In *Qualitative Reasoning and Decision Technologies*, N. Piera Carrete & M.G. Singh (eds), pp. 451-460, CIMNE: Barcelona, 1993

Belief Propagation in Qualitative Probabilistic Networks*

Marek J. Druzdzel
Carnegie Mellon University
Department of Engineering and Public Policy
Pittsburgh, PA 15213
marek+@cmu.edu

Max Henrion
Rockwell International Science Center
Palo Alto Laboratory
444 High St, Suite 400
Palo Alto, CA 94301
henrion@sumex-aim.stanford.edu

Abstract

Qualitative probabilistic networks (QPNs) [13] are an abstraction of influence diagrams and Bayesian belief networks replacing numerical relations by qualitative influences and synergies. To reason in a QPN is to find the effect of decision or new evidence on a variable of interest in terms of the sign of the change in belief (increase or decrease). We review our work on qualitative belief propagation, a computationally efficient reasoning scheme based on local sign propagation in QPNs. Qualitative belief propagation, unlike the existing graph-reduction algorithm, preserves the network structure and determines the effect of evidence on all nodes in the network. We show how this supports meta-level reasoning about the model and automatic generation of intuitive explanations of probabilistic reasoning.

1 Introduction

Probabilistic reasoning schemes are often criticized for the undue precision they require to represent uncertain knowledge in the form of numerical probabilities. In fact, such criticism is misconceived since probability theory is rooted in qualitative judgments of conditional independence and relative likelihood, and there are a wide variety of probabilistic schemes that do not require single point probabilities. These schemes range in specificity from purely qualitative schemes such as knowledge maps [7], I-maps [9], and qualitative probabilistic networks [13], to schemes allowing partial numerical specification, such as intervals rather than point probabilities [1, 11].

Our work is based on the qualitative probabilistic network (QPN) representation, introduced by Wellman [13]. QPNs are in essence a qualitative abstraction of Bayesian belief networks and influence diagrams. A QPN requires specification

*This research was supported in part by the Rockwell International Science Center.

of the graphical belief network, expressing probabilistic dependence and independence relations. In addition, it requires specification of the signs of influences and synergies among variables. A proposition a has a positive *influence* on a proposition b, if observing a to be true makes b more probable. Variable a is positively *synergistic* with variable b with respect to a third variable c, if the joint effect of a and b on the probability of c is greater than the sum of individual effects. QPNs generalize straightforwardly to multivalued and continuous variables.

QPNs can replace or supplement quantitative Bayesian belief networks where numerical probabilities are either not available or not necessary for the questions of interest. An expert may express his or her uncertain knowledge of a domain directly in the form of a QPN. This requires significantly less effort than a full numerical specification of a belief network. Alternatively, if we already possess a numerical belief network, then it is straightforward to identify the qualitative relations inherent in it, based on the formal probabilistic definitions of the properties. Examples of queries that can be resolved using QPNs include determining the effect of observations on the probability of a variable of interest. If a network contains decision nodes and a value node, such query with respect to the two can be used to identify dominating decision options [13].

Our main contributions to QPNs are: (1) identification of the conditions for *intercausal reasoning*, or *explaining away*, for example how confirmation of one cause of an observed effect may reduce (or increase) the probability of another cause [5, 6, 15], and (2) a computationally efficient scheme for qualitative belief updating that we call *qualitative belief propagation.* Qualitative belief propagation traces the effect of an observation e on other network variables by propagating the sign of change from e through the entire network. This differs from the graph-reduction approach [13] in that it does not modify the underlying network but rather labels all nodes with the sign of change. In cases where some of a QPN's variables are instantiated, qualitative belief propagation relies on a new qualitative property that we call *product synergy.* This is the condition for intercausal reasoning, or inference across a convergent node in a belief network (where arrows meet head to head). Formal presentation of our algorithm for qualitative belief propagation can be found in [4]. Formal presentation of the definition and properties of product synergy can be found in [5]. In this paper, we review the main concepts of qualitative belief propagation and discuss how this aids meta-level reasoning about the model and automatic generation of qualitative explanations of probabilistic updating. We leave out the formal definitions of the qualitative properties and theorems underlying qualitative belief propagation, focusing rather on the flavor of the method and the intuitions behind it. Section 2 reviews the main concepts of QPNs. Section 3 sketches briefly the belief propagation-based algorithm and discusses its advantages over the graph reduction approach. In Section 4, we show how our algorithm aids identification of sources of ambiguity in the network. Finally, in Section 5, we discuss possible applications of this work.

2 Qualitative Probabilistic Networks

Formally, a QPN is a pair $G = (V, Q)$, where V is a set of variables or nodes in the graph and Q is a set of qualitative relations among the variables [13]. All qualitative relations are expressed by signs '+, '−, '0, and '?, the last denoting ambiguity. There are two original types of qualitative relations in Q: qualitative influences and additive synergies. Their formal probabilistic definitions can be found in [13]. The *qualitative influences* define the sign of direct influence between two variables and correspond to arcs in a belief network. A positive qualitative influence between variables a and c, denoted by $S^+(a, c)$ expresses the fact that increasing the value of a, makes higher values of c more probable. If c is a binary variable, we define $C > \overline{C}$ (C means $c =$ **true**, $\overline{C}$ means $c =$ **false**. Negative qualitative influence, S^-, and zero qualitative influence, S^0, are defined analogously. The *additive synergy* is used with respect to two direct ancestors of a variable. A positive additive synergy, $Y^+(\{a, b\}, c)$, captures the property that the joint influence of a and b on c is greater than sum of their individual influences. Negative additive synergy, Y^-, and zero additive synergy, Y^0, are defined analogously.

We introduced the third qualitative property of QPNs, called *product synergy* in [6]. This was further studied in [15] and extended to support qualitative belief propagation in [5]. Product synergy captures the sign of conditional dependence between a pair of immediate predecessors of a node that has been observed or has indirect evidential support. The practical implication of product synergy is that under the specified circumstances, is forms a sufficient condition for a common pattern of reasoning known as *explaining away*. Explaining away is when given an observed effect and increase in probability of one cause, all other causes of that effect, that are negatively product synergistic to it, become less likely. Consider, for example, an automobile engine. Even though excessive oil consumption and oil leak through a cracked gasket can be assumed to be probabilistically independent, once we know that the oil level is low, this independence vanishes. Upon obtaining additional evidence for oil leak (e.g., observing greasy engine block), we find the likelihood of excessive oil consumption diminished — oil leak "explains away" excessive oil consumption. For a positive product synergy, the reverse is true, for example given low oil level, evidence for excessive oil consumption makes owner's negligence in replenishing oil more likely. Two direct predecessors of c, a and b exhibit a negative product synergy with respect to a value c_0 of c, denoted by $X^-(\{a, b\}, c_0)$, if given c_0 higher values of a make b less likely. Positive product synergy, X^+, and zero product synergy, X^0, are defined analogously.

If a qualitative property is not '+, '−, or '0, it is by default '? ($S^?$, $Y^?$, and $X^?$ respectively). As all the definitions of the qualitative properties are not-strict, both '+ and '− are consistent with '0; for the same reason '? is consistent with '0, '+, and '−. Any qualitative property that can be described by a '0 can be also described by '+, '−, or '?. Obviously, when specifying a network and doing any kind of reasoning, one prefers stronger conclusions to weaker ones and this is captured by the canonical order of signs: '0 is preferred to '+ and '−, and all three are preferred to '? [13]. These three qualitative properties can either be elicited directly from the expert along with the graphical belief network or derived from a numerical belief network, based on their formal definitions. It is worth noting that most popular probabilistic interactions exhibit unambiguous qualitative properties. It can be easily proven, for example, that bi-valued noisy-OR gates have always positive influences (S^+). For all pairs of their direct ancestors, they exhibit negative additive synergies (Y^-), negative product synergies (X^-) for the common effect present, and zero product synergies (X^0) for the common effect absent. Linear (Gaussian) models yield well defined qualitative influences (i.e., non–′?) and zero additive synergies (Y^0). Qualitative signs combine by means of sign multiplication ($\otimes$) and sign addition ($\oplus$) operators defined as follows:

$\otimes$	+	−	0	?
+	+	−	0	?
−	−	+	0	?
0	0	0	0	0
?	?	?	0	?

$\oplus$	+	−	0	?
+	+	?	+	?
−	?	−	−	?
0	+	−	0	?
?	?	?	?	?

Figure 1 shows an example of a QPN that captures the interaction of various variables related to low level of car engine oil. All variables in the example are

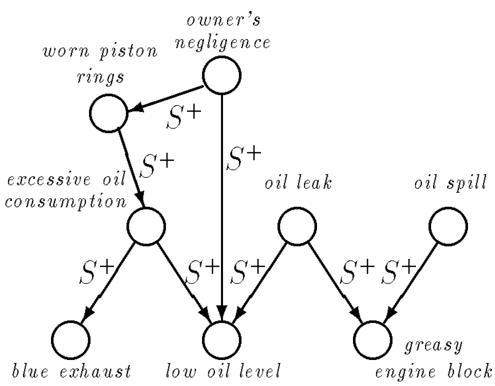

Figure 1: An example of a qualitative probabilistic network

propositional. *Owner's negligence* in replenishing oil leads to *low oil level* and in the long run to *worn piston rings* and *excessive oil consumption*. *Low oil level* can be also caused by an *oil leak* (e.g., through a cracked gasket). *Oil leak* and possible *oil spills* during adding or replacing oil can give the *engine block* a *greasy* look. Some of the variables in this network are directly observable; the others must be inferred. Links in a QPN are labeled by signs of the qualitative influences S^δ, each pair of links coming into a node is described by the signs of the synergies between them (not pictured). Note that all these relations are uncertain. *Excessive oil consumption* will usually lead to *blue exhaust*, but not always. But the fact that it makes *blue exhaust* more probable is denoted by a positive influence S^+. In the figures to follow, we will simply label the links with a ′+denoting S^+.

3 Qualitative Belief Propagation

Wellman [13] proposed an algorithm for reasoning in QPNs based on graph reduction, analogous to Shachter's [10] reduction algorithms for inference in quantitative belief networks. Given a query involving the influence of a node e on a node t, the algorithm applies repetitively arc-reversal and node-reduction operators until the graph is reduced to e and t connected by a single directed link from e to t. Since arc reversal loses qualitative information, the process can lead to ambiguity even where there is a definite qualitative relationship. Finding the optimal reduction sequence to minimize ambiguity has unknown computational complexity [14]. Further, the reduction of the original network structure makes it difficult to identify sources of ambiguity and to explain the inference process.

We proposed a computationally efficient algorithm for reasoning in QPNs that avoids these problems [4]. This algorithm, that we call *qualitative belief propagation*, is analogous to message-passing algorithms for quantitative belief networks (e.g., [8]) and traces the effect of an observation e on other network variables by propagating the sign of change from e through the entire network. This approach differs from the graph reduction-based algorithm in that it preserves the original structure of the network. Nothing is changed in the underlying graph, but every node on the path from e to t is given a label that characterizes the sign of impact. In this way, once the propagation is completed, one can easily read off the labeled graph how exactly the evidence impacts the target, i.e., what are the intermediate nodes through which e acts on t. The algorithm is formally described in [4]. Here, we restrict the exposition to the main ideas underlying the algorithm and give an example.

Belief propagation in singly connected networks has an intuitive meaning: the evidence flows from the observed variables outwards and never in the opposite direction. In the presence of multiple connections, this paradigm becomes problematic, as the evidence coming into a node can arrive from multiple directions. For any link that is part of a clique of nodes, it becomes impossible to point out in which direction the evidence flows. Numerical belief propagation through multiply connected graphs encounters the problem of a possible infinite sequence of local belief propagation and an unstable equilibrium that does not necessarily correspond to the new probabilistic state of the network [9, pages 195–223].

It turns out, that the qualitative properties of the QPNs allow for an interesting view of belief propagation. Qualitative influences and synergies are independent of any other nodes. This allows, for any two variables e and t, to disregard all such nodes and effectively decompose the flow of evidence from e to t into distinct trails (chains of links in the underlying graph) from e to t. On each of these trails, belief flows in only one direction, from e to t, and never in the opposite direction, exactly as it does in singly connected networks. Presence of different, parallel trails does not change any properties of a single trail and these can be determined by considering each trail in separation. Qualitative change in belief in a node t given a single evidence node e can be viewed as a sum of changes through individual evidential trails from e to t. It will be well determined only if the signs of these paths are consistent (i.e., the sign sum is not $'?$).

The algorithm for qualitative belief propagation is based on local message passing. The goal is to determine a sign for each node denoting the direction of change in belief for that node given new evidence for an observed node. Initially each node is set to $'0$, except the observed node which is set to the specified sign. A message is sent to each neighbor. The sign of each message becomes the sign product of its previous sign and the sign of the link it traverses. Each message keeps a list of the nodes it has visited and its origin, so it can avoid visiting any node more than once. Effectively, each message travels on one possible evidential trail. Each node, on receiving a message, updates its own sign with the sign sum of itself and the sign of the message. Then it passes a copy of the message to all unvisited neighbors that need to update their signs.

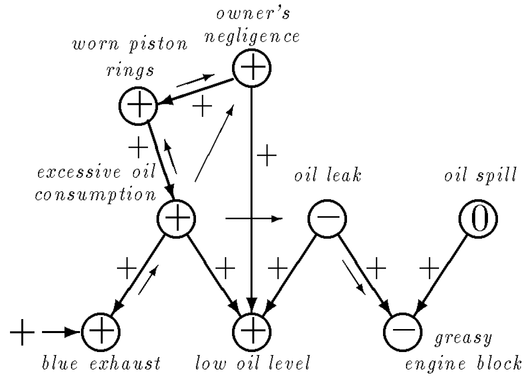

Figure 2: Algorithm for qualitative belief propagation: An example.

Figure 2 shows an example of how the algorithm works in practice. Suppose that we have observed *low oil level* and we want to know the effect of observing *blue exhaust* on other variables in the model. We set the signs of each of the nodes to ′0 and start by sending a positive sign to *blue exhaust*, which is our evidence node. *Blue exhaust* determines that its parent, node *excessive oil consumption*, needs updating, as the sign product of ′+ and the sign of the link ′+ is ′+ and is different from the current value at the node ′0. After receiving this message, *excessive oil consumption* sends a positive message to *worn piston rings*. Given that the node *low oil level* has been observed, *excessive oil consumption* will also send two intercausal messages: a positive message to *owner's negligence* and a negative message to *oil leak* (the signs are determined by the positive sign of *excessive oil consumption* and respective positive and negative signs of product synergies). No intercausal messages are passed between *owner's negligence* and *oil leak*, as the sign of the product synergy between the two given *low oil level* is zero. *Worn piston rings* will not send any further messages, as its only parent, *owner negligent in replenishing oil*, does not require any change from the already positive sign. The negative sign of *oil leak* will propagate to *greasy engine block*, that will not send any further messages and the algorithm will terminate. The final sign in each node expresses how the probability of this node is impacted by observing *blue exhaust*.

The character of the sign addition implies that each node can change its sign at most twice — first from ′0 to ′+, ′−, or ′? and then, if at all, only to ′?, which can never change to any other sign. Hence each node receives a request for updating its sign at most twice, and the total number of messages for the network to reach stability is less than twice the number of nodes. Each message carries a list of visited nodes, which contains at most the total number of nodes in the graph. Hence, the algorithm is quadratic in the size of the network. Unfortunately, this propagation algorithm does not generalize straightforwardly to quantitative belief networks, where belief updating is known to be NP-hard [2].

4 Identification of Sources of Ambiguity

Reasoning purely with signs can lead to ambiguities that cannot be resolved at this level of specificity. Generally, if the sign of change of a node through which propagation is conducted becomes ambiguous, the signs of all nodes that are located beyond that node will be ambiguous. A formal representation should not use more specificity than needed to support the reasoning and decision making required of it. A simple initial representation can provide insight into the problem and lead the system to a refinement of those parts only that are critical for answering the query, at each time doing no more work than necessary. As the time spent on analyzing a simple qualitative representation is usually negligible, the system will effectively allocate all of its efforts to the most important parts of the problem. To support this type of approach, the program needs to reflect on the model in order to locate the direct source of ambiguity and find the most fruitful direction of elaboration of the model. As qualitative belief propagation preserves the structure of the network, it directly supports this type of reasoning.

In a singly connected network, ambiguity can enter only through the ambiguity in a single influence or synergy ($S^?$ or $X^?$). In multiply connected networks, ambiguity can also be a result of conflicting influences coming to the target node from different trails (Figure 3–(a)). Some configurations of multiply connected networks always lead to conflicting trails. This is the case with what we call *am-*

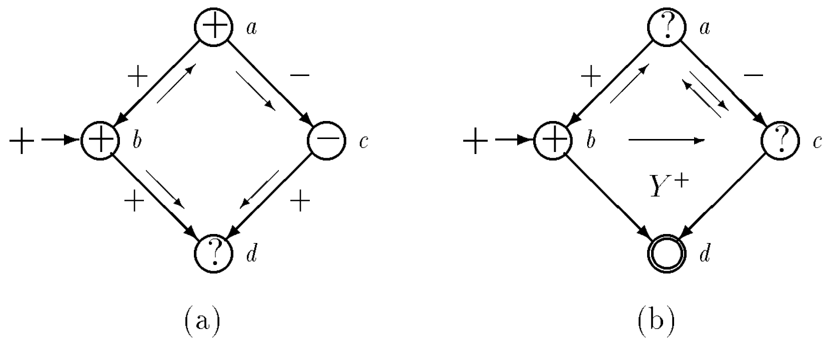

Figure 3: (a) Ambiguity resulting from conflicting trails: node d receives a positive message directly from b and a negative message from b through a and c. (b) Ambiguous loops: the product of links $b-a$, $a-c$, and the intercausal link $b-c$ (with d observed) is negative.

biguous loops (Figure 3–(a)). A loop is an active trail that starts and ends at the same node. If the sign product of all links in a loop (including intercausal links) is negative, sign propagation through such loop is guaranteed to lead to ambiguity ('? sign). It is possible to prove that at most one node of an ambiguous loop will have an unambiguous sign. We can prove it by contradiction. For a given source node, there may be at most one node in any loop that has no trails in that loop — if and only if all evidential trails to any node of the loop include that node. The sign of such a node will be independent on the signs of any nodes and links in the loop. For any other node in the loop, there will be at least two evidential trails, one going clockwise and one going counterclockwise. Suppose that a node n in the loop has an unambiguous sign. Without loss of generality, let us assume that it is '+. Now, there exist at least two distinct trails from n to any node m in the loop. Their product has to be negative because the loop is ambiguous (has a negative product) and the only way of getting a negative product is when one of the two trails is negative and one is positive. Their signs will be thus different by the sole virtue of the ambiguity of the loop. It follows that the sign of node n has to be $'?$, which contradicts the assumption. Note that if evidence enters the loop through more than one node, then all nodes in the loop will be $'?$, because the same argument will hold for each of these nodes. There will be always at least two evidential trails through which conflicting evidence comes to them. While most qualitative properties in the networks that we studied were well defined, it was not uncommon to observe ambiguities introduced by ambiguous loops in multiply connected networks with instantiated nodes. One reason for that is that the most common value of product synergy appears to be negative, which in loops often leads to conflicts with usually positive signs of links. Ambiguous loops need to be analyzed and elaborated upon as a whole.

Summarizing, we identified four elementary possible sources of ambiguity: ambiguous sign of a qualitative influence, ambiguous sign of an intercausal influence, conflicting influence signs from multiple paths, and ambiguity from propagation through an inherently ambiguous loop. These four sources of ambiguity can be easily identified in qualitative belief propagation and addressed by a higher level of specification (such as order of magnitude or full numeric specification) of the ambiguous part of the network.

5 Conclusions and Applications

This paper has summarized our work on qualitative belief propagation, a computationally efficient scheme for reasoning in qualitative probabilistic networks. Belief propagation is more powerful than graph reduction approach for two reasons: (1) it uses product synergy, which is a new qualitative property of probabilistic interactions, and (2) it offers a reasoning scheme, whose operators do not lead to loss of qualitative information and whose final results do not depend on the order of their application. Although examples of problems that can be resolved by belief propagation and not by graph reduction can be easily found, it is unfair to compare the strength of the two methods, as belief propagation uses an additional qualitative property, namely product synergy.

Wellman [12] describes several possible applications of QPNs, such as support for heuristic planning and identification of dominant decisions in a decision problem. Qualitative belief propagation supports these applications, and has the additional advantage over the graph-reduction approach in that it preserves the underlying graph and determines the sign of the node of interest along with the signs of all intermediate nodes. This supports directly two new applications of QPNs: meta-level reasoning about the model and automatic generation of qualitative verbal explanations of reasoning.

In case of sign-ambiguity a computer program can self-reflect about the model at a meta level and find the reason for ambiguity. Hence, the program can determine ways in which the least additional specificity could resolve the ambiguity. Automatic detection of possible ambiguities can be an aid in knowledge engineering. If a QPN specification of a domain is not sufficient to determine a sign of change, parts of the model can be identified where additional qualitative or quantitative information is needed.

Belief propagation appears to be easy to follow for people and it can be used for generation of qualitative verbal explanations of probabilistic reasoning. The individual signs along with the signs of influences can be translated into natural language sentences describing paths of change from the evidence to the variable in question (see Figure 4 for an example). Explanation of each step involves

```
Qualitative influence of greasy engine block on excessive oil consumption:
Greasy engine block is evidence for oil leak.
Oil leak explains low oil level, hence is evidence against excessive oil
consumption.
Therefore, greasy engine block is evidence against excessive oil
consumption.
```

Figure 4: Qualitative explanations: An example.

reference to a usually familiar causal or diagnostic interaction of variables. In general, explanations based on qualitative reasoning are easier to understand than explanations using numerical probabilities. So even where a quantified belief network is available, it may often be clearer to reduce it to the qualitative form, and base explanations on purely qualitative reasoning. We think that belief propagation-based scheme supports this in a straightforward and efficient manner by determining the signs of all nodes in the network during one execution of the algorithm. More details on generation of verbal explanations of reasoning based on belief propagation can be found in [3].

Another possible application is using qualitative reasoning about parts of the model during model building. Robustness of qualitative data provides an additional constraint on the elicited numerical probabilities and allows for a sound consistency check of the elicited numbers and the structure of the model. If the numbers elicited with respect to the model result in counterintuitive qualitative properties, we have a good reason to check the numbers. Also, if qualitative explanations of reasoning based on an early version of a model seem to be counterintuitive to the expert, it may be a good indication of possible problems in the model structure.

References

[1] John S. Breese and Kenneth W. Fertig. Decision making with interval influence diagrams. In P.P. Bonissone, M. Henrion, L.N. Kanal, and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence 6*, pages 467–478. Elsevier Science Publishers B.V. (North Holland), 1991.

[2] Gregory F. Cooper. The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence*, 42(2–3):393–405, March 1990.

[3] Marek J. Druzdzel. *Probabilistic Reasoning in Decision Support Systems: From Computation to Common Sense*. PhD thesis, Department of Engineer-

ing and Public Policy, Carnegie Mellon University, Pittsburgh, PA, January 1993.

[4] Marek J. Druzdzel and Max Henrion. Efficient reasoning in qualitative probabilistic networks. In *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI–93)*, Washington, D.C., July 1993.

[5] Marek J. Druzdzel and Max Henrion. Intercausal reasoning with uninstantiated ancestor nodes. In *Proceedings of the Ninth Annual Conference on Uncertainty in Artificial Intelligence (UAI–93)*, Washington, D.C., July 1993.

[6] Max Henrion and Marek J. Druzdzel. Qualitative propagation and scenario-based approaches to explanation of probabilistic reasoning. In P.P. Bonissone, M. Henrion, L.N. Kanal, and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence 6*, pages 17–32. Elsevier Science Publishers B.V. (North Holland), 1991.

[7] Ronald A. Howard. Knowledge maps. *Management Science*, 35(8):903–922, August 1989.

[8] Jin H. Kim and Judea Pearl. A computational model for causal and diagnostic reasoning in inference systems. In *Proceedings of the 8th International Joint Conference on Artificial Intelligence, IJCAI–83*, pages 190–193, Los Angeles, CA, 1983.

[9] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann Publishers, Inc., San Mateo, CA, 1988.

[10] Ross D. Shachter. Evaluating influence diagrams. *Operations Research*, 34(6):871–882, November–December 1986.

[11] Linda van der Gaag. Computing probability intervals under independency constraints. In P.P. Bonissone, M. Henrion, L.N. Kanal, and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence 6*, pages 457–466. Elsevier Science Publishers B.V. (North Holland), 1991.

[12] Michael P. Wellman. *Formulation of Tradeoffs in Planning Under Uncertainty.* Pitman & Morgan Kaufmann, London, 1990.

[13] Michael P. Wellman. Fundamental concepts of qualitative probabilistic networks. *Artificial Intelligence*, 44(3):257–303, August 1990.

[14] Michael P. Wellman. Graphical inference in qualitative probabilistic networks. *Networks*, 20(5):687–701, August 1990.

[15] Michael P. Wellman and Max Henrion. Explaining "explaining away". *IEEE Trans. PAMI*, 15(3):287–291, March 1993.