Proceedings of the Twenty-Third International Florida Artificial Intelligence Research Society Conference (FLAIRS 2010)

Generalized Non-impeding Noisy-AND Trees

Yang Xiang
University of Guelph, Canada

Abstract

To specify a Bayes net (BN), a conditional probability table (CPT), often of an effect conditioned on its n causes, needs assessed for each node. Its complexity is generally exponential in n. Noisy-OR reduces the complexity to linear, but can only represent reinforcing causal interactions. The non-impeding noisy-AND (NIN-AND) tree is the first causal model that explicitly expresses reinforcement, undermining, and their mixture. It has a linear complexity, in terms of both the number of parameters and the size of the tree topology. As originally proposed, the model allows only binary effect and cause variables. This work generalizes the model to multi-valued effect and causes, and analyzes key properties.

Introduction

To specify a BN, a CPT needs to be assessed for each non-root node. It is often advantageous to construct BNs along the causal direction, in which case a CPT is the distribution of an effect conditioned on its n causes. In general, assessment of a CPT has the complexity exponential on n.

Noisy-OR (Pearl 1988) is the most well known technique that reduces this complexity to linear. A number of extensions have also been proposed such as (Heckerman & Breese 1996; Galan & Diez 2000; Lemmer & Gossink 2004). However, noisy-OR, noisy-AND (Galan & Diez 2000), as well as related techniques, can only represent causal interactions that are reinforcing (Xiang & Jia 2007).

The NIN-AND tree (Xiang & Jia 2007) extends noisy-OR and provides the first causal model that explicitly expresses reinforcing and undermining causal interactions, as well as their mixture.[1] It requires specification of a set of probability parameters of a size linear in n, and a tree topology also of a size linear in n, which expresses the types of causal interactions among causes. The model uses default independence assumptions to gain the efficiency, but is also flexible enough to allow these assumptions to be relaxed. With the assumptions relaxed incrementally and more parameters are specified accordingly, any CPT can be encoded through a NIN-AND tree.

As originally proposed (Xiang & Jia 2007), the effect and cause variables in a NIN-AND tree are binary, which limits its scope of applicability. In this work, we draw from the generalization of noisy-OR from the binary case, such as (Henrion 1989; Diez 1993), and generalize the NIN-AND tree model to multi-valued effect and cause variables.

The remainder of the paper is organized as follows: We first review the binary NIN-AND tree models. We then introduce our terminology on graded multi-causal events. The basic processing units in a NIN-AND tree model, the NIN-AND gates, are generalized to graded multi-causal events. This is followed by the definition of the generalized NIN-AND tree model. We analyze its properties in relation to reinforcement and undermining, as well as the complexity for specifying a CPT using such a model.

Background on Binary NIN-AND Trees

This section is mostly based on (Xiang & Jia 2007). An *uncertain cause* is a cause that can produce an effect but does not always do so. Denote a binary effect variable by e and a set of binary cause variables of e by $X = \{c_1, ..., c_n\}$. Denote $e = true$ by e^+ and $e = false$ by e^-. Similarly, for each cause c_i, denote $c_i = true$ by c_i^+ and $c_i = false$ by c_i^-.

A *causal event* refers to an event that a cause c_i caused an effect e to occur successfully when all other causes of e are absent. Denote this causal event by $e^+ \leftarrow c_i^+$ and its probability by $P(e^+ \leftarrow c_i^+)$. The causal failure event, where e is false when c_i is true and all other causes of e are false, is denoted by $e^+ \not\leftarrow c_i^+$. Denote the causal event that a set $X = \{c_1, ..., c_n\}$ of causes caused e by $e^+ \leftarrow c_1^+, ..., c_n^+$ or $e^+ \leftarrow \underline{x}^+$. Denote the set of *all causes* of e by C.

The CPT $P(e|C)$ relates to probabilities of causal events as follows: If $C = \{c_1, c_2, c_3\}$, then $P(e^+|c_1^+, c_2^-, c_3^+) = P(e^+ \leftarrow c_1^+, c_3^+)$. C is assumed to include a leaky variable (if any) to capture causes that we do not wish to represent explicitly, and hence $P(e^+|c_1^-, c_2^-, c_3^-) = 0$.

Causes reinforce each other if collectively they are at least as effective in causing the effect as some acting by themselves. If collectively they are less effective, then they undermine each other. Note that if $C = \{c_1, c_2\}$ and c_1 and c_2 undermine each other, then all the following hold:

$$P(e^+|c_1^-, c_2^-) = 0,\ \ P(e^+|c_1^+, c_2^-) > 0,\ \ P(e^+|c_1^-, c_2^+) > 0,$$

$$P(e^+|c_1^+, c_2^+) < min(P(e^+|c_1^+, c_2^-), P(e^+|c_1^-, c_2^+)).$$

The following Def.1 defines the two types of causal interactions generally. Note that reinforcement and undermining can occur between individual variables as well as sets of variables. For instance, variables within each of two sets can be reinforcing, while the two sets can undermine each other. Hence, each W_i in Def.1 is not necessarily a singleton.

[1]Being unaware of this work and its precursor, (Maaskant & Druzdzel 2008) independently presented special cases of NIN-AND tree models.

Definition 1 *Let $R = \{W_1, W_2, ...\}$ be a partition of a set X of causes, $R' \subset R$ be any proper subset of R, and $Y = \cup_{W_i \in R'} W_i$. Sets of causes in R* **reinforce** *each other, iff*

$$\forall R' \; P(e^+ \leftarrow \underline{y}^+) \leq P(e^+ \leftarrow \underline{x}^+).$$

Sets of causes in R **undermine** *each other, iff*

$$\forall R' \; P(e^+ \leftarrow \underline{y}^+) > P(e^+ \leftarrow \underline{x}^+).$$

Disjoint sets of causes $W_1, ..., W_m$ satisfy *failure conjunction* iff

$$(e^+ \not\leftarrow \underline{w}_1^+, ..., \underline{w}_m^+) = (e^+ \not\leftarrow \underline{w}_1^+) \wedge ... \wedge (e^+ \not\leftarrow \underline{w}_m^+).$$

That is, collective failure is attributed to individual failures. They also satisfy *failure independence* iff

$$\begin{aligned} & P((e^+ \not\leftarrow \underline{w}_1^+) \wedge ... \wedge (e^+ \not\leftarrow \underline{w}_m^+)) \\ = \; & P(e^+ \not\leftarrow \underline{w}_1^+) \; ... \; P(e^+ \not\leftarrow \underline{w}_m^+). \end{aligned} \quad (1)$$

Disjoint sets of causes $W_1, ..., W_m$ satisfy *success conjunction* iff

$$e^+ \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+ = (e^+ \leftarrow \underline{w}_1^+) \wedge ... \wedge (e^+ \leftarrow \underline{w}_m^+).$$

That is, collective success requires individual effectiveness. They also satisfy *success independence* iff

$$\begin{aligned} & P((e^+ \leftarrow \underline{w}_1^+) \wedge ... \wedge (e^+ \leftarrow \underline{w}_m^+)) \\ = \; & P(e^+ \leftarrow \underline{w}_1^+) \; ... \; P(e^+ \leftarrow \underline{w}_m^+). \end{aligned} \quad (2)$$

It can be shown that causes are reinforcing when they satisfy failure conjunction and independence, and they are undermining when they satisfy success conjunction and independence. Undermining can be modeled by a direct NIN-AND gate as shown in the left of Fig. 1. Its root nodes (top)

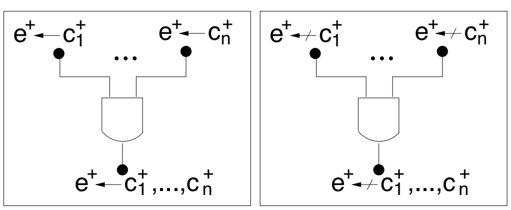

Figure 1: Direct (left) and dual (right) NIN-AND gates

are causal success events of single causes, and its leaf node (bottom) is the causal event in question, whose probability is computed by Eqn. (1). Reinforcement can be modeled by a dual NIN-AND gate (right). Its root nodes (top) are causal failure events of single causes, and its leaf node (bottom) is the causal failure event in question, whose probability is computed by Eqn. (2).

By combining direct and dual NIN-AND gates and organizing them into a tree topology, both reinforcement and undermining can be expressed in a single model, called a NIN-AND tree. Consider an example where $C = \{c_1, c_2, c_3\}$, c_1 and c_3 undermine each other, but collectively they reinforce c_2. Assuming the default conjunction and independence, their causal interaction, relative to the event

$$e^+ \leftarrow c_1^+, c_2^+, c_3^+,$$

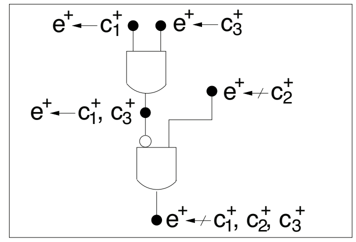

Figure 2: A NIN-AND tree causal model.

can be expressed by the NIN-AND tree shown in Fig. 2. The top gate is direct and the bottom gate (the *leaf gate*) is dual. The link downward from node $e^+ \leftarrow c_1^+, c_3^+$ has a white oval end (a *negation* link) and negates the event. All other links are *forward* links. Given an NIN-AND tree, the probability of the leaf event can be computed by Algorithm 1.

Algorithm 1 *GetCausalEventProb(T)*
Input: A NIN-AND tree T of leaf v and leaf gate g, with root probabilities specified.
for each node w directly inputting to g, do
  if $P(w)$ is not specified,
    denote the sub-NIN-AND-tree with w as the leaf by T_w;
    $P(w) = GetCausalEventProb(T_w)$;
  if (w, g) is a forward link, $P'(w) = P(w)$;
  else $P'(w) = 1 - P(w)$;
return $P(v) = \prod_w P'(w)$;

For the example in Fig. 2, after the following are specified,

$$P(e^+ \leftarrow c_1^+) = 0.85, \; P(e^+ \leftarrow c_2^+) = 0.8, \; P(e^+ \leftarrow c_3^+) = 0.7,$$

the probability $P(e^+ \not\leftarrow c_1^+, c_2^+, c_3^+) = 0.081$ can be derived. Using other NIN-AND tree models simplified from Fig. 2, the CPT in Table 1 can be derived. $P(e^+|c_1^+, c_2^-, c_3^+)$ is less than either $P(e^+|c_1^+, c_2^-, c_3^-)$ or $P(e^+|c_1^-, c_2^-, c_3^+)$ (undermining). $P(e^+|c_1^+, c_2^+, c_3^+)$ is larger than both $P(e^+|c_1^+, c_2^-, c_3^+)$ and $P(e^+|c_1^-, c_2^+, c_3^-)$ (reinforcement).

Table 1: The CPT of an example NIN-AND tree model.

$P(e^+\|c_1^-, c_2^-, c_3^-)$	0	$P(e^+\|c_1^+, c_2^-, c_3^+)$	0.595
$P(e^+\|c_1^+, c_2^-, c_3^-)$	0.85	$P(e^+\|c_1^+, c_2^+, c_3^-)$	0.97
$P(e^+\|c_1^-, c_2^+, c_3^-)$	0.8	$P(e^+\|c_1^-, c_2^+, c_3^+)$	0.94
$P(e^+\|c_1^-, c_2^-, c_3^+)$	0.7	$P(e^+\|c_1^+, c_2^+, c_3^+)$	0.919

Graded Multi-Causal Events

Let e be a multi-valued effect variable whose finite domain is denoted $D_e = \{e^0, e^1, , ..., e^\eta\}$, where $\eta \geq 1$. The value e^0 (through the superscript index 0) represents the absence of the effect condition. Each value e^j with a higher superscript index $j > 0$ represents the effect condition at a higher intensity. For instance, if e represents the fever condition of a patient, it may have a domain $\{c_i^0, c_i^1, c_i^2\}$ which corresponds to

$$\{normal, low\ fever, high\ fever\}.$$

Notation $e < e^j$ is well defined, when $0 < j \leq \eta$, to denote $e \in \{e^0, e^1, ..., e^{j-1}\}$, and so is $e \geq e^j$.

Let c_i ($i = 1, 2, ...$) be a multi-valued uncertain cause, whose finite domain is denoted $D_i = \{c_i^0, c_i^1, c_i^2, ...\}$. The value c_i^0 represents the absence of the condition signified by the variable c_i, and each value c_i^j with a higher superscript index $j > 0$ represents the condition at a higher intensity. Variables such as e and c_i are often referred to as *graded* (Diez 1993).

We denote a set of multi-valued cause variables of effect e (multi-valued) as $X = \{c_1, ..., c_n\}$. The set of *all causes* of e is denoted by C. Set C is assumed to include a leaky variable (if any) to capture causes not represented explicitly.

For multi-valued causes and effect, a *graded singular causal success* is an event that a cause c_i with value c_i^j ($j > 0$) caused the effect e to occur at a value e^k ($k > 0$) or higher, when every other cause c_m of e has the value c_m^0 (absent). Condition $k > 0$ means that the effect must be present. Denote this event by

$$e \geq e^k \leftarrow \{c_i^j\} \quad \textit{or simply} \quad e \geq e^k \leftarrow c_i^j$$

and its probability by $P(e \geq e^k \leftarrow c_i^j)$.

A *graded multi-causal success* involves a set X ($|X| > 1$) of causes of e, where each $c_i \in X$ has a value c_i^j ($j > 0$). That is, causes in X collectively caused the effect e to occur at a value e^k ($k > 0$) or higher, when every other cause $c_m \in C \setminus X$ has the value c_m^0. We denote the multi-causal success by

$$e \geq e^k \leftarrow \{c_1^{j_1}, ..., c_n^{j_n}\} \quad \textit{or simply} \quad e \geq e^k \leftarrow c_1^{j_1}, ..., c_n^{j_n}$$

or by the (somewhat abused) vector notion

$$e \geq e^k \leftarrow \underline{x}^+,$$

where superscript $+$ signifies that, for each $c_i \in X$, its value $c_i^{j_i} > c_i^0$.

A *graded singular causal failure* refers to an event where $e < e^k$ ($k > 0$) when a cause c_i has a value c_i^j ($j > 0$) and every other cause c_m of e has the value c_m^0. It is a failure event in the sense that c_i fails to produce the effect with an intensity e^k or higher. We denote the failure event by

$$e < e^k \leftarrow c_i^j.$$

In a *graded multi-causal failure*, a set X ($|X| > 1$) of causes of e are active when the effect $e < e^k$ ($k > 0$). That is, $e < e^k$, each $c_i \in X$ has a value c_i^j ($j > 0$), and each $c_m \in C \setminus X$ has the value c_m^0. We denote the failure event by

$$e < e^k \leftarrow c_1^{j_1}, ..., c_n^{j_n}$$

or by the vector notion

$$e < e^k \leftarrow \underline{x}^+.$$

Note that our terminology on multi-valued causal events differs from those based on inhibitors, e.g., (Pearl 1988; Heckerman & Breese 1996), and is more coherent with those in (Lemmer & Gossink 2004; Xiang & Jia 2007), although the latter deal with only binary cases.

The negation of event

$$e \geq e^k \leftarrow c_1^{j_1}, ..., c_n^{j_n}$$

is

$$e < e^k \leftarrow c_1^{j_1}, ..., c_n^{j_n}$$

and vice versa.

Probabilities of graded causal events can be converted to conditional probabilities and vice versa through the following proposition, whose proof is straightforward. For a set Y of causes, if $c_j = c_j^0$ for each $c_j \in Y$, we denote the instantiation of Y by $\underline{y}^0$.

Proposition 1 *Let e be an effect, $C = X \cup Y$ ($X \cap Y = \emptyset$) be the set of all causes of e, X be instantiated to $\underline{x}^+$, and Y be instantiated to $\underline{y}^0$. Then the following hold, where $k > 0$.*

1. $P(e \geq e^k \leftarrow \underline{x}^+) = 1 - P(e < e^k \leftarrow \underline{x}^+)$.
2. $P(e^0|\underline{x}^+, \underline{y}^0) = 1 - P(e \geq e^1 \leftarrow \underline{x}^+)$.
3. $P(e^\eta|\underline{x}^+, \underline{y}^0) = P(e \geq e^\eta \leftarrow \underline{x}^+)$.
4. *For $k < \eta$,*

   $$P(e^k|\underline{x}^+, \underline{y}^0) = P(e \geq e^k \leftarrow \underline{x}^+) - P(e \geq e^{k+1} \leftarrow \underline{x}^+).$$
5. $P(e \geq e^k \leftarrow \underline{x}^+) = \sum_{j=k}^{\eta} P(e^k|\underline{x}^+, \underline{y}^0)$.

The first equation in Proposition 1 deals with negation of a causal event. The next three convert causal probabilities to conditional probabilities. The last one converts conditional probabilities to a causal probability. These conversions are useful in processing the probabilities of input and output events of generalized NIN-AND trees as will be presented below.

Generalized NIN-AND Gates

Definition 2 *Disjoint sets of causes $W_1, ..., W_m$ of effect e satisfy* **graded success conjunction** *iff*

$$e \geq e^k \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+$$
$$= (e \geq e^k \leftarrow \underline{w}_1^+) \wedge ... \wedge (e \geq e^k \leftarrow \underline{w}_m^+),$$

where $k > 0$.

Definition 3 *Disjoint sets of causes $W_1, ..., W_m$ of effect e satisfy* **graded success independence** *iff events*

$$e \geq e^k \leftarrow \underline{w}_1^+, \quad ..., \quad e \geq e^k \leftarrow \underline{w}_m^+$$

are independent of each other, where $k > 0$. That is, the following equation holds,

$$\begin{aligned} & P(e \geq e^k \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+) \\ = \ & P(e \geq e^k \leftarrow \underline{w}_1^+) \ ... \ P(e \geq e^k \leftarrow \underline{w}_m^+). \end{aligned} \quad (3)$$

We depict the interaction of causes that satisfy graded success conjunction and graded success independence by a graphical model as shown in Fig. 3. The success conjunction is represented by the AND gate. The success independence is signified by the disconnection of input events other than through the gate. Since the causes are uncertain causes, the AND gate is noisy. Common noisy-AND gates, e.g., (Galan & Diez 2000), are *impeding* in that the probability of a causal event is zero unless the set of active causes is equal to C. The probability of the output event of the gate in Fig. 3 is determined by Eqn. (3) from probabilities of the

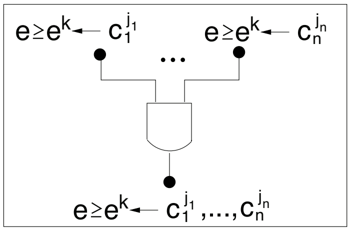

Figure 3: A generalized direct NIN-AND gate.

input events, no matter $X = C$ or not. Hence, the gate is *non-impeding*. To distinguish it from the binary case (see the background section) as well as the case introduced below, we term the gate in Fig. 3 as a *generalized direct non-impeding noisy-AND gate* or a generalized direct NIN-AND gate.

Definition 4 *Disjoint sets of causes $W_1, ..., W_m$ of effect e satisfy* **graded failure conjunction** *iff*

$$e < e^k \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+$$
$$= (e < e^k \leftarrow \underline{w}_1^+) \wedge ... \wedge (e < e^k \leftarrow \underline{w}_m^+),$$

where $k > 0$.

Definition 5 *Disjoint sets of causes $W_1, ..., W_m$ of effect e satisfy* **graded failure independence** *iff failure events*

$$e < e^k \leftarrow \underline{w}_1^+, \quad ..., \quad e < e^k \leftarrow \underline{w}_m^+$$

are independent of each other, where $k > 0$. That is, the following equation holds,

$$\begin{aligned} & P(e < e^k \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+) \\ = \; & P(e < e^k \leftarrow \underline{w}_1^+) \, ... \, P(e < e^k \leftarrow \underline{w}_m^+). \end{aligned} \quad (4)$$

We depict the interaction of causes that satisfy graded failure conjunction and graded failure independence by a graphical model as shown in Fig. 4. The failure conjunction is

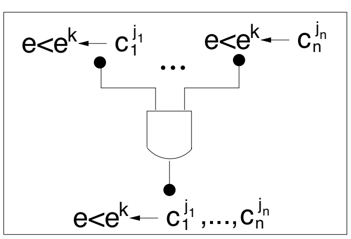

Figure 4: A generalized dual NIN-AND gate.

represented by the AND gate, and the failure independence is signified by the disconnection of input events other than through the gate. The probability of the output event of the gate is determined by Eqn. (4) from probabilities of the input events. The gate in Fig. 4 differs from that in Fig. 3 in that all input and output events are causal failure events. Hence, we refer to it as a *generalized dual NIN-AND gate*.

Def. 2 through 5 are relative to sets of causes. Figs. 3 and 4 are special cases where these sets are singletons. A more general example appears in Fig. 5 below.

Reinforcing and Undermining Properties

We analyze the reinforcing and undermining behaviors of generalized NIN-AND gates, which differ from those of binary NIN-AND gates. We first give a more refined definition of reinforcing and undermining.

Definition 6 *Let D_e be the domain of effect e and S_e be a subset of D_e, where either S_e contains a single element e^k ($k > 0$), or it contain all values $\geq e^k$.*

Let $R = \{W_1, W_2, ...\}$ be a partition of a set X of causes of effect e, $R' \subset R$ be any proper subset of R, and $Y = \cup_{W_i \in R'} W_i$. Denote $V = C \setminus X$ and $Z = C \setminus Y$.

Sets of causes in R **reinforce** *each other* **relative to** S_e*, iff*

$$\forall R' \; P(e \in S_e | \underline{y}^+, \underline{z}^0) \leq P(e \in S_e | \underline{x}^+, \underline{v}^0).$$

Sets of causes in R **undermine** *each other* **relative to** S_e*, iff*

$$\forall R' \; P(e \in S_e | \underline{y}^+, \underline{z}^0) > P(e \in S_e | \underline{x}^+, \underline{v}^0).$$

Note that Def. 6 is defined based on conditional probabilities rather than (causal) probabilities of causal events as Def. 1. This is because reinforcement and undermining are best described through comparison of conditional probabilities. In the binary case, the conversion between conditional and causal probabilities is trivial, but it is less so in the multi-valued case.

In the following, we show that a generalized direct NIN-AND gate models undermining relative to certain S_e's.

Proposition 2 *Let $W_1, ..., W_m$ be disjoint sets of causes of effect e and $e \geq e^\eta \leftarrow \underline{w}_1^+, \quad ..., \quad e \geq e^\eta \leftarrow \underline{w}_m^+$ be the root (input) events of a generalized direct NIN-AND gate g. Let $P(e \geq e^\eta \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+)$ be the probability of the leaf (output) event of g. Then, for $i = 1, ..., m$, we have*

$$P(e \geq e^\eta \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+) < P(e \geq e^\eta \leftarrow \underline{w}_1^+).$$

Proposition 2 says that a generalized direct NIN-AND gate models undermining relative to the most intensive value of the effect, i.e., $S_e = \{e^\eta\}$. It follows directly from Eqn. (3).

Proposition 3 *Let $W_1, ..., W_m$ be disjoint sets of causes of effect e and $e \geq e^1 \leftarrow \underline{w}_1^+, \quad ..., \quad e \geq e^1 \leftarrow \underline{w}_m^+$ be the root (input) events of a generalized direct NIN-AND gate g. Let $P(e \geq e^1 \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+)$ be the probability of the leaf (output) event of g. Then, for $i = 1, ..., m$, we have*

$$P(e \geq e^1 \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+) < P(e \geq e^1 \leftarrow \underline{w}_1^+).$$

Proposition 3 says that a generalized direct NIN-AND gate models undermining relative to the collection of active values of the effect, i.e., $S_e = \{e^1, ..., e^\eta\}$. It follows directly from Eqn. (3).

Consider an example where $C = \{c_1, c_2\}$, $|D_1| = 2$, $|D_2| = |D_e| = 3$, and

$$P(e^1 | c_1^1, c_2^0) = 0.3, P(e^2 | c_1^1, c_2^0) = 0.45,$$
$$P(e^1 | c_1^0, c_2^1) = 0.35, P(e^2 | c_1^0, c_2^1) = 0.22,$$
$$P(e^1 | c_1^0, c_2^2) = 0.4, P(e^2 | c_1^0, c_2^2) = 0.5.$$

Using suitable generalized direct NIN-AND gates with all root events singular, we can derive the following by Eqn. (3),

$$P(e \geq e^1 | c_1^1, c_2^2) = 0.675,$$

$$P(e \geq e^2 | c_1^1, c_2^2) = 0.225,$$

from which we obtain the following by Proposition 1,

$$P(e^0 | c_1^1, c_2^2) = 0.325,$$
$$P(e^1 | c_1^1, c_2^2) = 0.45,$$
$$P(e^2 | c_1^1, c_2^2) = 0.225.$$

Undermining holds relative to $S_e = \{e^2\}$ because

$$P(e^2 | c_1^1, c_2^2) = 0.225 < 0.45 = P(e^2 | c_1^1, c_2^0)$$

and $0.225 < 0.5 = P(e^2 | c_1^0, c_2^2)$. It also holds relative to $S_e = \{e^1, e^2\}$ because

$$P(e \geq e^1 | c_1^1, c_2^2) = 0.675 < 0.75 = P(e \geq e^1 | c_1^1, c_2^0)$$

and $0.675 < 0.9 = P(e \geq e^1 | c_1^0, c_2^2)$. However, undermining does not hold relative to $S_e = \{e^1\}$ because

$$P(e^1 | c_1^1, c_2^2) = 0.45 > 0.3 = P(e^1 | c_1^1, c_2^0)$$

and $0.45 > 0.4 = P(e^1 | c_1^0, c_2^2)$.

Similarly, we show below that a generalized dual NIN-AND gate models reinforcement relative to both the most intensive value of the effect and the collection of active values of the effect.

Proposition 4 *Let $W_1, ..., W_m$ be disjoint sets of causes of effect e and $e \geq e^k \leftarrow \underline{w}_1^+, \quad ..., \quad e \geq e^k \leftarrow \underline{w}_m^+$ be the root (input) events of a generalized dual NIN-AND gate g, where either $k = \eta$ or $k = 1$. Let $P(e \geq e^k \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+)$ be the probability of the leaf (output) event of g. Then, for $i = 1, ..., m$, we have*

$$P(e \geq e^k \leftarrow \underline{w}_1^+, ..., \underline{w}_m^+) > P(e \geq e^k \leftarrow \underline{w}_1^+).$$

Consider the above example. Using suitable generalized dual NIN-AND gates with all root events singular, we can derive the following by Eqn. (4),

$$P(e < e^1 | c_1^1, c_2^2) = 0.025,$$
$$P(e < e^2 | c_1^1, c_2^2) = 0.275,$$

from which we obtain the following by Proposition 1,

$$P(e^0 | c_1^1, c_2^2) = 0.025,$$
$$P(e^1 | c_1^1, c_2^2) = 0.25,$$
$$P(e^2 | c_1^1, c_2^2) = 0.725.$$

Reinforcement holds relative to $S_e = \{e^2\}$ because

$$P(e^2 | c_1^1, c_2^2) = 0.725 > 0.45 = P(e^2 | c_1^1, c_2^0)$$

and $0.725 > 0.5 = P(e^2 | c_1^0, c_2^2)$. It also holds relative to $S_e = \{e^1, e^2\}$ because

$$P(e \geq e^1 | c_1^1, c_2^2) = 0.975 > 0.75 = P(e \geq e^1 | c_1^1, c_2^0)$$

and $0.975 > 0.9 = P(e \geq e^1 | c_1^0, c_2^2)$. However, reinforcement does not hold relative to $S_e = \{e^1\}$ because

$$P(e^1 | c_1^1, c_2^2) = 0.25 < 0.3 = P(e^1 | c_1^1, c_2^0)$$

and $0.25 < 0.4 = P(e^1 | c_1^0, c_2^2)$.

In summary, a generalized direct NIN-AND gate expresses undermining and a generalized dual NIN-AND gate expresses reinforcement, relative to both the most intensive value and the collection of active values of the effect.

Generalized NIN-AND Trees

The following definition generalizes the binary NIN-AND tree models to multi-valued effect and causes.

Definition 7 *A* **generalized NIN-AND tree** *is a directed tree for a multi-valued effect e and a set $X = \{c_1, ..., c_n\}$ of multi-valued causes, parameterized by a* **boundary** *value e^k ($k > 0$) of e and an instantiation $\underline{x}^+ = \{c_1^{j_1}, ..., c_n^{j_n}\}$ of X, where $j_i > 0$ ($i = 1, ..., n$).*

1. *There are two types of nodes. An* **event** *node (a black oval) has an in-degree ≤ 1 and an out-degree ≤ 1. A* **gate** *node (a generalized NIN-AND gate) has an in-degree ≥ 2 and an out-degree 1.*
2. *There are two types of links, each connecting an event and a gate along the input-to-output direction of gates. A* **forward** *link (a line) is implicitly directed. A* **negation** *link (with a white oval at one end) is explicitly directed.*
3. *Each terminal node is an event labeled by a graded causal event $e \geq e^k \leftarrow \underline{y}^+$ or $e < e^k \leftarrow \underline{y}^+$. There is a single* **leaf** *(no child) where $\underline{y}^+ = \underline{x}^+$, and the gate it connects to is the* **leaf gate**. *For each* **root** *(no parent; indexed by i), $\underline{y}_i^+ \subset \underline{x}^+$, $\underline{y}_j^+ \cap \underline{y}_k^+ = \emptyset$ for $j \neq k$, and $\bigcup_i \underline{y}_i^+ = \underline{x}^+$.*
4. *Inputs to a gate g are in one of two cases:*
   (a) *Each is either connected by a forward link to a node labeled $e \geq e^k \leftarrow \underline{y}^+$, or by a negation link to a node labeled $e < e^k \leftarrow \underline{y}^+$. The output of g is connected by a forward link to a node labeled $e \geq e^k \leftarrow \cup_i \underline{y}_i^+$.*
   (b) *Each is either connected by a forward link to a node labeled $e < e^k \leftarrow \underline{y}^+$, or by a negation link to a node labeled $e \geq e^k \leftarrow \underline{y}^+$. The output of g is connected by a forward link to a node labeled $e < e^k \leftarrow \cup_i \underline{y}_i^+$.*

Fig. 5 is an example of a generalized NIN-AND tree for $C = \{c_1, c_2, c_3\}$ where $|D_e| = |D_1| = |D_2| = |D_3| = 3$.

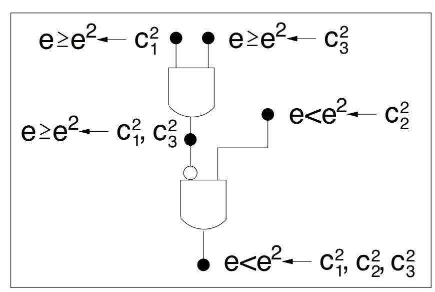

Figure 5: A generalized NIN-AND tree.

The probability of the leaf event of a generalized NIN-AND tree can be evaluated using Algorithm 1 GetCausalEventProb. From the model in Fig. 5, $P(e \geq e^2 \leftarrow c_1^2) = 0.85$, $P(e \geq e^2 \leftarrow c_2^2) = 0.8$, and $P(e \geq e^2 \leftarrow c_3^2) = 0.7$, it can be derived

$$P(e < e^2 \leftarrow c_1^2, c_2^2, c_3^2) = 0.081.$$

Using Proposition 1, the probabilities of root events in a generalized NIN-AND tree can be obtained from conditional probabilities that involve only a single active cause ($c_i \neq c_i^0$ and $c_j = c_j^0$ for $j \neq i$). After the probability of the leaf event is derived for each relevant graded causal event, the corresponding CPT can be obtained by applying Proposition 1 to the probabilities of the leaf events.

Properties of Generalized NIN-AND Trees

Theorem 1 establishes that generalized NIN-AND trees model both reinforcement and undermining correctly.

Theorem 1 *Let T be a generalized NIN-AND tree where the probability for each root node is specified in the range $(0, 1)$. Let $P(v)$ be returned by GetCausalEventProb(T).*

Then $P(v)$ combines the given probabilities according to reinforcement and undermining expressed by the topology of T, with each generalized direct NIN-AND gate corresponding to undermining and each generalized dual NIN-AND gate corresponding to reinforcement, relative to both $S_e = \{e^n\}$ and $S'_e = D_e \setminus \{e^0\}$.

Proof: $GetCausalEventProb$ evaluates first the output event for each gate node whose inputs are root events. If the root events are graded causal successes, then the gate is a generalized direct NIN-AND gate. By Propositions 2 and 3, the probability of the output event reflects the result of undermining, relative to both S_e and S'_e. Otherwise, the root events are graded causal failures, and the gate is a generalized dual NIN-AND gate. By Proposition 4, the probability of the output event reflects the result of reinforcement, relative to both S_e and S'_e.

After the evaluation, root nodes (and links incident to them) no longer participate in further evaluations and can be deleted. The remaining subtree is still a generalized NIN-AND tree with the depth reduced by one. $GetCausalEventProb$ repeats the above computation until the depth reduces to zero. The statement is true for the evaluation at each depth and hence the theorem holds. □

The following theorem establishes that specification of CPT using generalized NIN-AND trees is efficient.

Theorem 2 *Let $C = \{c_1, ..., c_n\}$ be the set of all causes of effect e that satisfy the graded success (failure) conjunction and independence. Denote $|D_e|$ by $\eta + 1$ and $|D_i|$ by $\beta + 1$ ($i = 1, ..., n$). Let P^* be a set of conditional probabilities*

$$P^* = \{P(e^k|c_1^0, ..., c_{i-1}^0, c_i^m, c_{i+1}^0, ..., c_n^0) \mid k > 0, m > 0\}.$$

Then, the following hold.

1. *The CPT $P(e|X)$ can be derived from P^* using generalized NIN-AND trees.*
2. *The complexity to specify P^* is $O(\eta\ (\beta_1 + ... + \beta_n))$.*

Proof: Other than $P(e^0|c_1^0, ..., c_n^0) = 1$, each other conditional probability in the CPT $P(e|X)$ can be derived from probabilities of output events of at most two generalized NIN-AND trees through Proposition 1. For each probability in P^*, the effect is present and exactly one cause is active. P^* contains all such probabilities. Hence, P^* is sufficient to specify the probabilities of input events of all relevant generalized NIN-AND trees through Proposition 1, and condition 1 follows.

Condition 2 amounts to a simple counting. □

Assuming $\eta = \beta_i$ for $i = 1, ..., n$, the above complexity becomes $O(n\ \eta^2)$ and is hence linear in n. A number of other important properties of generalized NIN-AND tree models are mentioned briefly below although their elaboration is beyond space limit.

Although the probability of each graded multi-causal event requires the use of a separate generalized NIN-AND tree, all of them can be derived from a single generalized NIN-AND tree. Hence, the complexity to specify the necessary tree topologies is also linear in n. Default conjunction and independence assumptions embedded in the model can be relaxed if necessary through additional numerical parameters than what is included in Theorem 2. By doing so, any CPT can be encoded through a generalized NIN-AND tree. Finally, generalized NIN-AND trees revert to binary NIN-AND trees when all variables are binary.

Conclusion

In this work, we generalize the binary NIN-AND tree causal models to multi-valued effect and causes. The generalized NIN-AND trees model explicitly reinforcement and undermining among causes relative to the most intensive level as well as the collection of active levels of effect. Specification of CPTs using generalized NIN-AND trees is shown to be efficient. Hence, this result will allow numerical parameters in Bayes nets to be specified efficiently through the intuitive concepts of reinforcement and undermining.

Acknowledgements

The author thanks comments from anonymous reviewers. Financial support from NSERC, Canada is acknowledged.

References

Diez, F. 1993. Parameter adjustment in Bayes networks: The generalized noisy or-gate. In Heckerman, D., and Mamdani, A., eds., *Proc. 9th Conf. on Uncertainty in Artificial Intelligence*, 99–105. Morgan Kaufmann.

Galan, S., and Diez, F. 2000. Modeling dynamic causal interaction with Bayesian networks: temporal noisy gates. In *Proc. 2nd Inter. Workshop on Causal Networks*, 1–5.

Heckerman, D., and Breese, J. 1996. Causal independence for probabilistic assessment and inference using Bayesian networks. *IEEE Trans. on System, Man and Cybernetics* 26(6):826–831.

Henrion, M. 1989. Some practical issues in constructing belief networks. In Kanal, L.; Levitt, T.; and Lemmer, J., eds., *Uncertainty in Artificial Intelligence 3*. Elsevier Science Publishers. 161–173.

Lemmer, J., and Gossink, D. 2004. Recursive noisy OR - a rule for estimating complex probabilistic interactions. *IEEE Trans. on System, Man and Cybernetics, Part B* 34(6):2252–2261.

Maaskant, P., and Druzdzel, M. 2008. An ici model for opposing influences. In Jaeger, M., and Nielsen, T., eds., *Proc. 4th European Workshop on Probabilistic Graphical Models*, 185–192.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.

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