic irritation)).

The qualitative knowledge of the modeller is acquired in three steps:

1. For a case, the modeller specifies the actually known information (for example, "dislocation of vessels" and "injury", see Figure 2). Next, he specifies a diagnostic hypothesis (for example, the hypothesis that the patient might suffer from a "space requirement" process, like for example inner bleeding). Thirdly, he specifies what information he would look for next (for example, "haemodynamic irritation" and "vomiting"). Independencies are obtained from this dialog in the following way: Information not considered relevant to the hypothesis by the modeller, given the history data and symptoms, is independent of the hypothesis. In our example, "permanent headache" was *not* selected, so "permanent headache" and "space requirement" are considered independent, given "injury" and "dislocation of vessels": p(space requirement | injury, dislocation of vessels, permanent headache) = p(space requirement | injury, dislocation of vessels).
2. The modeller states the hypothesis that the graph is consistent with the information specified by her or him in the diagnostic dialog. The system analyses this hypothesis using the d-separation criterion [11]. This may lead to one of the following results: a) the graph and the in-/ dependencies are consistent, b) edges have to be removed from the graph in order to be consistent with the in-/ dependencies, c) edges have to be added to the graph, or d) edges have to be removed from and added to the graph as well.
3. On further request, the modeller may ask the system for modification proposals and an explanation of these proposals. In the example stated, the system proposes for Figure 2 to add an edge from "space requirement" to "haemodynamic irritation" because the modeller specified that "haemodynamic irritation" is informative for "space requirement" given information about "dislocation of vessels" and "injury".

After qualitative refinement of the model, numerical probability tables can be entered, and the model may be applied to inference tasks. Quantification is a major knowledge acquisition bottleneck (see next section). With respect to applications MEDICUS generates qualitative recommendations based on quantitative propagation for diagnosis, consulting, and therapy planning in selected subdomains of environmental medicine. The models are currently developed together with our cooperation partners who are domain experts but not statisticians. The system lists the syndrome hypotheses most probable in the light of actual evidence, and it recommends what symptoms, external factors, and actions to consider next.

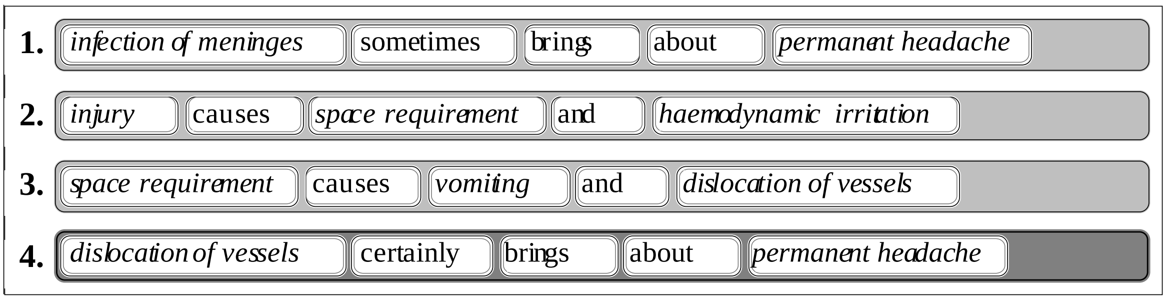


Figure 1: Four sentences created in the linguistic model editor

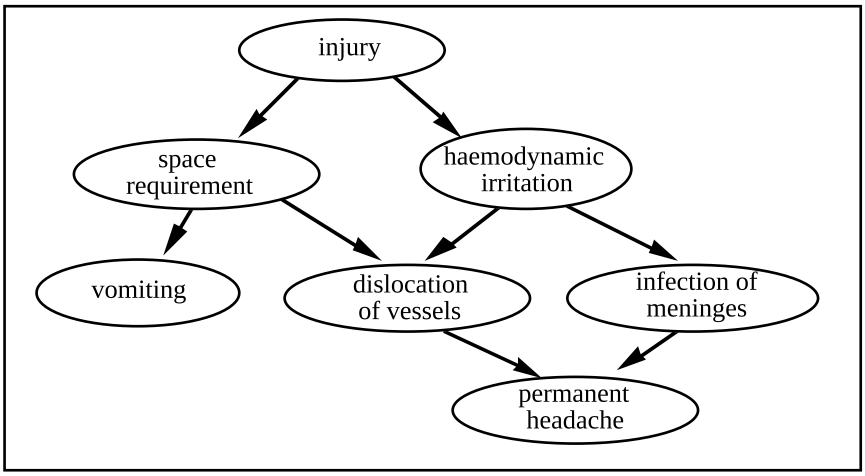


Figure 2: Graph representation generated for the sentences of Figure 1

3. The Compilation of Qualitative Fuzzy Statements about Stochastic Relations into their Quantitative Counterparts

As already indicated, methods have to be developed and applied that enable, encourage, and support novices to model in uncertain domains. The most nonreactive way would be to exploit the information already given by the modeller in natural language modelling assertions stated in the linguistic model editor (Figure 1). Many empirical investigations of the semantics of adverb phrases like "probably", "perhaps", "maybe", etc., and modal verb forms like "should", "will", "may", etc. (as already mentioned above) try to assign membership functions to these linguistic forms, for example by presenting different "wheel of fortune" configurations to subjects. For a list of linguistic terms, the subjects then have to indicate how well each term describes the "wheel of fortune" configuration presented. For example, a wheel of fortune with a winning area of 20% and a loosing area of 80% is better described by the statement "It is possible that I will win" than by the statement "It is very likely that I will win."
But as indicated, these studies do not address the multivariate case. We see two approaches to do this:

- an extension of the membership functions approach to multivariate distributions. This leads to what we call a "likelihood approach".
- a solely probability-based approach ("distribution approach").

4. Extending membership functions to relational terms in the form of likelihoods

We can apply the "wheel of fortune paradigm" to conditional events. This can be achieved by a wheel of fortune configuration as depicted in Figure 3: If spinning wheel A leads to the event "A+", then the wheel "B after A+" is spun, otherwise the wheel "B after A-" is spun. (For simplicity, we only consider binary variables here, although MEDICUS handles multivalued variables as well.).

Next, membership functions of these wheel-of-fortune configurations to a linguistic characterisation like for example "B is a typical consequence of A" have to be obtained empirically. It is possible to assign a probabilistic semantics to these membership functions: "voting semantics" [1]. This means that you can interpret the values of the fuzzy membership function after some normalization as conditional probabilities representing the subjective view of a voting committee. The values of the fuzzy membership function are after some normalization equivalent to the conditional (voting) probability p(hypothesis | evidence). Under a bayesian view this conditional voting probability is equivalent to the subjective estimate of the aposteriori probability for the hypothesis when some evidence has been observed.

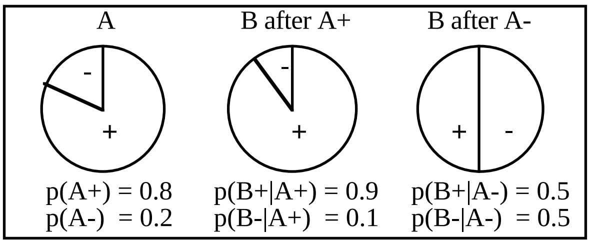


Figure 3: Wheel of fortune for relations between two binary variables A and B

According to probabilistic concepts of causality [13], a positive influence of variable A on variable B can be expressed by p(B+ | A+) > p(B+ | A-), and a negative influence as p(B- | A+) > p(B- | A-). Thus it seems reasonable to hypothesise that fuzzy relations express a certain relationship between p(B+ | A+) and p(B+ | A-), or between p(B- | A+) and p(B- | A-), respectively. This relation can be expressed by the likelihood p(B+ | A+) / p(B+ | A-). Thus the idea seems reasonable that the values of membership to a "fuzzy" concept describing the relation between two variables A and B can be expressed as a function of the likelihood p(B+ | A+) / p(B+ | A-). Furthermore, by taking the likelihood fraction we achieve a reduction by one dimension compared to the number of the underlying conditional probabilities or "wheels of fortunes".

Thus the likelihood ratio that can be derived from *two* wheels of fortune ("B after A+" and "B after A-") is a bivariate analogue to the univariate situation with *one* wheel of fortune pursued in the studies mentioned above. This straightforward extension to multivariate distributions was proposed and sketched in [14].

If we interpret membership functions as probabilities according to the voting semantics, then the probability that a certain matter of facts F is expressed by the verbal phrase V, p(V | F), is proportional to the membership value μV(F) of F for V. If we interpret V as evidence E and F as the hypothesis H, then p(V | F) = p(E | H) is the "causal" probability in Bayesian terms, and μE(H) is the corresponding membership function in fuzzy terms.

Using this assumption, probabilities p(verbal phrase | likelihood) = p(V | F) can be obtained, like for example p("B+ is considered a cue for A+" | p(B+ | A+) / p(B+ | A-) = x). In addition, an apriori distribution of the likelihood, p(p(B+ | A+) / p(B+ | A-) = x), is needed. Then the desired "diagnostic" probabilities p(H | E) = p(F | V) = p(likelihood | verbal phrase) = p([p(B+ | A+) / p(B+ | A-)] | "B+ is considered a cue for A+") can be obtained. The mode of this "diagnostic" probability distribution is the likelihood ratio that best represents the verbal phrase in question: pmax (H | E).

If there is more than one verbal phrase for a certain relationship, for example verbal descriptions provided by different experts (like "B+ is considered a cue for A+", "A+ may cause B+" and so on), then we can combine evidence by the aposteriori probabilities p(H | E1, E2) = p([p(B+ | A+) / p(B+ | A-)] | "B+ is considered a cue for A+", "A+ may cause B+", ...).

In the binary case, the conditional probabilities p(B+ | A+), p(B+ | A-) etc. needed for the Bayesian network can be obtained from the desired aposteriori "diagnostic" probabilities pmax (likelihood | verbal phrase) = pmax (H | E). (For example, if p(B+ | A+) / p(B+ | A-) is known to be c1 and p(B- | A+) / p(B- | A-) is known to be c2, then p(B+ | A-) = (1-c2) / (c1-c2), p(B+ | A+) = (c1-c1c2) / (c1-c2), and so on).

This "likelihood" approach has a serious drawback so that it becomes infeasible in multivariate situations. You have to try many different angles of the "wheels of fortune" and likelihoods to get the most "typical" diagnostic probability pmax (H | E). and the most "typical" likelihood H.

5. An alternative Bayesian-net-based approach with distributional hypotheses

The second approach is based on Bayesian networks and on distributions. The basic idea is i) to let subjects rate the adequacy of selected verbal relational terms as descriptions for some preselected distributions or frequency tables, ii) to backpropagate verbal evidence for distributions of hypotheses, and iii) to compute expected values for their cells with the help of the aposteriori probabilities of those distributions, thus obtaining a "tailored" distribution representing the given set of verbal statements most adequately. (This is of course a hypothesis that can be and has to be tested empirically.) The main advantage of this approach - and its main difference to the approach described above - is that it saves lots of assessment steps: Only a few, selected frequency distributions have to be judged by the subjects.

The other difference between the approaches - likelihoods based on the "wheel of fortune paradigm" vs. distributions - is less fundamental because we could use the wheel of fortune for generating distributions as well, or use distributions instead of fortune wheels for computing likelihoods. But we think that distributions have some empirical advantages over the "wheel of fortune" approach. In the knowledge acquisition and model construction phase, they might be easier to understand for subjects used to work with data material.

We start by assuming several distribution hypotheses (see Figure 4) about the stochastic relation between two binary variables. We assume a set of minimal independent hypotheses so that we can span the whole distribution space. This means that we can generate all distributions by a linear combination of our hypotheses exhaustively. We assume that hypotheses have equal apriori probabilities. The expected value of the apriori hypotheses is a noninformative stochastic relation between the stochastic variables: a distribution of independent variables. The hypotheses are realisations of the hypothesis variable which is the root node in a two-layered Bayesian net.

Figure 5 shows some example distribution hypotheses for two binary variables X and Y. (In this example we used probabilities.) On the right of Figure 5 the expected distribution H is depicted.

The leaf nodes of the net of Figure 4 represent the various empirical evidences for the hypotheses. The evidences are the various verbal statements which can be used to describe stochastic relations: "X correlates rather high with Y", "X has a strong influence on Y", "X weakens the influence from Y on Z",... The root node is linked to leaf nodes by arcs which have to be denoted by conditional probabilities p(evidence | hypothesis). These "causal" probabilities have to be acquired by empirical studies before it is possible to use the Bayes net for the calculation of the expected distribution conditional to the empirical evidence. In fuzzy terms these probabilities correspond to *some* realisations of the membership function $\mu V(F)$.

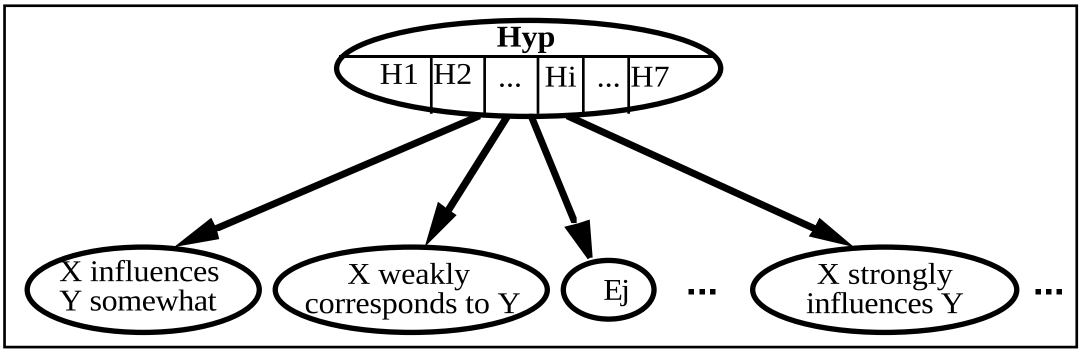
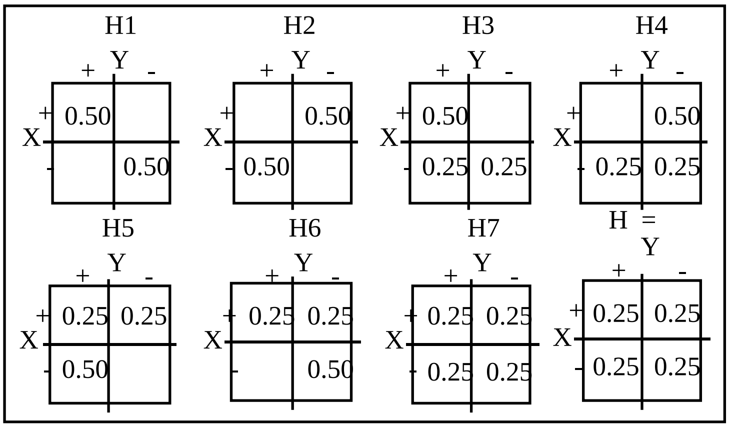


Figure 4: Bayesian network with a concept node (root node) containing seven distribution hypotheses, and nodes representing various verbal relational terms (leaf nodes)

In summary, the expected aposteriori distribution (EAD) can be computed by a three step procedure: (1) acquiring empirical verbal evidence about the stochastic relation between variables (by letting subjects confirm or disconfirm sentences like for example: "It is true that 'X strongly influences Y' describes the situation in hypothesis H2."), (2) propagating this evidence back to the root hypothesis node, (3) combining the hypotheses distributions with their aposteriori probabilities so that the resulting EAD is the expected value H' of the hypothesis distributions computed with the aposteriori probabilities of these hypotheses. This EAD is a subjective estimate. Conditional probabilities obtained from it are plugged into the Bayesian dag as long as we do not have better and objective estimates for the quantitative information. Note that we only need a small number of distributions Hi to be judged by subjects (step (1)).

But there is a problem with this approach: If the distributional hypotheses look like the ones in Figure 5, then not all possible EADs can be generated. If we do not want to constrain the space of possible EADs apriori, we need the four canonical distributions shown in Figure 6.

But it seems reasonable that subjects would have severe trouble judging the adequacy of sentences like the ones in Figure 4 for these distributions. Therefore we propose a three-layerd network as shown in Figure 7. The topmost node contains the hypotheses H' of Figure 6. The middle node contains the hypotheses H of Figure 5. The relationships between the upper two layers again have to be determined empirically by having subjects rate the similarity of each of the hypotheses H to one hypothesis H'i at a time. This information would fully specify the net of Figure 7, and we can compute our EAD by calculating the expected distribution on H'.



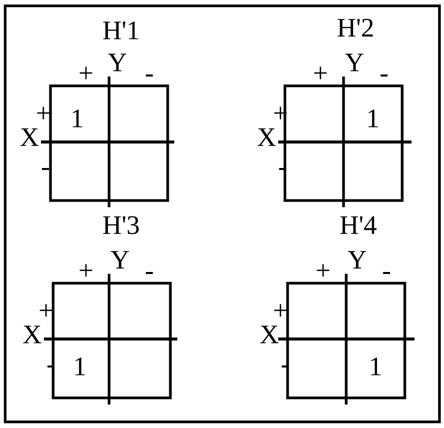


Fig. 5: Bivariate distribution hypotheses. Fig. 6: Distributions necessary for generating all possible EADs

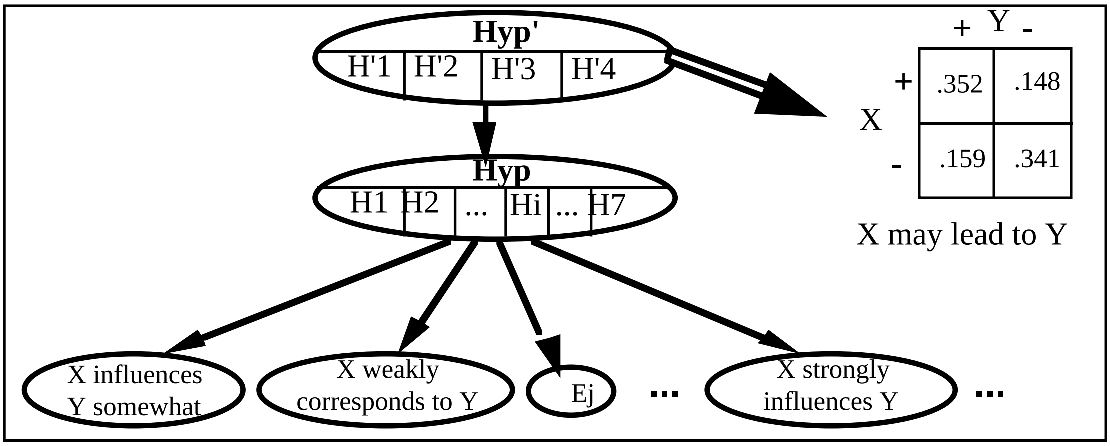


Figure 7: Net of Figure 4 extended by a second concept node

6. Summary and what to do

Figure 8 summarises the last two sections. The input to our approach consists of a given stochastic relation and a (set of) verbal description(s) of this relation. The stochastic relation may be represented by "wheels of fortune" or by a multidimensional probability / frequency table. The table can of course also be generated from the wheel-of-fortune configurations. The information given in the table may be expressed by likelihoods, leading to a dimension reduction. The multidimensional table (or likelihood information as well) is part of a minimal and exhaustive set of hypotheses of a two-layered or three-layered Bayesian net. After acquiring causal probability estimates empirically (the probabilities of verbal statements given each of the hypotheses) for this minimal exhaustive set, aposteriori distributions are computed for the hypotheses that are used for generating an expected hypothesis ("expected table" or "expected likelihood"). This expected hypothesis serves as the base for computing the conditional probabilities needed for the Bayesian network of the modelled domain.

We performed initial empirical studies with the "distribution approach" as depicted in Figure 7. The results seem promising. We found out that a modified approach according to figure 4 with the canonical set hypotheses (figure 6) is empirically tractable and sufficient for our purposes. One of the next steps is to validate the expected distributions generated. This could for example be done by letting subjects specify distributions for given verbal relational statements, or by having them select the most suitable distribution from a set containing the generated expected distribution.

The approach sketched above should of course not be restricted to bivariate distributions of binary variables. We are working at extensions in these directions too.

References

[1] BALDWIN, MARTIN & PILSWORTH, FRIL - Fuzzy and Evidential Reasoning in AI, Wiley, 1995

[2] ELSTEIN, A.L., SHULMAN, L.S., SPRAFKA, S.A. (1978). Medical Problem Solving - An Analysis of Clinical Reasoning, Cambridge: Harvard University Press.

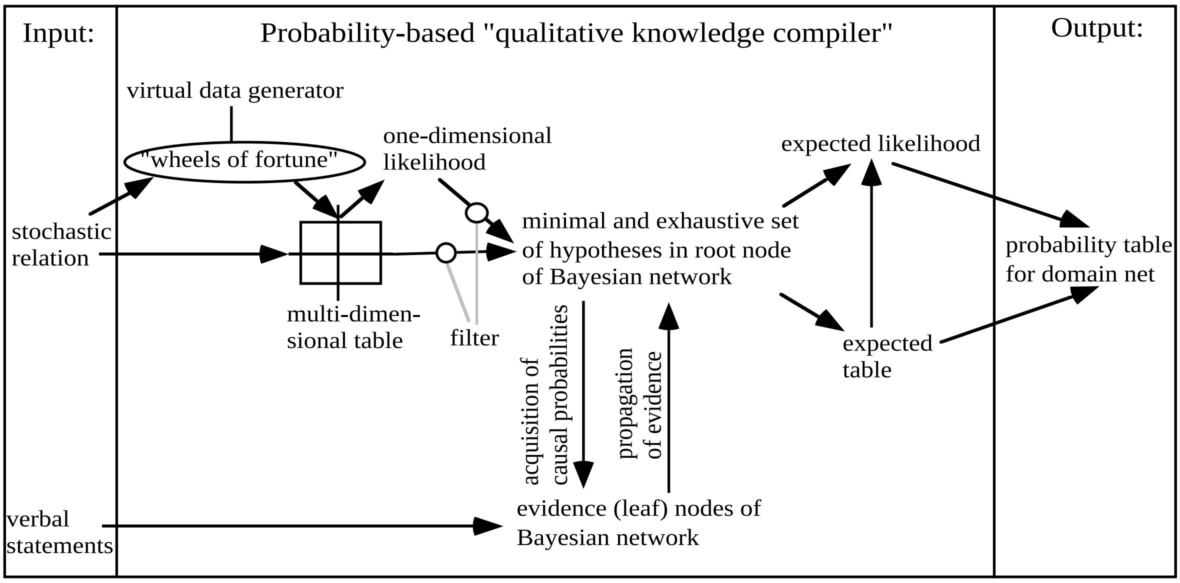


Figure 8: Summary of the "likelihood approach" and the "distribution approach"
for compiling verbal relational assertions into quantitative conditional probabilities

[3] FOLCKERS, J., MÖBUS, C., SCHRÖDER, O., THOLE, H.-J. (1996). An Intelligent Problem Solving Environment for Designing Explanation Models and for Diagnostic Reasoning in Probabilistic Domains. In C. FRASSON, G. GAUTHIER, A. LESGOLD (eds): Intelligent Tutoring Systems. Proceedings ITS 96, Berlin: Springer, 353-362.

[4] HECKERMAN, D.E. (1991). Probabilistic Similarity Networks, Cambridge: MIT Press.

[5] HENRION, M. (1987). Uncertainty in Artificial Intelligence: Is Probability Epistemologically and Heuristically Adequate? In J.L. MUMPOWER, L.D. PHILIPPS, O. RENN, V.R.R. UPPULURI (eds), Expert Judgements and Expert Systems. Berlin: Springer (NATO ASI Series F: Computer and Systems Science), 106-129.

[6] JUNGERMANN, H., THÜRING, M. (1993). Causal Knowledge and the Expression of Uncertainty, in G. STRUBE et al. (eds), The Cognitive Psychology of Knowledge. Amsterdam: Elsevier, 53-73.

[7| KIPPER, B. JAMESON, A. (1994). Semantics and Pragmatics of Vague Probability Expressions, University of Saarbrücken, Sonderforschungsbereich 314 Künstliche Intelligenz - Wissensbasierte Systeme, Bericht Nr. 105

[8] MÖBUS, C., SCHRÖDER, O. (1997). Acquiring qualitative and quantitative knowledge from verbal statements and dialogues in probabilistic domains. In W. BANDILLA, F. FAULBAUM (eds), Advances in Statistical Software 6, 1997, Stuttgart: Lucius & Lucius, 267-274

[9] NAKAO, M.N., AXELROD, S. (1983). Numbers are Better Than Words, Am. J. of Med., 74, 1061f

[10] NEAPOLITAN, R.E. (1990). Probabilistic Reasoning in Expert Systems, New York: Wiley.

[11] PEARL, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufman.

[12] SPETZLER, C.S., STAEL von HOLSTEIN, C.-A.S. (1975). Probability Encoding in Decision Analysis. Management Science, 22 (3), 340-358.

[13] SUPPES, P. (1970). A Probabilistic Theory of Causality, Amsterdam: North-Holland.

[14] SCHRÖDER, O., MÖBUS, C., THOLE, H.-J. (1996). Acquiring Knowledge from Linguistic Models in Complex, Probabilistic Domains. In P. BRNA, A. PAIVA, J. SELF (eds): Proc. of the European Conference on Artificial Intelligence in Education EuroAIED, Lisbon, Portugal: Colibri, 206-212.

[15] TEIGEN, K.H. (1988). The Language of Uncertainty, Acta Psychologica, 68, 27-38.

[16] TEIGEN, K.H., BRUN, W.(1995). Yes, But it is Uncertain: Direction and Communicative Intention of Verbal Probabilistic Terms, Acta Psychologica, 88, 233-258.

[17] WALDMANN, M.R., HOLYOAK, K.J. (1992). Predictive and Diagnostic Learning Within Causal Models Journal of Experimental Psychology: General, 121, 222-236.

[18] WALLSTEN, T.S., BUDESCU, D.V., RAPOPORT, A., ZWICK, R., FORSYTH, B.(1986). Measuring the Vague Meanings of Probability Terms, J. Exp. Psych.: General, 115, 4, 348-365.

[19] WALLSTEN, T.S., BUDESCU, D.V. (1995). A Review of Human Linguistic Probability Processing: General Principles and Empirical Evidence, The Knowledge Engineering Review, 10, 1, 43-62.

Acknowledgements

We thank Wilfried BOHLKEN for assisting in the mathematical and the empirical work.