Axiomatizing Noisy-OR

Fabio G. Cozman*

May 25, 2004

Abstract

The Noisy-OR function is extensively used in probabilistic reasoning, and usually justified with heuristic arguments. This paper investigates sets of conditions that imply the Noisy-OR function.

1 INTRODUCTION

This paper examines the foundations of a popular pattern of probabilistic reasoning, the *Noisy-OR* combination function. This function has been used extensively in artificial intelligence, both in purely probabilistic models and in models that combine probability and logic.

When building a probabilistic model, one must often deal with a variable X that depends directly on several variables $Y_1, \ldots, Y_n$. The Bayesian network fragment in Figure 1 shows a possible situation, where variables Y_i are parents of X. In this case call X a *collider* [16].

To specify a Bayesian network, each collider must be associated with a probability distribution $p(X|Y_1, \ldots, Y_n)$. Assume all variables are binary with values T (for "true") and F (for "false"). Then the complete specification of $p(X|Y_1, \ldots, Y_n)$ requires 2^n probability values. An attractive strategy is to find methods that specify $p(X|Y_1, \ldots, Y_n)$ using fewer parameters.

The Noisy-OR function is a compact representation for the distribution of colliders. The idea is to start with n probability values p_i, where p_i is the probability that $\{X = T\}$ conditional on $\{Y_i = T\}$ and $\{Y_j = F\}$ for $j \neq i$. That is,

$$p_i = p\big(X = T|Y_i = T, \{Y_j = F\}_{j=1, j\neq i}^{n}\big) \, .$$

The probabilities p_i are called *link probabilities*. Suppose each variable Y_i is examined and, if it is T, then there is chance $(1 - p_i)$ that it is flipped to F; if Y_i is F, then is stays with F. Denote

*Escola Politécnica, Univ. de São Paulo, São Paulo, Brazil. Email: fgcozman@usp.br

by Y_i' the result of flipping (or not) Y_i. Finally, suppose that

$$X = (Y_1' \vee Y_2' \vee \cdots \vee Y_n');$$

that is, X is the OR combination of the Y_i'. A Noisy-OR function is thus a disjunction of "noisy" versions of Y_i. The distribution of X conditional on $Y_1, \ldots, Y_n$ is

$$p(X = T|Y_1, \ldots, Y_n) = 1 - \prod_{i:Y_i=T} (1 - p_i). \tag{1}$$

Given its history of good service, the Noisy-OR has been the object of intense investigation in the literature. However it does not seem that the following question has been asked so far: Is there a simple set of conditions on colliders that forces the Noisy-OR function to be adopted? A solid foundation for the Noisy-OR function is currently an important issue, as this function is a critical component of models that merge logical rules and probabilistic data. A justification of the Noisy-OR function based on a set of axioms should be useful in pointing out exactly what is assumed by the model, and suggesting alternative models through modification/removal of axioms. An axiomatic characterization of the Noisy-OR is the purpose of this paper.

Section 2 discusses some of the motivating factors that have led to widespread adoption of the Noisy-OR function, focusing on the "causal" and "rule-based" applications of the function. Sections 3, 4 and 5 present two possible axiomatizations for the Noisy-OR function. Section 6 concludes the paper.

2 ARGUMENTS FOR NOISY-OR

The Noisy-OR combination function is quite intuitive, as it essentially postulates an OR combination among slightly corrupted versions of the variables $Y_1, \ldots, Y_n$. This rationale is quite appealing in at least two situations.

- Suppose each Y_i is interpreted as a "cause" of X. Each active Y_i is then causing X with probability p_i. Causes are active or not independently of each other, and X happens when one or more of its causes are active.
- Suppose each edge $X \leftarrow Y_i$ is interpreted as an implication, much like a Prolog rule. As in Prolog, a sequence of rules $X \leftarrow Y_i$ with identical consequent X may then mean the

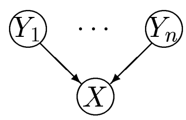

Figure 1: A collider X with parents $Y_1, \ldots, Y_n$. For the purposes of this paper, nodes and variables are equivalent.

disjunction of the antecedents. Each p_i may be viewed as the probability that a rule is actually fired, independently of other rules.

The Noisy-OR function was proposed by Pearl (apparently at the same time as similar proposals appeared in other fields) [14]. The motivation for the proposal was to reduce the elicitation effort involved in building a Bayesian network. This argument for Noisy-OR was detailed by Henrion [10], and a new ingredient was added: the *leak* probability p_L that X is T even when all Y_i are F. Other extensions of the Noisy-OR model ensued, for example Noisy-AND, Noisy-MAX [6, 17].

A particularly relevant characteristic of the Noisy-OR function is its *explaining away* property. Assuming variables $Y_1, \ldots, Y_n$ are (unconditionally) independent of each other, explaining away occurs when, for any distinct Y_i and Y_j,

$$p(Y_i = T | X = T, Y_j = T, \mathbf{Y}_{-ij}) < p(Y_i = T | X = T, \mathbf{Y}_{-ij}) \,, \tag{2}$$

where $\mathbf{Y}_{-ij}$ indicates an arbitrary instantiation of all variables $\{Y_k\}_{k=1,k\neq i,k\neq j}^n$.

The idea behind property (2) is simple: upon observing $\{Y_j = T\}$, belief in $\{Y_i = T\}$ decreases as an explanation for the observed event $\{X = T\}$, regardless of any configuration of the parents of X.

Wellman and Henrion define the explaining away property in a slightly different form [21], requiring that

$$p(Y_i = T | X = T, Y_j = T, \mathbf{Y}_{-ij}) \leq p(Y_i = T | X = T, \mathbf{Y}_{-ij}) \,. \tag{3}$$

However, the use of equality in the definion of the explaining away property is perhaps not so advisable, as it allows for situations where the observation of Y_j (after observation of $\{X = T\}$) simply does not change the belief in Y_i.

The Noisy-OR function satisfies Expression (2) for any value of the link probabilities in the open interval $(0, 1)$; if the more liberal property (3) is adopted, then the Noisy-OR function leads to explaining away for any value of the link probabilities. Explaining away is one of several qualitative patterns of probabilistic reasoning that can be referred to as *synergy* patterns among parents of a variable [2, 20, 21]. We note that Lucas has also put forward a thorough analysis of the interaction between qualitative patterns and several "noisy" combination functions [12].

The general properties of the Noisy-OR function and its generalizations were captured by Heckerman and Breese in their definition of *causal independence* [9]. Here the causal interpretation is obviously present. The idea is that each Y_i is a cause associated with an inhibitory variable I_i. The variable I_i is binary with $p(I_i = T) = p_i$. The actual effect of Y_i is given by the conjunction $Y_i \wedge I_i$. There is causal independence in the sense that the variables I_i are (unconditionally) independent of each other and of all variables $Y_1, \ldots, Y_n$. Finally, a deterministic function g takes $\{Y_i \wedge I_i\}_{i=1}^n$ as input and produces X. Figure 2 depicts the structure of the causal independence model.

Zhang and Poole have also investigated the structure in Figure 2, additionally assuming that g is itself a combination of two-place functions [22]. The Noisy-OR function and its most popular generalizations all satisfy this additional assumption. In fact, one of the important properties of the Noisy-OR function is that it greatly simplifies inference algorithms due to its factorization in smaller terms.

A slightly different stream of research has focused on the use of Noisy-OR functions for combination of logical/probabilistic rules. The underlying theme is to combine probabilistic rules such as

$$X(u) \xleftarrow{p} Y(u,v),$$

where X and Y are now relations among objects of some domain, and p is interpreted as the probability that the rule "fires" [8, 11, 13, 15]. The difficulty here is that $X(u)$ (for a given u) may be associated with several rules, one for each instantiation of v. Thus one must decide how to combine the available probabilistic information on $X(u)$, $Y(u, v_1), \ldots, Y(u, v_n)$. The most popular combination function is exactly the Noisy-OR [8, 11, 13]. Consequently, the Noisy-OR function has become a central element in the interface between logical and probabilistic reasoning.

While several arguments for the Noisy-OR function are based on causal models, many applications of this function have no causal basis (particularly in rule-based domains). The next sections discuss the justification of the Noisy-OR function from a neutral perspective.

3 PROPERTIES OF COMBINATION FUNCTIONS

Assume that $Y_1, \ldots, Y_n$ are binary variables that are (unconditionally) independent of each other, and X is a binary variable that directly depends on $Y_1, \ldots, Y_n$. All binary variables have values T and F.

In his presentation of the Noisy-OR function, Pearl argues that the following property is desirable for any combination function [14]:

Accountability: The probability of $\{X = T\}$ must be zero if all Y_i are set to F.

Another property advocated by Pearl is *exception independence*. The idea is that any Y_i may be inhibited with probability p_i under certain conditions:

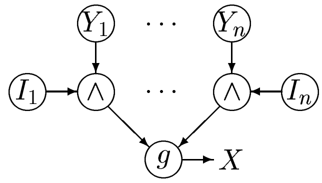

Figure 2: Inhibitory variables and the combination function g.

Exception independence: The value of X is affected by each Y_i only through $(Y_i \wedge I_i)$, where the $I_1, \ldots, I_n$ are inhibitory variables that are (unconditionally) independent of each other and of the variables $Y_1, \ldots, Y_n$.

This is exactly the causal-independence property discussed in Section 2 [9].

Exception independence implies the structure in Figure 2, where g is a function of the $\{Y_i \wedge I_i\}_{i=1}^n$. Now, g may be any function, including a probabilistic one. Thus there is not yet a definite, unique form for $p(X|Y_1, \ldots, Y_n)$.

An additional, and rather substantive, condition that is present in the literature is to take g as a *deterministic* function. This is certainly a strong condition that cannot be defended from first principles; it is probably better to take it as a property that effectively *defines* the meaning of combination functions. Hence:

Determinism: The value of X is produced by a deterministic function $g(Y_1 \wedge I_1, \ldots, Y_n \wedge I_n)$.

These properties make no claims that causal effects are at work. Even though they certainly make sense in a causal model, the properties are supposed to be neutral with respect to causality.

An eminently reasonable requirement on g is that its application be "associative" and "commutative." This requirement has been proposed by Zhang and Poole as part of the *definition* of causal independence [22], and is satisfied in Noisy-OR, Noisy-AND, Noisy-MAX, and other "Noisy-XX" models in the literature. Lucas has also examined this requirement in detail [12].

Associativity: For any $\{Y_1, \ldots, Y_n\}$, $g(Y_1, \ldots, Y_n) = g(\ldots g(g(Y_i, Y_j), Y_k) \ldots, Y_l)$ (for distinct i,j,k,l), and the value of g does not change for any permutation of its inputs.

Associativity and determinism restrict considerably the possible functions g, but there are still several possible functions. The following theorem clarifies this issue; here and in the remainder of the paper, $A \oplus B$ denotes the XOR operation for A and B.

Theorem 1 *Accountability, exception independence, determinism, and associativity, are satisfied only by four two-place functions* $g(A, B)$*:* $\{F, A \wedge B, A \oplus B, A \vee B\}$.

Proof. The effect of the associativity and determinism is to restrict our attention to two-place boolean functions $g(A, B)$. There are sixteen boolean functions $g(A, B)$, but not all of them are valid. Functions with $g(F, F) = 1$ must be discarded, because these functions lead to $p(X|A = F, B = F) = 1$ and violate accountability. Four other functions are not commutative; the remaining four functions are listed in the theorem. QED

We note that Theorem 1 has essentially been proved by Lucas [12].

Exception independence, associativity and determinism are "structural" properties that define the nature of combination functions. To proceed, further conditions must be adopted — conditions that capture the intended meaning for combination functions. The next two sections explore this path.

4 EXPLAINING AWAY AND REVERSE INDEPENDENCE

A possible condition to impose on any combination function is that the function satisfies the explaining away property (2). That is, multiple parents $Y_1, \ldots, Y_n$ should "compete" for X. Consider the following requirement, inspired by Wellman and Henrion analysis of qualitative reasoning [21]:

Explaining Away: The property (2) is satisfied as long as $Y_1, \ldots, Y_n$ are (unconditionally) independence of one another, when link probabilities and prior probabilities for the Y_i are in the open interval (0,1).

The Noisy-OR function satisfies accountability, exception independence, determinism, associativity and explaining away. Do these five properties *imply* the Noisy-OR function? That is, do they work as an *axiomatization* of the Noisy-OR function? The answer to this question requires the treatment of conditioning events with zero probability — as these events appear in the explaining away pattern for several possible g functions. The discussion of zero probability events is deferred to the Appendix; the main conclusion is:

Theorem 2 *Accountability, exception independence, associativity, determinism and explaining away are satisfied only by two two-place functions* $g(A, B)$*:* $\{A \oplus B, A \vee B\}$.

It should be noted that Theorem 2 is only obtained for the strict definition of explaining away that uses Expression (2). If the less strict Expression (3) is used, then all four functions mentioned in Theorem 1 satisfy the conditions of Theorem 2.

Thus the explaining away property does not single out the Noisy-OR function. So, instead of the explaining away property, consider the following property, first explored by Agosta [1]. Suppose X is produced by a combination function from (unconditionally) independent parents $Y_1, \ldots, Y_n$. Given the event $\{X = F\}$, shouldn't the parents $Y_1, \ldots, Y_n$ still be independent of one another? Because X was just obseved to be "inactive," one might argue that independence relations should be preserved.[1] The property of interest is:

Reverse independence If X is produced by a combination of (unconditionally) independence parents $Y_1, \ldots, Y_n$, then the parents are independent conditional on $\{X = F\}$.

[1] Note that such conditional independence relations are not implied by d-separation, as d-separation would not sanction any independence among parents conditional on $\{X = F\}$.

$g(A,B)$	$p(T\|F,F)$	$p(T\|F,T)$	$p(T\|T,F)$	$p(T\|T,T)$
F	0	0	0	0
$A \wedge B$	0	0	0	pq
$A \oplus B$	0	q	p	$p+q-2pq$
$A \vee B$	0	q	p	$p+q-pq$

Table 1: The probabilities $p(X = T|Y_1, Y_2)$ for the four functions in Theorem 1.

Agosta denotes reverse independence by CICI (for *conditional inter-causal independence*) [1], but perhaps we should not focus only on causal interpretations for such a property.

It turns out that reverse independence alone is not enough to imply the Noisy-OR function:

Theorem 3 *Accountability, exception independence, associativity, determinism and reverse independence are satisfied only by two two-place functions* $g(A, B)$*:* $\{F, A \vee B\}$.

The proof of this theorem is discussed in the Appendix.

Explaining away and reverse independence do not uniquely imply the Noisy-OR function when adopted separately. However, we uniquely obtain the Noisy-OR function by adopting both properties:

Theorem 4 *Accountability, exception independence, associativity, determinism, explaining away and reverse independence are only satisfied by the Noisy-OR function.*

Thus we have identified one possible path for axiomatizing the Noisy-OR function, in the form of the six conditions in Theorem 4 (a similar conclusion can be obtained using results by Lucas [12]). One may wonder whether an alternative axiomatization is possible with less conditions; one such axiomatization is presented in the next section.

5 CUMULATIVITY

Consider then a collider X with two parents Y_1 and Y_2. Table 1 shows the probability values for $p(X = T|Y_1, Y_2)$ for the four possible functions g indicated in Theorem 1. The link probabilities are $p = p(X = T|Y_1 = T, Y_2 = F)$ and $q = p(X = T|Y_1 = F, Y_2 = T)$.

Among all functions in Table 1, the Noisy-XOR function has a somewhat unpleasant feature: the probability of $\{X = T\}$ conditional on $\{Y_1 = T, Y_2 = T\}$ may be *smaller* than the probability of $\{X = T\}$ conditional on either $\{Y_1 = T, Y_2 = F\}$ or $\{Y_1 = F, Y_2 = T\}$. It seems reasonable to expect that a combination function should guarantee that the more inputs are active, the higher is the probability of the collider X.

Consider then the following property:

Cumulativity: If two configurations $\mathbf{Y}_1$ and $\mathbf{Y}_2$ of parents $Y_1, \ldots, Y_n$ are identical, except that some variables are set to T in $\mathbf{Y}_1$ and to F in $\mathbf{Y}_2$, then $p(X = T|\mathbf{Y}_1) > p(X = T|\mathbf{Y}_2)$ for link probabilities in the open interval $(0, 1)$.

While the explaining away and reverse independence properties are found in the literature and reflect standard facts about the Noisy-OR, cumulativity is a new (albeit straightforward) assumption on combination functions.

Note that, once cumulativity is assumed, accountability loses most of its appeal. In fact, accountability is not even necessary in the presence of cumulativity:

Theorem 5 *Exception independence, associativity, determinism and cumulativity are only satisfied by Noisy-OR functions.*

Proof. By direct verification of Table 1, the Noisy-OR is the only possible function for the case of two parents. Expression (1) then shows that the Noisy-OR function satisfies the condition for any number of parents. Functions that produce T for all inputs equal to F violate cumulativity and need not be considered. QED

6 CONCLUSION

The Noisy-OR function has been one of the most effective tools for elicitation of probabilistic models, particularly models that have a causal basis. This function has recently been adopted as a central element in approaches that merge logical rules and probabilistic information. As applications of Noisy-OR move away from purely causal models, it is important it is to have a solid foundation for this function. The purpose of this paper is to present justifications of the Noisy-OR function that are based on sets of simple yet appealing properties.

To summarize, there are two sets of properties that imply the Noisy-OR function:

- Accountability, exception independence, associativity, determinism, explaining away and reverse independence.
- Exception independence, associativity, determinism and cumulativity.

The first set of properties contains conditions that have been long associated with the Noisy-OR function, usually in connection to causal models. The second set is more compact, and is perhaps appropriate as a common foundation for both “causal” and “rule-based” applications of the Noisy-OR function. These sets of properties can now be used to investigate combination functions that go beyond the Noisy-OR, either by dropping some properties, or by modifying others.

A PROOFS FOR SECTION 4

This section discusses the proof of Theorems 2, 3 and 4. A few arguments depend on decisions regarding conditioning on events of zero probability.

Start with the simplest situation: take the structure in Figure 1 and assume that only two parents, Y_1 and Y_2, are present. Table 1 contains the probabilities $p(X|Y_1, Y_2)$ for the four functions in Theorem 1, where the link probabilities are $p = p(X = T|Y_1 = T, Y_2 = F)$ and $q = p(X = T|Y_1 = F, Y_2 = T)$.

To verify the explaining away condition, it is necessary to compute $p(Y_1 = T|X = T, Y_2 = T)$ and $p(Y_1 = T|X = T)$; explaining away occurs when the former is strictly smaller than the latter. Denote by α the probability of the event $\{Y_1 = T\}$ and by β the probability of the event $\{Y_2 = T\}$. Recall that Y_1 and Y_2 are assumed (unconditionally) independent by the explaining away condition.

For $g(A, B) = A \vee B$, we obtain:

$$p(Y_1 = T|X = T, Y_2 = T) = \frac{(q - pq + p)\alpha}{(q - pq + p)\alpha + q(1 - \alpha)},$$

$$p(Y_1 = T|X = T) = \frac{\alpha((q - pq + p)\beta + p(1 - \beta))}{\alpha((q - pq + p)\beta + p(1 - \beta)) + (1 - \alpha)q}.$$

For $g(A, B) = A \oplus B$, we obtain:

$$p(Y_1 = T|X = T, Y_2 = T) = \frac{(q - 2pq + p)\alpha}{(q - 2pq + p)\alpha + q(1 - \alpha)},$$

$$p(Y_1 = T|X = T) = \frac{\alpha((q - 2pq + p)\beta + p(1 - \beta))}{\alpha((q - 2pq + p)\beta + p(1 - \beta)) + (1 - \alpha)q}.$$

Thus, both Noisy-OR and Noisy-XOR satisfy the explaining away condition for two parents (as long as the relevant probabilities are different from 0 and 1; this condition is assumed in the remainder of this section). The same process of direct verification shows that $g(A, B) = A \wedge B$ fails the explaining away property, as $p(Y_1 = T|X = T, Y_2 = T) = p(Y_1 = T|X = T) = 1$ for this function.

The remaining situation is the function $g(A, B) = F$. In this case $p(A = T|X = T)$ must be defined for the zero probability event $\{X = T\}$. Note also that a few other combination functions generate conditioning on zero probabilities; for example, the function $g(A, B) = A \wedge B$ leads to zero probability conditioning for $p(Y_1 = T|X = T, Y_2 = F)$. Conditioning on zero probabilities is a delicate situation that can be handled by several methods [3, 4, 5, 7, 18, 19]. Here the difficulty is that $\{X = T, Y_2 = F\}$ may be a logical impossibility (consider the situation where $g(A, B) = F$; then $\{X = T\}$ is logically impossible). One solution is to remove any function that can lead to such inconveniencies. A possibly more elegant approach is as follows. Suppose we have a leak probability ϵ that X will be T, independent of any other event, and then we

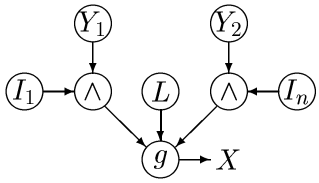

Figure 3: Leak variable L with probability ϵ for two parents Y_1 and Y_2.

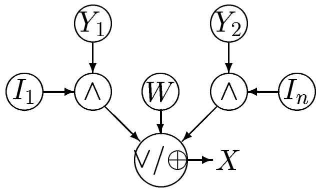

Figure 4: Variable X with several parents: all parents $X_3, \ldots, X_n$ are lumped into W.

examine the behavior of probabilities as $\epsilon \to 0$. Thus we consider the structure in Figure 3. Here no conditioning with zero probabilities occurs. For $g(A, B) = F$, we obtain:

$$p(Y_1 = T|X = T, Y_2 = T) = p(Y_1 = T|X = T) = \alpha,$$

and we thus remove this function from consideration. The probabilities $p(Y_1 = T|X = T, Y_2 = T)$ and $p(Y_1 = T|X = T)$ can be computed for all four functions in Theorem 1 with leak probability ϵ; the conclusion is the same (after tedious algebraic manipulations): for two parents, only Noisy-OR and Noisy-XOR satisfy the conditions in Theorem 2.

It remains to be shown that the explaining away condition is satisfied by Noisy-OR and Noisy-XOR functions with more than two parents. To show this, we can lump every formula other than X and Y into a variable W that is active with probability $p(W)$. The resulting structure is presented in Figure 4.

Now for the Noisy-OR function we have

$$p(X|Y_1, Y_2) = \sum_W p(X|Y_1, Y_2)(1 - p(W)) + p(W),$$

where $p(X|Y_1, Y_2)$ is the Noisy-OR combination of Y_1 and Y_2. It is possible possible to show, after tedious algebraic manipulations, that explaining away occurs for any $p(W)$ with non-extreme values. Likewise, for the Noisy-XOR function we have

$$p(X|Y_1, Y_2) = \sum_W p(X|Y_1, Y_2)(1 - 2p(W)) + p(W),$$

where $p(X|Y_1, Y_2)$ is not the Noisy-XOR combination of Y_1 and Y_2. Explaining away again occurs for any $p(W)$ with non-extreme values. These results are still valid even if we add a leak probability ϵ, effectively merging the structures in Figures 3 and 4. We thus obtain Theorem 2.

$g(A,B)$	$p(Y_1 = T \vert X = F, Y_2 = T)$	$p(Y_1 = T \vert X = F)$
F	α	α
$A \wedge B$	$\alpha(1-pq)/(1-\alpha pq)$	$\alpha(1-\beta pq)/(1-\alpha\beta pq)$
$A \oplus B$	$\frac{\alpha(1-p-q+2pq)}{\alpha(1-\alpha p-q+2\alpha pq)}$	$\frac{\alpha(1-p-\beta q+2\beta pq)}{\alpha(1-\alpha p-\beta q+2\alpha\beta pq)}$
$A \vee B$	$\alpha(1-p)/(1-\alpha p)$	$\alpha(1-p)/(1-\alpha p)$

Table 2: The probabilities $p(Y_1 = T|X = F, Y_2 = T)$ and $p(Y_1 = T|X = F)$ for the four functions in Theorem 1.

As for Theorem 3, compute $p(Y_1 = T|X = F, Y_2 = T)$ and $p(Y_1 = T|X = F)$ from Table 1 for the four functions in Theorem 1. There is no conditioning on zero probability events, and the resulting probabilities are given in Table 2. Direct verification leads to Theorem 3 for structures with two parents; more parents can be handled using the same structure depicted in Figure 4.

Finally, Theorems 2 and 3 directly imply Theorem 4.

Acknowledgements

This work has received generous support from HP Labs and HP Brazil. The work has also been supported by CNPq (through grant 3000183/98-4).

References

[1] J. M. Agosta. "Conditional inter-causally independent" node distributions, a property of "noisy-or" models. In *Proc. Conf. on Uncertainty in Artificial Intelligence*, pp. 9–16, San Francisco, 1991. Morgan Kaufmann.

[2] J. H. Bolt, S. Renooij, and L. C. van der Gaag. Upgrading ambiguous signs in QPNs. In *Conf. on Uncertainty in Artificial Intelligence*, pp. 73–80, San Francisco, California, 2003. Morgan Kaufmann.

[3] G. Coletti. Coherent numerical and ordinal probabilistic assessments. *IEEE Transactions on Systems, Man and Cybernetics*, 24(12):1747–1753, 1994.

[4] F. G. Cozman. Algorithms for conditioning on events of zero lower probability. In *Proc. of the Fifteenth Int. Florida Artificial Intelligence Research Society Conf.*, pp. 248–252, Pensacola, Florida, 2002.

[5] B. de Finetti. *Theory of probability, vol. 1-2*. Wiley, New York, 1974.

[6] F. J. Diez. Parameter adjustment in Bayes networks: The generalized noisy OR-gate. In *Proc. Conf. on Uncertainty in Artificial Intelligence*, pp. 99–105, San Francisco, California, 1993. Morgan Kaufmann.

[7] A. Gilio and R. Scozzafava. Conditional events in probability assessment and revision. *IEEE Transactions on Systems, Man and Cybernetics*, 24(12):1741–1746, 1994.

[8] S. Glesner and D. Koller. Constructing flexible dynamic belief networks from first-order probalistic knowledge bases. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pp. 217–226, 1995.

[9] D. Heckerman and J. S. Breese. Causal independence for probability assessment and inference using Bayesian networks. Technical Report MSR-TR-94-08, Microsoft Research, March 1994.

[10] M. Henrion. Some practical issues in constructing belief networks. In *Uncertainty in Artificial Intelligence 3*. North-Holland, Amsterdam, 1989.

[11] M. Jaeger. Relational Bayesian networks. In *Proc. Conf. on Uncertainty in Artificial Intelligence*, pp. 266–273, San Francisco, California, 1997. Morgan Kaufmann.

[12] P. Lucas. Bayesian network modelling by qualitative patterns. In *ECAI*, 2002.

[13] L. Ngo and P. Haddawy. Answering queries from context-sensitive probabilistic knowledge bases. *Theoretical Computer Science*, 171(1–2):147–177, 1997.

[14] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, California, 1988.

[15] D. Poole. Probabilistic Horn abduction and Bayesian networks. *Artificial Intelligence*, 64:81–129, 1993.

[16] P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search (second edition)*. MIT Press, 2000.

[17] S. Srinivas. A generalization of the noisy-OR model. In *Proc Conf. on Uncertainty in Artificial Intelligence*, pp. 208–215, San Francisco, 1993. Morgan Kaufmann.

[18] B. Vantaggi. Graphical representation of asymmetric graphoid structures. In *Third Int. Symp. on Imprecise Probabilities and Their Applications*, pp. 560–574. Carleton Scientific, 2003.

[19] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[20] M. P. Wellman. Qualitative probabilistic networks for planning under uncertainty. *Uncertainty in Artificial Intelligence 2*, pp. 197–208, 1988.

[21] M. P. Wellman and M. Henrion. Explaining 'explaining away'. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(3):287–307, 1993.

[22] N. L. Zhang and D. Poole. Exploiting causal independence in Bayesian network inference. *Journal of Artificial Intelligence Research*, pp. 301–328, 1996.