An Independence of Causal Interactions Model for Opposing Influences

Paul P. Maaskant[1] & **Marek J. Druzdzel**[2,3]

[1] Department of Media and Knowledge Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands
[2] Faculty of Computer Science, Białystok Technical University, Wiejska 45A, 15-351 Białystok, Poland
[3] Decision Systems Laboratory, School of Information Sciences and Intelligent Systems Program, University of Pittsburgh, Pittsburgh, PA 15260, USA

Abstract

We introduce the DeMorgan gate, an Independence of Causal Interactions (ICI) model that is capable of modeling opposing influences, i.e., a mixture of positive and negative influences of parents on a child. The model is a noisy version of a conjunctive normal form of Boolean functions and is an extension and a combination of the popular Noisy-OR and Noisy-AND models, preserving their intuitive semantics. We report the results of a simple experiment testing the usefulness of the proposed model for elicitation of conditional probability distributions.

1 Introduction

Bayesian networks (BNs) (Pearl, 1988) offer a sound framework for reasoning in uncertain problem domains. A Conditional Probability Table (CPT) in a BN specifies the relation between a variable and its immediate predecessors (parents) in the graph. A fundamental problem of CPTs is their exponential growth in the number of parent variables. For nodes with more than a handful of parents, common in practical models, eliciting the CPT from human experts is daunting. So is learning it from data, as there are typically not enough cases to learn every distribution in a CPT reliably.

Independence of Causal Influences (ICI) models (see Díez and Druzdzel (2008) for a comprehensive review of the existing ICI models), provide a solution by assuming that parent variables cause the effect independently of each other. The benefit of this assumption is such that the number of required parameters is linear, rather than exponential, in the number of parent variables. One of the main practical limitations of the existing ICI models is that they cannot model opposing influences, i.e., combinations of influences that increase, and decrease the posterior probability of the child variable. Existing attempts to address this problem are, we believe, weak. In this paper, we propose a new ICI model based on a combination of OR and AND gates in a Conjunctive Normal Form (CNF) of a Boolean expression that combines opposing influences, yet retains a clear parametrization.

A few words about the notation. We will use uppercase letters to denote random variables (e.g., X) and lowercase letters to denote their states (e.g., x). Because all variables in this paper will be Boolean, a variable X will take only two states, x and $\overline{x}$. Bold uppercase letters will denote sets of random variables (e.g., $\mathbf{X}$) and bold lowercase letters (e.g., $\mathbf{x}$) will denote value assignments to sets of random variables. We will use $\Pr(X)$ to denote the probability distribution over a variable X.

2 Foundations

2.1 The Noisy-OR Model

The Noisy-OR model (Pearl, 1988; Henrion, 1989) is a probabilistic extension of the logical OR relation. Its variables, e.g., X, are binary and can be either *present*, denoted as x, or *absent*, denoted as $\overline{x}$. Each *present* parent event can independently produce the child effect. Noisy-OR's amechanistic property assumes that if none of the parent variables $X_1, ..., X_n$ are present, then neither is the child variable Y, i.e.,

$$\Pr(\overline{y}|\overline{x}_1, \ldots, \overline{x}_n) = 1 \ . \tag{1}$$

We define the probability that x_i produces y as

$$\Pr(y|\overline{x}_1, ..., x_i, \ldots, \overline{x}_n) = z_i \ . \tag{2}$$

The ICI assumption allows us to derive the probability of $\mathbf{X}$ producing y as

$$\Pr(y|\mathbf{X}) = 1 - \prod_{X_i = x_i} (1 - z_i) \ .$$

Repeating this process for all possible parent configurations gives us the CPT. Henrion (1989) extended the Noisy-OR model by introducing a leak probability z_L, yielding

$$\Pr(y|\mathbf{X}) = 1 - (1 - z_L) \prod_{X_i = x_i} (1 - z_i) \ . \tag{3}$$

The leak variable z_L represents the probability that the child variable is in its *present* state, even when all the parents are *absent*. An intuitive interpretation of leak is that it represents the effect of unmodeled causes of Y.

2.2 The Amechanistic Property

Amechanistic ICI models (Heckerman and Breese, 1996) are a subclass of ICI models that make two additional assumptions: (1) each variable has a *typical* state, referred to as the *distinguished* state. This is usually the *default* state for that variable. (2) If all parent variables are in their distinguished states, then so is the child variable. For example, if all possible causes of coughing are absent, then coughing is absent as well. In the course of elicitation, we can reduce the mental load needed to imagine a causal influence and estimate its strength by assuming that all other nodes are in the state that is not active and does not interfere with the cause that we are focusing on. The parameters for an amechanistic ICI model can be obtained by asking simple and clear questions and, effectively, such a model is particularly suited for parametrization by human experts.

In the Noisy-OR gate, assumption (2) is captured by Eq. 1, while Eq. 2 expresses the question needed for parameter elicitation. This enables us to directly elicit the probabilistic strength of the causal influence of a parent variable X_i on a child variable Y by asking simple questions, such as "What is the probability of coughing if a patient has pneumonia and no other factors that may cause coughing are present?" We believe that the unquestionable popularity of the Noisy-OR model is in part due to its amechanistic property. Some proposals for canonical gates do not have the amechanistic property and this, we believe, is their major weakness at the outset.

3 Causal Interactions

We describe below four fundamental types of causal interactions between an individual parent X and a child node Y.

Cause This is the most common type of interaction, modeled in the Noisy-OR gate: X is a causal factor and has a positive influence on Y. This influence, just as is the case in the Noisy-OR gate, does not need to be perfect. For example, smoking is quite likely a causal factor in lung cancer. Yet, incidence of lung cancer among smokers, while much larger than incidence of lung cancer among non-smokers, is still within a few percent. Hence, the conditional probability of lung cancer given that a person is a smoker is still fairly low.

The distinguished state of a cause is the state in which the cause has no effect on the child. For example, being a non-smoker has no effect on lung cancer.

Barrier This is a negated counterpart of a cause: X is a factor that decreases the probability of Y. For example, regular exercise decreases the probability of heart disease. While it is a well established factor with a negative influence on heart disease, it is unable by itself to prevent heart disease. One way of looking at a barrier is that it is dual to a cause: Absence of the barrier event is a causal factor for the child, i.e., $\overline{x}$ is a cause. One might go around the very existence of barriers in knowledge engineering by using negated versions of the variables that represent them. In the example above, one might define a variable *Lack of regular exercise*, which would behave as a cause of the variable *Heart disease*. This, however, might become cumbersome if *Regular exercise* participated in other interactions in a model. It might happen, for example, that it is a parent of both *Heart disease* and *Good physical shape*. Because *Regular exercise* decreases the probability of one and increases the probability of the other, barrier, which is a negated cause, is a useful modeling construct.

The distinguished state of a barrier is also the state in which the barrier has no effect on the child. For example, exercise may be thought as not influencing the risk of heart disease and it is the distinguished state in this interaction. We should point out here that the concept of a distinguished state is relative to an interaction and the same variable can have different distinguished states in different interactions that it participates in.

Requirement X is required for Y to be present. There are perfect requirements, such as being a female is a requirement for being pregnant but there are also requirements that are in practice not absolutely necessary. For example, a sexual intercourse is generally believed to be a requirement for pregnancy, but it is not a strict requirement, as pregnancy may be also caused by artificial insemination.

The distinguished state of a requirement is the state that is necessary for the effect to take place at all. For example, being a female is a requirement for becoming pregnant and it is the distinguished state in this interaction.

Inhibitor X inhibits Y. For example, rain may inhibit wild land fire or use of a condom during intercourse with an infected individual may inhibit contracting the HPV virus. Like in the other types of interactions, the parent may be imperfect in inhibiting the occurrence of the child. Fire may start even if there is rain and effectiveness of a condom in protecting from the HPV virus is only around 70%. Similarly to the relationship between causes and barriers, inhibitors are dual to requirements: Absence of an inhibitor event is a requirement for the child.

The distinguished state of an inhibitor is the state that has no effect on the child, i.e., the inhibiting factor being absent. For example, *Rain* is an inhibitor of *Wild land fire*. Its distinguished state is *No rain*, in which case the fire may happen.

4 The DeMorgan Model

We start with the deterministic version of the DeMorgan gate and later extend it to accommodate noise.

4.1 Deterministic DeMorgan Gate

Promoting Influences Promoting influences (causes and barriers) are modeled well by the Noisy-OR gate, which is a noisy version of the following Boolean formula

$$Y = X_1 \vee X_2 \vee \ldots \vee X_m \ , \tag{4}$$

where Xs stand for causes or barriers. The distinguished state of Y is *absent*.

Inhibiting Influences Presence of any inhibitor U_i is sufficient to cancel the child effect. We can express the effect that a set of inhibiting influences (requirements and inhibitors) have on Y by the following Boolean function

$$\overline{Y} \equiv U_1 \vee U_2 \vee \ldots \vee U_n \ , \tag{5}$$

where Us stand for requirements and inhibitors.

Eq. 5 is similar to Eq. 4 but now the parents cancel the child event instead of producing it. Variable Y is *absent* if at least a single U_i is present. We assume that the distinguished state of Us is *absent*. However, contrary to our previous example, we assume the distinguished state of the child Y to be *present*. This assumption is based on common sense: We cannot cancel an event that is not present.

Combining Influences We can combine promoting and inhibiting influences by first applying one of De Morgan's laws to Eq. 5. We get

$$Y \equiv \overline{U}_1 \wedge \overline{U}_2 \wedge \ldots \wedge \overline{U}_n \ . \tag{6}$$

A logical proposition that combines (4) and (6) can only be true for a particular parent configuration if both (4) and (6) are also true for that same configuration. This implies that both equations form a conjunction

$$Y = (X_1 \vee X_2 \vee \ldots \vee X_m) \wedge \overline{U}_1 \wedge \overline{U}_2 \wedge \ldots \wedge \overline{U}_n \ .$$

We now have the logical proposition that we need in order to define the interaction for the combined model with Xs and Us and the child variable Y. Note that the proposition on the right hand-side of Eq. 4.1 is in the CNF. Because each of the conjuncts, save one, consists of a single variable, we can build a simple model to represent this proposition, using only a single OR and a single AND gate.

4.2 Modeling Uncertainty

Noise for Promoting Influences Noise for promoting influences is best modeled by mimicking the Noisy-OR gate, i.e., specifying a causal strength parameter v_i for each of the promoting influences and adding a leak parameter v_L. v_i is the probability of y given that parent i is not in its distinguished state and all other parents are in their distinguished states (please note that we need to choose a different distinguished state for causes and barriers). Eq. 3, the formula for leaky Noisy-OR gives the probability of y as a function of v_is and v_L (the formula uses zs instead of vs).

Noise for Inhibiting Influences The distinguished state of a parent that models an inhibiting influence is its *absent* state, i.e., when the parent is absent it is *certain not to inhibit the child event*, but when present, *it will inhibit the child event with some probability.* When a parent U_i is in its non-distinguished state, we assign a probability d_i that it will inhibit the child event. We include this uncertainty in the network by adding a noise variable W_i that has the following behavior

$$\Pr(w_i|U_i) = \begin{cases} 0 & \text{if } U_i = \overline{u}_i \\ d_i & \text{if } U_i = u_i \end{cases} ,$$

and the child variable Y is equivalent to the conjunction of variables $\overline{W}_1, \ldots, \overline{W}_n$, as given by Eq. 6. We have shown by De Morgan's laws that the Eq. 5 is the logical equivalent of Eq. 6. Eq. 3 gives us the probability of Y occurring

$$\Pr(y|\mathbf{U}) = \begin{cases} \prod_{u_i \in +\mathbf{u}} (1 - d_i) & \text{if } +\mathbf{u} \neq \emptyset \\ 1 & \text{if } +\mathbf{u} = \emptyset \end{cases} .$$

We define $+\mathbf{u}$ to be the subset of $\mathbf{U}$ that contains all parents that are in their non-distinguished states.

It makes little sense to ask for the effect of rain on a bonfire, when the latter is absent. By analogy, we cannot determine d_i directly if we assume that the distinguished state of Y is *absent*. Therefore, we determine d_i relative to an arbitrary set of promoting influences (with a joint effect v on Y) or even the leak parameter, although it seems that elicitation will be more reliable for larger values of v.[1] Suppose we know the effect of a promoting influence X_i, denoted as v_i, and the effect of both X_i and inhibiting influence U_j, denoted as q_j, i.e.,

$$p = 1 - (1 - v_L)(1 - v_i) \ ,$$
$$q_j = (1 - (1 - v_L)(1 - v_i))(1 - d_j) \ .$$

We have $d_j = 1 - q_j/p$.

Derivation of the CPT The total effect of simultaneous presence of noisy promoting and inhibiting causes in a leaky noisy DeMorgan gate can be combined into a CPT as follows:

$$\Pr(y|\mathbf{X}, \mathbf{U}) = (1 - (1 - v_L) \prod_{x_i \in +\mathbf{x}} (1 - v_i)) \prod_{u_j \in +\mathbf{u}} (1 - d_j) \ .$$

[1] Thus, there are combinatorially many questions that we can ask in order to obtain d_j, something not unheard of in probability elicitation.

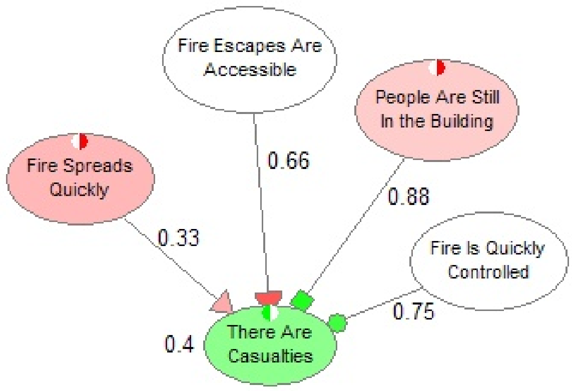

Figure 1: DeMorgan model example network

5 Knowledge Engineering for the DeMorgan Gate

A knowledge engineer has to most of all ensure that a gate elicited can be viewed as DeMorgan gate. The conditions that have to be fulfilled for the DeMorgan gate are similar to those listed for other canonical gates by Díez and Druzdzel (2008): Each parent must be able to cause or to inhibit the child node through a separate causal mechanism and there may be no significant interactions among these mechanisms.

Now, for each type of interaction, q_i, the parameter associated with the causal link from a parent X_i corresponds to the probability of the effect y happening if all parents but X_i are in their distinguished states. The leak parameter v_L expresses the probability of y given that all parents are in their distinguished states.

Consider the following network based on DeMorgan gate with one cause (*Fire Spreads Quickly*), one barrier (*Fire Escapes Are Accessible*), one requirement (*People Are Still In the Building*), and one inhibitor (*Fire Is Quickly Controlled*).

We will now give example questions to be asked of an expert. Please note that there is a natural discrepancy between what one has to say formally and what sounds clear to a human. Each of the questions listed below can be adjusted to the needs of particular context, i.e., their elements can be rephrased or omitted if they do not make sense.

The leak parameter "What is the probability of casualties if the fire does not spread quickly, fire escapes are not accessible, people are still in the building, and fire is not quickly controlled? Please note that casualties may happen due to other, unmodeled causes."

Cause "What is the probability of casualties if the fire spreads quickly, fire escapes are not accessible, people are still in the building, fire is not quickly controlled, and no other unmodeled causal factors are present?"

Barrier "What is the probability of casualties if the fire does not spread quickly, fire escapes are accessible, people are still in the building, fire is not quickly controlled, and no other unmodeled causal factors are present?"

Requirement "What is the probability of casualties if the fire does not spread quickly, fire escapes are not accessible, there are no people in the building, fire is not quickly controlled, and no other unmodeled causal factors are present?" Please note that the possible casualties are due to the fact that information concerning absence of people in the building may be false or the casualties may be that of the fire fighters.

Inhibitor "What is the probability of casualties if the fire does not spread quickly, fire escapes are not accessible, there are people in the building, fire is quickly controlled, and no other unmodeled causal factors are present?"

6 Empirical Evaluation

To validate the DeMorgan model, we conducted an experiment based on the methodology for evaluating probability elicitation schemes introduced by Wang et al. (2002). Its main advantage is that it controls for a-priori domain knowledge on the part of the subjects. The subjects are first asked to learn an abstract domain, which they have never seen before (typically an abstract interactive computer game). Since every subject may have a different set of experiences in the course of their interaction with the new domain, recording these provides us with a gold standard of the frequency observed by the subject. A perfect elicitation scheme should retrieve these frequencies and the experimental setup aims at comparing elicitation schemes on how well they do so.

6.1 Subjects

Our subjects were 24 students in a graduate course *Decision Analysis and Decision Support Systems* in the School of Information Sciences, University of Pittsburgh. The students were familiar with, although not experts in, decision analysis, probability theory, and BNs. For their participation, they received a small course credit and a handful of M&Ms.

6.2 Experiment Design

The subjects were asked to play a simple, fictional computer game resembling a black box with four propositional inputs (X_1, X_2, X_3, and X_4) and one propositional output (Y) with states *Success* and *Failure*. Their task was to obtain *Success* at the output by means of selecting a combination of inputs. Subjects were allowed 160 trials, each trial consisting of three phases: (1) selecting values for X_1 through X_4, (2) pressing a key, and (3) observing the value of Y. The value of Y was chosen randomly by means of sampling from a DeMorgan gate, although the subjects were not aware of it. Two of the inputs (assigned randomly) were causes and the remaining two inputs were barriers. Model parameters were randomly chosen for each of the subject from the intervals $[0.5, 0.9]$ (causes), $[0.3, 0.9]$ (barriers) and $[0.1, 0.3]$ (the leak). Each subject faced thus a different probabilistic model driving the game.

Because of a relatively small number of subjects, we used a within-subject design. At the conclusion of the training phase, the subjects were asked (1) to give the full CPT ($\Pr(Y|X_1, X_2, X_3, X_4)$), consisting of 16 entries, and (2) indicate which inputs were promoting and which were inhibiting influences, assess their strengths and the leak probability. The order of the two elicitations was randomized to compensate for a possible carry-over effect.

6.3 Experiment Results

We used the probability distribution *observed* by each subject as the gold standard of what the subject knew. For each value OBS_i of the observed CPT, we calculated the maximum a-posteriori estimate given the subject's 160 observations, using a Beta prior distribution with a very small equivalent sample size (in order to avoid zero probabilities), i.e.,

$$OBS_i = \frac{s_i + 0.01}{t_i + 0.02} \; ,$$

where s_i denotes the number of successful trials, and t_i the total number of trials for input configuration i.

Of interest to us was the elicitation error, i.e., the difference between the observed CPT and the elicited probability distributions. We measured the error by the averaged Euclidean and Hellinger distances (Kokolakis and Nanopoulos, 2001). Because both measures are defined for single distributions, we averaged errors across all 16 distributions in the CPT.

The subjects each took between 30 and 45 minutes to complete the experiment. We judged one of the subjects to be an outlier, and excluded the subject from further analysis. This subject likely confused the concept of inhibiting with promoting, as she reported very low probabilities in cases where she observed very high probabilities and vice versa.

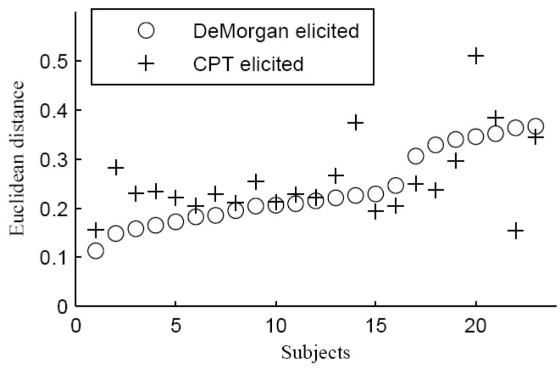

Figure 2: Raw data (sorted by increasing distance for the DeMorgan model)

Figure 2 shows raw data, i.e., the Euclidean distance for each subject: (1) the distance between the observed CPT and the CPT generated by the elicited DeMorgan model, sorted from the smallest to the largest distance, and (2) the distance between the observed CPT and the directly elicited CPT. We would like to point out that the range of distances is lower for the DeMorgan gate. Table 1 shows the average Euclidean and Hellinger distances across all subjects. A one-tailed paired t-tests performed on both distance measures yielded $p \approx 0.14$ for the Euclidean, and $p \approx 0.29$ for the Hellinger distance, showing no significant difference in accuracy at the commonly used $\alpha = 0.05$ significance level. Although the accuracy gain in favor of the DeMorgan model was not statistically significant, our results suggest that the CPT generated by DeMorgan model is at least as accurate as a directly elicited CPT. This becomes a non-trivial advantage when the number of parent variables is larger. And so, for a family with 10 parent variables, we have 21 questions for the DeMorgan model, versus 1,024 questions needed to elicit the CPT directly.

Measure	DeMorgan	CPT
Averaged Euclidean Distance	0.2382	0.2566
Weighted Hellinger Distance	0.2481	0.2563

Table 1: Averaged Euclidean and Hellinger distances.

7 Related Work

Inhibitors are mentioned by Pearl (1988), who calls them *global inhibitors* and lays the foundations for both requirements and inhibitors, as proposed in DeMorgan gate. Pearl stops short, however, from combining logical OR and AND gates with negation, which is what DeMorgan gate does.

Srinivas (1993) generalizes the Noisy-OR model to multiple states and proposed a model that is known as the "feeding lines model," embodying a world of possible functions that tie a node to its parents. It is quite likely that there exist functions among all possible that will combine positive and negative influences. Srinivas' proposal for an extension of Noisy-OR has never been adopted and we are not aware of any work extending the "feeding lines model."

Heckerman and Breese (1994) and later Lucas (2005) discuss the foundations of ICI models and draw attention to so called *decomposable* ICI models. Lucas analyzes in depth canonical models based on Boolean functions, reminding that there are 2^{2^n} different n-ary Boolean functions and so is the potential number of causal interactions. The DeMorgan model is decomposable, although it does not decompose into identical functions. It is indeed one of a huge number of possibilities, but as we argue in this paper, it may well be one of few that are intuitive and potentially readily adopted in practice ICI models.

A proposal for combining positive and negative influences has been the CAusal STrength (CAST) model (Chang et al., 1994), which is an extension of BNs that is able to model simultaneous opposing influences. Although very popular, particularly in government and military applications, a major weakness of the CAST model is its unclear parametrization. Parents can influence a child variable in both of their states and do not have a distinguished state, hence, are not amechanistic.

Lemmer and Gossink's *recursive Noisy-OR model* (Lemmer and Gossink, 2004) deals with positive and negative influences, although not in the same model, i.e., a model includes either all positive or all negative influences.

Finally, Xiang and Jia (Xiang and Jia, 2007) proposed a general model based on combining Noisy-AND gates with negation, apparently developed independently from this proposal. That model is capable of modeling positive and negative influences similarly to our proposal.

8 Conclusions

An important property of the DeMorgan model is that it is able to model any logical interaction between inputs, when their influences on the output are independent, i.e., when they are ICI. In particular, DeMorgan gate can handle a combination of positive and negative influences, while preserving both probabilistic soundness and the amechanistic property, critical in probability elicitation. Probabilistic soundness ensures that it is mathematically correct, and propositional logic, that lies at its foundations, ensures that our model is meaningful and intuitive for humans.

The results of our experiment indicate that elicitation of a small DeMorgan model is at least as accurate as direct elicitation of a CPT. Yet, the DeMorgan model requires a number of parameters that is linear, rather than exponential, in the number of parent variables. We expect that the DeMorgan model will show a great advantage over direct elicitation especially for larger models. We have embedded the DeMorgan model in SMILE and QGeNIe, a qualitative interface to SMILE, our probabilistic reasoning engine, and made it available to the community (http://genie.sis.pitt.edu/). QGeNIe is useful in rapid modeling of problems involving propositional variables. We are currently working on extending the DeMorgan model to multi-valued variables along the lines of the Noisy-MAX and Noisy-MIN gates.

Acknowledgments

This work has been supported by the Air Force Office of Scientific Research grant FA9550-06-1-0243 and by Intel Research. All experimental data were obtained using SMILE, a Bayesian inference engine developed at the Decision Systems Laboratory and available at http://genie.sis.pitt.edu/. We would like to thank Adam Zagorecki for stimulating discussions on the topic of amechanistic ICI models and help with our experiment. We thank anonymous reviewers for the PGM–06 for valuable suggestion that improved the paper. Majority of this work was performed while the first author was at the Decision Systems Laboratory.

References

K.C. Chang, P.E. Lehner, A.H. Levis, Abbas K. Zaidi, and X. Zhao. 1994. On causal influence logic. Technical report, George Mason University, Center of Excellence for C3I.

F. Javier Díez and Marek J. Druzdzel. 2008. Canonical probabilistic models for knowledge engineering. Unpublished manuscript, available at http://www.ia.uned.es/~fjdiez/papers/canonical.html.

David Heckerman and John S. Breese. 1994. A new look at causal independence. In *Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence (UAI–94)*, pages 286–292, San Mateo, CA. Morgan Kaufmann Publishers, Inc.

David Heckerman and John S. Breese. 1996. Causal independence for probability assessment and inference using Bayesian networks. *IEEE, Systems, Man, and Cybernetics*, 26:826–831.

Max Henrion. 1989. Some practical issues in constructing belief networks. In L.N. Kanal, T.S. Levitt, and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence 3*, pages 161–173. Elsevier Science Publishing Company, Inc., New York, N. Y.

G. Kokolakis and P.H. Nanopoulos. 2001. Bayesian multivariate micro-aggregation under the Hellinger's distance criterion. *Research in Official Statistics*, 4(1):117–126.

John F. Lemmer and D.E. Gossink. 2004. Recursive noisy-OR: A rule for estimating complex probabilistic causal interactions. *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics*, 34(6):2252–2261.

Peter J.F. Lucas. 2005. Bayesian network modeling through qualitative patterns. *Artificial Intelligence*, 163(2):233–263.

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

Sampath Srinivas. 1993. A generalization of the noisy-OR model. In *Proceedings of the Ninth Annual Conference on Uncertainty in Artificial Intelligence (UAI–93)*, pages 208–215, San Francisco, CA. Morgan Kaufmann Publishers.

Haiqin Wang, Denver H. Dash, and Marek J. Druzdzel. 2002. A method for evaluating elicitation schemes for probabilistic models. *Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, 32(1):38–43.

Y. Xiang and N. Jia. 2007. Modeling causal reinforcement and undermining for efficient CPT elicitation. *IEEE Transactions on Knowledge and Data Engineering*, 19(12):1708–1718.