International Journal of Approximate Reasoning 63 (2015) 32–47

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

An intercausal cancellation model for Bayesian-network engineering

Steven P.D. Woudenberg [a,*], Linda C. van der Gaag [a], Carin M.A. Rademaker [b]

[a] Department of Information and Computing Sciences, Utrecht University, The Netherlands
[b] Department of Clinical Pharmacy, University Medical Center, Utrecht University, The Netherlands

ARTICLE INFO

Article history:
Received 24 April 2014
Received in revised form 22 May 2015
Accepted 26 May 2015
Available online 5 June 2015

Keywords:
Bayesian networks
Knowledge modelling
Causal interaction
Intercausal cancellation

ABSTRACT

When constructing Bayesian networks with domain experts, network engineers often use the noisy-OR model, and causal interaction models more generally, to alleviate the burden of probability elicitation: the use of such a model serves to reduce the number of probabilities to be elicited on the one hand, and on the other hand forestalls experts having to give assessments for probabilities with compound conditions which they feel are hard to envision. Recently, we have shown that ill-considered use of the noisy-OR model specifically can substantially decrease a network's performance, especially in domains in which causal mechanisms include cancellation effects. Motivated by this observation, we designed a new causal interaction model, with the same engineering advantages as the noisy-OR model, to describe such effects. We detail properties of our intercausal cancellation model, and compare it against existing causal interaction models. We further illustrate the application of our model in the real-world domain of pharmacology.

1. Introduction

Bayesian networks of realistic size easily require thousands of probabilities for the quantification of their graphical structures. For many applications, these probabilities need to be assessed by domain experts. This assessment task often proves quite daunting, and is typically impeded by the experts feeling uncomfortable with providing numbers to describe their knowledge and experience [6,17]. To alleviate the burden on both the experts and the engineers involved, researchers have developed tailored elicitation techniques [7] and have designed causal interaction models to be used for the network quantification task [3,9,10,14,16,23].

A causal interaction model in essence is a parameterised conditional probability table for the effect variable of a causal mechanism in a Bayesian network. The best known among these models are the noisy-OR model [16] and its generalisations. For their parameters, these models require assessments for just a small number of conditional probabilities for the effect variable of a mechanism. All other probabilities required for a fully specified conditional probability table are calculated from these parameters through simple mathematical functions that assume a specific pattern of interaction among the causes of the effect. The use of the model thereby incurs a substantial reduction of the number of probabilities to be assessed explicitly. In addition, it serves to calculate the probabilities with compound conditions which experts typically find hardest to assess.

* Corresponding author.
E-mail addresses: S.P.D.Woudenberg@uu.nl (S.P.D. Woudenberg), L.C.vanderGaag@uu.nl (L.C. van der Gaag).

http://dx.doi.org/10.1016/j.ijar.2015.05.011
0888-613X/

Over the years, causal interaction models, and the noisy-OR model specifically, have gained considerable popularity among network engineers. Evidence has built up that these models can be used for many real-world applications of Bayesian networks, even if their underlying assumptions are not fully met in reality [1,5]. In our recent studies however, we demonstrated that ill-considered use of the noisy-OR model can result in poorly calibrated probability values [19,21]. Especially for causal mechanisms involving cancellation effects among their cause variables, will the noisy-OR calculated probabilities deviate substantially from the true probabilities. We encountered such cancellation effects upon describing the interactions among pharmaceutical substances in our application domain of pharmacology. Using noisy-OR calculated probabilities in this domain would result in these cancellation effects not being properly modelled and possibly in erroneous conclusions being drawn. In this paper, we develop an intercausal cancellation model which provides a parameterised conditional probability table for the effect variable of a causal mechanism embedding cancellation effects among its cause variables. We distinguish between different types of cancellation, ranging from full to partial and from one-sided to mutual cancellation, motivated by the examples encountered in our domain of pharmacology, and show that our model embeds the various types of cancellation discerned. Our cancellation model is designed from first principles just like the noisy-OR model and shares the same engineering advantages. In our application domain, the intercausal cancellation model served to accommodate easily-accessible knowledge about patterns of interaction among multiple pharmaceutical substances and thereby proved to substantially alleviate the quantification task for our application. We compared the results obtained from our model against those found with related variants of the noisy-OR model, and concluded that for capturing the patterns of cancellation in our application domain, our model showed considerable engineering advantages over these existing models.

The paper is organised as follows. In Section 2, we introduce our notational conventions and review the principles underlying the noisy-OR model. We elaborate on our motivation for defining an intercausal cancellation model in Section 3. Our basic model with its underlying assumptions and properties is defined in Section 4, and some extensions are discussed in Section 5. In Section 6, we demonstrate application of our model to two real-world examples from our domain of pharmacology. We compare the intercausal cancellation model against existing models in Section 7. The paper is concluded in Section 8.

2. Preliminaries

In this section we briefly review Bayesian networks in general and thereby introduce our notational conventions. We further describe the noisy-OR model and its underlying principles in detail.

2.1. Bayesian networks in general

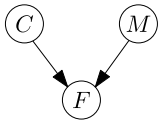
We consider a joint probability distribution Pr over a set of binary random variables. The variables are denoted by capital letters, and each variable V has the values $\bar{v}$ and v to describe the absence and presence respectively, of some concept. (Sub)sets of variables are denoted by bold-face capital letters, and joint value combinations for sets of variables are written in bold-face small letters.

We assume that the probability distribution Pr is represented by a Bayesian network. This representation includes a directed acyclic graph in which the nodes capture the variables involved and in which the arcs describe the dependency relation among them. The arcs of the graph are often looked upon as modelling a causal relationship between the connected variables. Although we will not make any claims with respect to a causal interpretation, we will adopt the terminology involved: if the graph includes an arc pointing from a variable C to a variable E, we say that C is a cause variable for E, and E is an effect variable of C. Within the graph, we now consider a causal mechanism consisting of n cause variables $\mathbf{C} = \{C_1, \ldots, C_n\}, n \geq 0$, and a single effect variable E. The strengths of the probabilistic influences of the cause variables on the effect variable E are expressed by means of a conditional probability table $\Pr(E \mid \mathbf{c})$ for E, given all possible joint value combinations $\mathbf{c}$ for the set of cause variables $\mathbf{C}$; we note that the table specifies a total of 2^n conditional probability distributions for E.

A Bayesian network describes a unique joint probability distribution and hence provides for computing any prior or posterior probability of interest over its variables. The problem of exact probabilistic inference with a Bayesian network has been proven to be NP-hard [2]; more recent research showed #P-hardness for the problem [18]. For networks of which the graph has bounded treewidth however, exact inference can be solved in polynomial time, that is, polynomial in the size of the network. Various algorithms have been designed for this purpose, among which the junction-tree propagation algorithm [11,13] is the most efficient to date.

2.2. The noisy-OR model

When constructing a Bayesian network for a real-world problem, the first step is to configure a directed graph to describe the probabilistic dependencies between the relevant variables. Subsequently, conditional probability tables are to be specified for all variables, that is, for the effect variables of all causal mechanisms in the graph. For each effect variable, such a table details an exponential number of conditional probability distributions, that is, exponential in the number of cause variables involved. Experience shows that obtaining all required numbers is a daunting and time-consuming task, especially



Cold (C)	Malaria (M)	Fever (F)	
		yes	no
yes	yes	0.96	0.04
yes	no	**0.60**	0.40
no	yes	**0.90**	0.10
no	no	0.00	1.00

Fig. 1. An example causal mechanism describing having a fever (modelled by the variable F) as an effect of suffering from a cold (C) and/or malaria (M) (*left*), and the noisy-OR established probability table for the effect variable F (*right*).

if these numbers are to be elicited from domain experts. The noisy-OR model was the first model been designed to alleviate this burden for mechanisms with binary variables [10,16].

The noisy-OR model essentially is a parameterised conditional probability table for the effect variable E of a causal mechanism. It captures the information that each modelled cause in essence suffices to give rise to the effect e, but that some implicit processes may inhibit the effect to occur; it can thus be looked upon as modelling a logical OR with uncertain perturbation effects. The parameters of the noisy-OR model are

$$\Pr(e \mid \bar{c}_1, \ldots, \bar{c}_{i-1}, c_i, \bar{c}_{i+1}, \ldots, \bar{c}_n)$$

for $i = 1, \ldots, n$, of the effect arising in the presence of just a single cause; these parameter probabilities express that in the presence of cause c_i, the occurrence of the effect e may be inhibited by some unmodelled process, with probability $1 - \Pr(e \mid \bar{c}_1, \ldots, \bar{c}_{i-1}, c_i, \bar{c}_{i+1}, \ldots, \bar{c}_n)$. The probability $\Pr(e \mid \bar{c}_1, \ldots, \bar{c}_n)$ of the effect e arising if none of its causes are present, is taken to be zero by the model. To arrive at a fully specified probability table for the variable E, the conditional probability of the effect e arising in the presence of a specific combination $\mathbf{c}$ of multiple causes is calculated by the model as

$$\Pr(e \mid \mathbf{c}) = 1 - \prod_{j \in J} (1 - \Pr(e \mid \bar{c}_1, \ldots, \bar{c}_{j-1}, c_j, \bar{c}_{j+1}, \ldots, \bar{c}_n))$$

where J is the set of indices of the cause variables C_j which are marked as being present in the value combination $\mathbf{c}$.

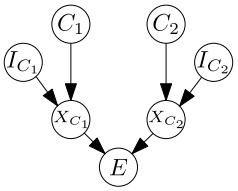
Example 2.1. Fig. 1 depicts, on the left, the graph of a (highly simplified) causal mechanism. The mechanism describes having a cold (variable C) and having malaria (M) as the possible causes of having a fever (F). On the right, the figure shows the conditional probability table that results from applying the noisy-OR model for the effect variable F of the mechanism. The probabilities printed in bold are the two parameter probabilities of the table; in a real application these probabilities would have to be assessed by experts or learned from data. The probability $\Pr(f \mid \bar{c}, \bar{m})$ further is set to zero and the remaining probability $\Pr(f \mid c, m)$ is calculated by the noisy-OR model from the two parameter probabilities. We note that the probability of a fever occurring in a person who suffers from both a cold and malaria, is calculated to be higher than the probabilities of the fever to arise in persons suffering from either one of the diseases.

Underlying the noisy-OR model are the properties of accountability and exception independence. The property of accountability states that the effect e in a causal mechanism cannot occur if all modelled causes are absent; it thus states that $\Pr(e \mid \bar{c}_1, \ldots, \bar{c}_n) = 0$. The property of exception independence pertains to the exception mechanisms which are left implicit in the causal mechanism. An exception mechanism is an unmodelled process which governs the inhibition of an essentially deterministic causal relation; we say that a causal relation which asserts that the effect should arise in the presence of the associated cause, is inhibited if the effect does *not* occur despite the cause's presence. The property of exception independence now states that the exception mechanisms underlying such inhibitions are mutually independent apriori. To demonstrate how the calculation rule of the noisy-OR model is derived from the two properties of accountability and exception independence, we construct an auxiliary mechanism in which the property of exception independence has been made explicit [16]; the graph of this mechanism is depicted in Fig. 2. In addition to the effect and cause variables, the mechanism includes the variables X_{C_i} to describe the deterministic processes through which the causes c_i achieve the effect e, and the variables I_{C_i} to represent the inhibiting exception mechanisms. To the right of the mechanism, the figure shows the conditional probability tables for the variables I_{C_i} and X_{C_i}, and the associated table for the effect variable. We note that the probability table for E is deterministic, as are the tables for the variables X_{C_i}; the uncertainty of the effect to arise in the presence of a particular cause is attributed entirely to the inhibiting exception mechanism for this cause. In the sequel, we will use the phrase *conceptual mechanism* to refer to the mechanism from Fig. 2; we would like to note that the conceptual mechanism serves as a proof construct only.

From the conceptual mechanism described above, the conditional probability distributions for the effect variable E given all possible joint value combinations for the two cause variables C_1 and C_2, are now readily established:

$$\Pr(e \mid \bar{c}_1, \bar{c}_2) = 0$$

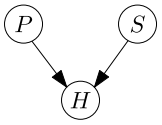
$$\Pr(e \mid c_1, \bar{c}_2) = 1 - \Pr(i_{C_1})$$



$\Pr(i_{C_1}) = 1 - \Pr(e \mid c_1, \bar{c}_2)$ $\quad$ $\Pr(i_{C_2}) = 1 - \Pr(e \mid \bar{c}_1, c_2)$

$\Pr(x_{C_1} \mid c_1, i_1) = 0$	$\Pr(x_{C_2} \mid c_2, i_2) = 0$	$\Pr(e \mid x_{C_1}, x_{C_2}) = 1$
$\Pr(x_{C_1} \mid c_1, \bar{i}_1) = 1$	$\Pr(x_{C_2} \mid c_2, \bar{i}_2) = 1$	$\Pr(e \mid \bar{x}_{C_1}, x_{C_2}) = 1$
$\Pr(x_{C_1} \mid \bar{c}_1, i_1) = 0$	$\Pr(x_{C_2} \mid \bar{c}_2, i_2) = 0$	$\Pr(e \mid x_{C_1}, \bar{x}_{C_2}) = 1$
$\Pr(x_{C_1} \mid \bar{c}_1, \bar{i}_1) = 0$	$\Pr(x_{C_2} \mid \bar{c}_2, \bar{i}_2) = 0$	$\Pr(e \mid \bar{x}_{C_1}, \bar{x}_{C_2}) = 0$

Fig. 2. The conceptual mechanism with the property of exception independence made explicit: process and inhibitor variables are included in the graph (*left*) and incorporated in the associated probability distribution (*right*).



Pain medication (P)	Sleep (S)	Headache (H)	
		yes	no
yes	yes	0.15	0.85
yes	no	0.50	**0.50**
no	yes	0.30	**0.70**
no	no	1.00	0.00

Fig. 3. A causal mechanism which describes having a headache (modelled by the variable H), and taking pain medication (P) and sleeping (S) as its possible remedies (*left*), along with the inversed noisy-OR established probability table for the effect variable H (*right*).

$$\Pr(e \mid \bar{c}_1, c_2) = 1 - \Pr(i_{C_2})$$
$$\Pr(e \mid c_1, c_2) = 1 - \Pr(i_{C_1}) + 1 - \Pr(i_{C_2}) - (1 - \Pr(i_{C_1})) \cdot (1 - \Pr(i_{C_2}))$$

We observe that these distributions constitute a noisy-OR calculated conditional probability table for E: the probabilities $\Pr(e \mid c_1, \bar{c}_2) = 1 - \Pr(i_{C_1})$ and $\Pr(e \mid \bar{c}_1, c_2) = 1 - \Pr(i_{C_2})$ are the two parameter probabilities of the model, and the probability $\Pr(e \mid c_1, c_2)$ equals $1 - (1 - \Pr(e \mid c_1, \bar{c}_2)) \cdot (1 - \Pr(e \mid \bar{c}_1, c_2))$ as prescribed by the model. We would like to note that the conceptual mechanism is used only for defining the noisy-OR model and is never used in practice.

To conclude, we observe that the noisy-OR model as described above is tailored to causal mechanisms in which the effect is absent apriori. The model is readily inverted however, to apply to causal mechanisms in which the effect is present apriori and is annihilated by the various causes. With the parameter probabilities $\Pr(\bar{e} \mid \bar{c}_1, \ldots, \bar{c}_{j-1}, c_j, \bar{c}_{j+1}, \ldots, \bar{c}_n)$ for $j = 1, \ldots, n$, the model then computes

$$\Pr(\bar{e} \mid \mathbf{c}) = 1 - \prod_{j \in J} \left(1 - \Pr(\bar{e} \mid \bar{c}_1, \ldots, \bar{c}_{j-1}, c_j, \bar{c}_{j+1}, \ldots, \bar{c}_n)\right)$$

where J again is the set of indices of the cause variables C_j which are marked as being present in the value combination $\mathbf{c}$.

Example 2.2. Fig. 3 depicts a (highly simplified) causal mechanism for the context of a headache clinic. The mechanism represents having a headache (variable H), and describes taking pain medication (P) and sleeping (S) as possible remedies. In the absence of any remedy, the probability of having a headache is set to $\Pr(h \mid \bar{p}, \bar{s}) = 1.0$ by the property of accountability of the inverted noisy-OR model. The probabilities $\Pr(\bar{h} \mid p, \bar{s})$ and $\Pr(\bar{h} \mid \bar{p}, s)$ constitute the model's parameter probabilities. The probability $\Pr(\bar{h} \mid p, s)$ of the headache having disappeared after taking pain medication and sleeping, is calculated by the inverted noisy-OR model from the two parameter probabilities as prescribed. We observe that the combined presence of the two remedies results in a stronger influence on the effect disappearing when compared to the influence of just a single remedy, as also seen with the more standard use of the noisy-OR model.

In the sequel we will refer to the use of the noisy-OR model without explicitly mentioning whether the standard or the inverted model is meant, as long as ambiguity cannot occur.

3. Motivation

Causal interaction models, have met with wide-spread use among Bayesian-network engineers because of their clear advantages for the quantification task. Experiences with the noisy-OR model specifically have shown that Bayesian networks with model-calculated probability tables perform comparably to expert-quantified ones, even when the domain knowledge does not match the patterns of intercausal interaction imposed by the model. These experiences in fact have led to the suggestion that Bayesian networks are quite robust against poorly calibrated values induced in their conditional probability tables by the noisy-OR model, and that the model's use might often be warranted for mere pragmatic reasons [1,5].

In a recent study, we investigated the effects of using the noisy-OR model in further detail [19,21]. More specifically, we employed sensitivity analysis techniques to study the shifts in output probability that are occasioned by changing the true probabilities for the effect variable of a causal mechanism to noisy-OR computed ones. For an example causal mechanism with two cause variables C_1 and C_2 and an effect variable E, Fig. 4 depicts the output probability of interest $\Pr(c_1 \mid c_2, e)$ as

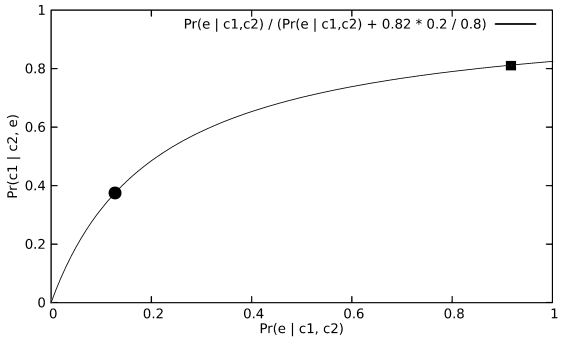


Fig. 4. The sensitivity function describing the output probability $\Pr(c_1 \mid c_2, e)$ in terms of the probability $\Pr(e \mid c_1, c_2)$; the dot indicates the output probability given the true value of $\Pr(e \mid c_1, c_2)$, and the square marks the output given the noisy-OR calculated value for $\Pr(e \mid c_1, c_2)$.

a function of the probability $\Pr(e \mid c_1, c_2)$, given fixed values for $\Pr(e \mid c_1, \bar{c}_2)$ and $\Pr(e \mid \bar{c}_1, c_2)$. With the depicted function, the square indicates the output probability given the noisy-OR calculated value for $\Pr(e \mid c_1, c_2)$ and the dot marks the output probability given the true value of $\Pr(e \mid c_1, c_2)$. The figure shows that the two values established for the output probability $\Pr(c_1 \mid c_2, e)$ differ substantially, which could easily result in erroneous conclusions being drawn.

The difference between the two values calculated for the output probability described above originates from a cancellation effect embedded in the true probabilities of the mechanism. While the presence of each cause separately is modelled as increasing the probability of e, the true probability of e given the simultaneous presence of both causes shows that the separate influences are cancelled out to quite some extent:

$$\Pr(e \mid c_1, \bar{c}_2) = 0.150 \qquad \Pr(e \mid \bar{c}_1, c_2) = 0.850$$
$$\Pr(e \mid \bar{c}_1, \bar{c}_2) = 0.000 \qquad \Pr(e \mid c_1, c_2) = 0.075$$

The use of the noisy-OR model for the mechanism however, assumes that the simultaneous presence of both causes results in an increased influence on the effect and establishes the probability $\Pr(e \mid c_1, c_2)$ to be equal to 0.8725.

Upon describing the interaction effects among pharmaceutical substances in the real-world domain of pharmacology, we encountered various patterns of cancellation. The presence of these cancellation patterns forestalled the use of the (inverted) noisy-OR model for specifying the conditional probability tables involved: while the effects of two substances upon concurrent intake would mutually decrease in reality, the noisy-OR model would indicate an increased overall effect, which would result in erroneous reasoning patterns upon propagation. Since providing conditional probabilities for an effect in view of a combination of substances proved nearly impossible for our domain expert, we decided that it would be highly advantageous for our application to have available a causal interaction model to help describe cancellation effects. We expect that such a model will in fact find wider use in a range of biomedical, chemical and environmental domains.

Since the introduction of the noisy-OR model, numerous variants and generalisations have been developed. Two well-known and commonly used generalisations are the leaky noisy-OR model [3,10], which does not require the accountability assumption to hold, and the noisy-MAX model [3,10], which applies to causal mechanisms involving non-binary variables. Also, variants of the noisy-OR model have been developed which are capable of describing negative causal influences and/or negative causal interactions. Examples include the DeMorgan model [15], the XOR model, the (inhibited) recursive noisy-OR model [8,12] and the NIN-AND tree [23,24], of which the DeMorgan model is a special case. While their existence clearly illustrates the demand for models capturing intercausal cancellation, we found that none of these models covered the range of cancellation patterns that we required for our application. We will return to this observation in Section 7.

4. The intercausal cancellation model

To describe possible patterns of cancellation among the causes of a common effect, we define a new causal interaction model with the same engineering advantages as the noisy-OR model. In this section we develop the basic cancellation model with two cause variables; in Section 5 we will discuss extensions of this basic model, among which is the extension to multiple cause variables.

4.1. A conceptual mechanism describing cancellation

We begin by developing a conceptual mechanism for describing patterns of cancellation from first principles. We recall that the conceptual mechanism defining the noisy-OR model associates with each cause variable C_i an inhibitor variable I_{C_i} to describe the uncertainty of the causal effect; the model's rule for calculating the joint effect of multiple simultaneous causes then builds upon an independence assumption for these inhibitor variables. For the design of our intercausal cancellation model, we exploit the same idea and develop a conceptual mechanism with an additional, explicitly modelled, annihilator variable for capturing the cancellation effects among the various causes; this conceptual mechanism again serves

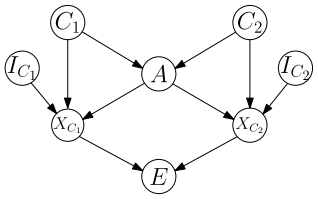


Fig. 5. The graph of the conceptual mechanism defining the intercausal cancellation model with two cause variables and with their inhibitor variables made explicit.

as a proof construct only. For ease of exposition, we focus on mechanisms with just the two cause variables C_1 and C_2 and the common effect variable E.

The graph of the conceptual mechanism defining our intercausal cancellation model is depicted in Fig. 5. As before, the two variables X_{C_1} and X_{C_2} represent the processes through which the causes c_1 and c_2 achieve their common effect e. The two inhibitor variables I_{C_1} and I_{C_2} again capture the exception mechanisms that may inhibit the causal influences. The new variable A serves to describe the cancellation effect among the two causes, and will be called the annihilator variable for C_1 and C_2. From the graph we read that the inhibitor variables are mutually independent apriori, as with the conceptual mechanism for the noisy-OR model. The process variables X_{C_i} however, have become dependent of both cause variables in the mechanism through the variable A.

The conditional probability tables for the effect variable E and for the annihilator variable A are defined as

$$\begin{aligned} &\Pr(e \mid \bar{x}_{C_1}, \bar{x}_{C_2}) = 0 \qquad \Pr(a \mid \bar{c}_1, \bar{c}_2) = 0 \\ &\Pr(e \mid x_{C_1}, \bar{x}_{C_2}) = 1 \qquad \Pr(a \mid c_1, \bar{c}_2) = 0 \\ &\Pr(e \mid \bar{x}_{C_1}, x_{C_2}) = 1 \qquad \Pr(a \mid \bar{c}_1, c_2) = 0 \\ &\Pr(e \mid x_{C_1}, x_{C_2}) = 1 \qquad \Pr(a \mid c_1, c_2) = 1 \end{aligned}$$

The conditional probability table for E is the same as with the noisy-OR model, and shows that each of the processes x_{C_1} and x_{C_2} suffices to cause the effect e, which is absent apriori. The probability table for the variable A expresses that the annihilator is activated only if both causes are present simultaneously.

Before specifying the conditional probability tables for the process variables X_{C_i}, we recall that a causal process may be inhibited despite the presence of its associated cause; the inhibitor variables I_{C_i} were introduced to capture such inhibition in the conceptual mechanism for the noisy-OR model. We further recall that upon practical application of the model, the inhibitors' probabilities are effectively absorbed into the conditional probability table of the effect variable. We now take a similar approach for the conceptual mechanism for our cancellation model and absorb the probabilities of the inhibitor variables I_{C_i} into the associated process variables X_{C_i}, which results in the same model as depicted in Fig. 5 yet without the variables I_{C_i} and I_{C_j}. The conditional probability tables for the process variables X_{C_1} and X_{C_2} now become

$$\begin{aligned} &\Pr(x_{C_1} \mid \bar{c}_1, \bar{a}) = 0 \qquad \Pr(x_{C_2} \mid \bar{c}_2, \bar{a}) = 0 \\ &\Pr(x_{C_1} \mid \bar{c}_1, a) = 0 \qquad \Pr(x_{C_2} \mid \bar{c}_2, a) = 0 \\ &\Pr(x_{C_1} \mid c_1, \bar{a}) = p_1 \qquad \Pr(x_{C_2} \mid c_2, \bar{a}) = p_2 \\ &\Pr(x_{C_1} \mid c_1, a) \leq p_1 \qquad \Pr(x_{C_2} \mid c_2, a) \leq p_2 \end{aligned}$$

In these tables, the conditional probabilities $\Pr(x_{C_i} \mid c_i, \bar{a})$ capture, for each cause separately, the probability that the process x_{C_i} is evoked by the presence of its associated cause; the numerical values p_1 and p_2 thus incorporate the probabilities of inhibition as described above. These two probabilities are parameter probabilities of our intercausal cancellation model, just as with the noisy-OR model.

In the tables for the process variables X_{C_i}, we have not yet specified concrete numerical values for the remaining probabilities $\Pr(x_{C_i} \mid c_i, a)$. We note that these probabilities now incorporate not just the inhibition probability but also the degree to which the cause's influence on the effect is cancelled out by the presence of the other cause. The restrictions indicated for these probabilities are designed to reflect the meanings of the process and annihilator variables: in the presence of the annihilator, the causal processes x_{C_1} and x_{C_2} are less likely to be evoked than in its absence. To arrive at fully specified probability tables, we now set

$$\begin{aligned} &\Pr(x_{C_1} \mid c_1, a) = p_1 - \alpha_{1|12} \cdot p_1 \\ &\Pr(x_{C_2} \mid c_2, a) = p_2 - \alpha_{2|12} \cdot p_2 \end{aligned}$$

with $0 \leq \alpha_{1|12}, \alpha_{2|12} \leq 1$. The cancellation parameters $\alpha_{1|12}$ and $\alpha_{2|12}$ capture the degree of cancellation among the two causes, where $\alpha_{1|12}$ represents the degree to which the effect of cause c_1 is cancelled out by the simultaneous presence of c_2; $\alpha_{2|12}$ has a similar interpretation. For the parameter $\alpha_{1|12}$, we find that $\alpha_{1|12} = 0$ implies $\Pr(x_{C_1} \mid c_1, a) = p_1$, which expresses that the influence of the cause c_1 on e is not cancelled out at all by the simultaneous presence of the cause c_2;

$\alpha_{1|12} = 1$, and hence $\Pr(x_{C_1} \mid c_1, a) = 0$, expresses full cancellation of the effect of c_1. Any choice with $0 < \alpha_{1|12} < 1$ captures a partial cancellation effect: the closer $\alpha_{1|12}$ is to 1, the stronger the cancellation due to the presence of c_2 on the effect of c_1 will be. We would like to note that in the absence of any cancellation, that is, when $\alpha_{1|12} = \alpha_{2|12} = 0$, our cancellation model reduces to the traditional noisy-OR model. In addition to full and partial cancellation, also one-sided and mutual cancellation are readily modelled through the cancellation parameters. For example, the combination of parameter values $\alpha_{1|12} = 1$ and $\alpha_{2|12} = 0$ describes a full one-sided cancellation of the effect of cause c_1, while the effect of cause c_2 on e is unaffected.

From the conceptual mechanism developed above, we now derive the conditional probability distributions for the effect variable E given all possible value combinations for the two cause variables C_1 and C_2. By building upon properties of independence, applying well-known rules of probability, and using the conditional probability tables described above, we find that

$$\begin{aligned}\Pr(\bar{e} \mid c_1, c_2) &= \Pr(\bar{e} \mid \bar{x}_{C_1}, \bar{x}_{C_2}) \cdot \Pr(\bar{x}_{C_1} \mid c_1, a) \cdot \Pr(\bar{x}_{C_2} \mid c_2, a) \cdot \Pr(a \mid c_1, c_2) \\ &= \left(1 - p_1 + \alpha_{1|12} \cdot p_1\right) \cdot \left(1 - p_2 + \alpha_{2|12} \cdot p_2\right)\end{aligned}$$

from which we establish that

$$\begin{aligned}\Pr(e \mid c_1, c_2) &= 1 - \left(1 - p_1 + \alpha_{1|12} \cdot p_1\right) \cdot \left(1 - p_2 + \alpha_{2|12} \cdot p_2\right) \\ \Pr(e \mid c_1, \bar{c}_2) &= p_1 \\ \Pr(e \mid \bar{c}_1, c_2) &= p_2 \\ \Pr(e \mid \bar{c}_1, \bar{c}_2) &= 0\end{aligned}$$

We would like to note that, as with the conceptual mechanism for the noisy-OR model, the mechanism detailed above is used only for defining our intercausal cancellation model and is not used in practice. Upon practical application, just the parameterised conditional probability table $\Pr(E \mid C_1, C_2)$ for the effect variable E is used. We note that for a full specification of this table, values need to be obtained for the parameter probabilities p_1 and p_2 involved, and for the cancellation parameters $\alpha_{1|12}$ and $\alpha_{2|12}$.

4.2. On the cancellation parameters

Upon applying the intercausal cancellation model defined above, a Bayesian-network engineer has to choose appropriate values for the cancellation parameters involved. When no cancellation effects are expected, all cancellation parameters are set to zero; the model then further requires the same parameters as the noisy-OR model. When intercausal cancellation effects do occur in an application domain, modelling these effects requires non-zero values for one or more of the regulatory cancellation parameters. As these parameters generally have little meaning to the experts involved and therefore cannot be elicited directly, an in-depth discussion of the processes induced by the various causes is required to obtain sufficient insight in the effects to be captured. If the discussion with the domain experts reveals that causes exhibit (mutual or one-sided) full cancellation of their influences, then the choice of parameter values is readily made. With respect to partial cancellation, our experiences in the field of pharmacology revealed that the associated regulatory parameters were easier to estimate with our domain expert than the actual compound conditional probabilities for the effect variable. In fact, probabilities of the effect arising in the presence of rare combinations of causes often were nearly impossible for her to give, while she could fairly easily explain the underlying causal mechanisms and the cancellation involved. We will return to these experiences in Section 6.

To support practical application of our cancellation model, we propose a maximum-entropy choice of value for the cancellation parameters, designed specifically for describing (mutual or one-sided) partial cancellation; this maximum-entropy parameter value equals $\frac{1}{2}$. For describing a two-sided partial cancellation using this choice of value, we would thus take $\alpha_{1|12} = \alpha_{2|12} = \frac{1}{2}$ and find

$$\begin{aligned}\Pr(e \mid c_1, c_2) &= 1 - \left(1 - p_1 + \alpha_{1|12} \cdot p_1\right) \cdot \left(1 - p_2 + \alpha_{2|12} \cdot p_2\right) \\ &= 1 - \left(1 - p_1 + \tfrac{1}{2} \cdot p_1\right) \cdot \left(1 - p_2 + \tfrac{1}{2} \cdot p_2\right) \\ &= \tfrac{1}{2} \cdot \left(p_1 + p_2 - \tfrac{1}{2} \cdot p_1 \cdot p_2\right)\end{aligned}$$

from which the conditional probability table for the effect variable E of our two-cause mechanism is established to be

$$\begin{aligned}\Pr(e \mid c_1, c_2) &= \tfrac{1}{2} \cdot \left(p_1 + p_2 - \tfrac{1}{2} \cdot p_1 \cdot p_2\right) \\ \Pr(e \mid c_1, \bar{c}_2) &= p_1 \\ \Pr(e \mid \bar{c}_1, c_2) &= p_2 \\ \Pr(e \mid \bar{c}_1, \bar{c}_2) &= 0\end{aligned}$$

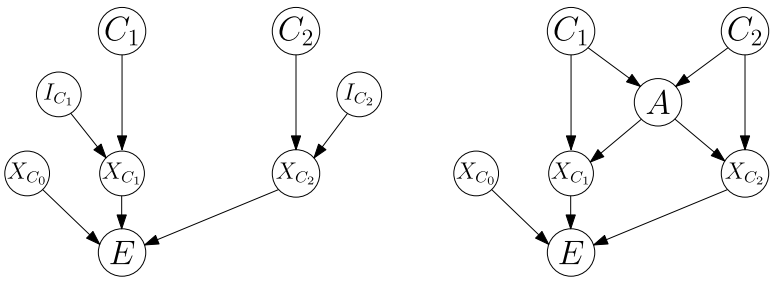


Fig. 6. The graph of the conceptual mechanism defining the leaky noisy noisy-OR model (*left*) and the intercausal cancellation representation of the leaky noisy-OR model (*right*).

where $p_1 = \Pr(e \mid c_1, \bar{c}_2)$ and $p_2 = \Pr(e \mid \bar{c}_1, c_2)$ are the remaining parameter probabilities of the model. We observe that using the maximum-entropy choice of value for the two cancellation parameters will result in a conditional probability $\Pr(e \mid c_1, c_2)$ within the interval $[0; \frac{3}{4}]$.

5. Extensions of the cancellation model

For wider applicability of the intercausal cancellation model, we present two extensions of the basic model introduced above. In Section 5.1 we extend the model to apply to causal mechanisms in which the common effect can occur spontaneously, and in Section 5.2 we unfold the extension of our cancellation model to render it applicable to mechanisms with more than two cause variables.

5.1. The leaky intercausal cancellation model

A frequently used extension of the noisy-OR model is the leaky noisy-OR model [3,10] which applies to causal mechanisms for which the property of exception independence holds, yet for which accountability does not hold, that is, for which $\Pr(e \mid \bar{c}_1, \ldots, \bar{c}_n) > 0$. The so-called leak probability $\Pr(e \mid \bar{c}_1, \ldots, \bar{c}_n)$ can in essence be represented using the standard noisy-OR model by including an additional cause variable C_0 in the causal mechanism under study; this variable is then fixed at always being present. In the leaky noisy-OR model however, no explicit cause variable C_0 is used. Instead, in the conceptual mechanism underlying this model, the effect of the additional variable is absorbed into its associated process variable X_{C_0}, as indicated in Fig. 6 on the left. The leaky noisy-OR model has the leak probability as an extra parameter in addition to the n parameter probabilities from the noisy-OR model, and computes the probabilities $\Pr(e \mid \mathbf{c})$ for arbitrary value combinations $\mathbf{c}$ for the cause variables through

$$\Pr(e \mid \mathbf{c}) = 1 - (1 - \Pr(e \mid \bar{c}_1, \ldots, \bar{c}_n)) \cdot \prod_{j \in J} (1 - \Pr(e \mid \bar{c}_1, \ldots, \bar{c}_{j-1}, c_j, \bar{c}_{j+1}, \ldots, \bar{c}_n))$$

where J, as before, is the set of indices of the cause variables C_j which are marked as being present in $\mathbf{c}$. We note that, although two approaches of the leaky noisy-OR model are in use which differ in the inclusion of the leak probability in the parameter probabilities, we have chosen to restrict the discussion to the above variant; our observations apply to both approaches alike however.

Incorporating a leak probability into our intercausal cancellation model can be achieved in a similar way as with the noisy-OR model. In essence, an additional cause variable C_0 is assumed, the effect of which is again absorbed into its associated process variable. As the additional cause is always present, we assume that any cancellation effects it may have with the explicitly modelled causes are already embedded in the estimated parameter probabilities; we assume moreover that it has no further cancellation effects in the presence of more than one of these causes. The graph of the conceptual mechanism which now underlies our leaky intercausal cancellation model is depicted in Fig. 6 on the right. From its parameters, the leaky cancellation model with the two cause variables C_1 and C_2 establishes the probability $\Pr(e \mid c_1, c_2)$ to be

$$\Pr(e \mid c_1, c_2) = 1 - (1 - \Pr(e \mid \bar{c}_1, \bar{c}_2)) \cdot (1 - p_2 + \alpha_{1|12} \cdot p_2) \cdot (1 - p_1 + \alpha_{2|12} \cdot p_1)$$

where the cancellation parameters now capture the degree of cancellation of a cause's process in the presence of the other explicit cause and the background presence of the implicit cause.

5.2. Multiple cause variables

So far, we considered causal mechanisms with two binary cause variables only. We now extend our intercausal cancellation model to take multiple cause variables into consideration. To this end, we consider a mechanism with the three binary cause variables C_1, C_2 and C_3, the presence of each of which can cause the effect e to arise; the extension of the model to

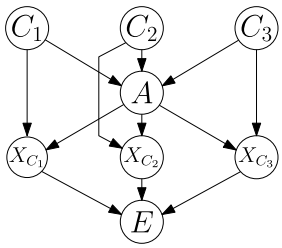


Fig. 7. The three-cause conceptual mechanism for the intercausal cancellation model.

more than three binary cause variables follows straightforwardly. Fig. 7 shows the conceptual mechanism which is used as a proof construct for deriving the conditional probability table of the effect variable E.

The three-cause conceptual mechanism includes a single annihilator variable A which, as before, serves to indicate which cancellation effects occur for which combination of causes. While with two cause variables just a single possible cancellation effect had to be captured, with three cause variables a total of four cancellation effects have to be indicated, that is, one for every combination of two or more simultaneously present causes. The annihilator variable therefore can no longer be binary, but has to take five values to model the cancellation effects. The conditional probability table capturing the meaning of the annihilator variable thus includes 40 probabilities; for space-saving reasons, we specify just the probabilities which are equal to 1:

$$\begin{aligned}
&\Pr(A=0 \mid \bar{c}_1, \bar{c}_2, \bar{c}_3) = 1 \qquad \Pr(A=1 \mid \bar{c}_1, c_2, c_3) = 1\\
&\Pr(A=0 \mid \bar{c}_1, \bar{c}_2, c_3) = 1 \qquad \Pr(A=2 \mid c_1, \bar{c}_2, c_3) = 1\\
&\Pr(A=0 \mid \bar{c}_1, c_2, \bar{c}_3) = 1 \qquad \Pr(A=3 \mid c_1, c_2, \bar{c}_3) = 1\\
&\Pr(A=0 \mid c_1, \bar{c}_2, \bar{c}_3) = 1 \qquad \Pr(A=4 \mid c_1, c_2, c_3) = 1
\end{aligned}$$

where the value $A=0$ indicates the absence of a cancellation effect, and the values $1,\ldots,4$ of the annihilator variable indicate the cancellation effect induced by a specific combination of causes. The remaining entries of the conditional probability table for the variable A are equal to zero.

We now turn to the conditional probability tables for the intermediate process variables, and address the table for the variable X_{C_1}; the conditional probability tables for the other process variables are constructed analogously. As before, if the cause modelled by the variable C_1 is absent, the process modelled by X_{C_1} will not be initiated, that is, $\Pr(x_{C_1} \mid \bar{c}_1, A) = 0$ for all values of the annihilator variable A. The conditional probability table for X_{C_1} further specifies:

$$\begin{aligned}
&\Pr(x_{C_1} \mid c_1, A=0) = \Pr(e \mid c_1, \bar{c}_2, \bar{c}_3)\\
&\Pr(x_{C_1} \mid c_1, A=1) = 0\\
&\Pr(x_{C_1} \mid c_1, A=2) = \Pr(e \mid c_1, \bar{c}_2, \bar{c}_3) - \alpha_{1|13} \cdot \Pr(e \mid c_1, \bar{c}_2, \bar{c}_3)\\
&\Pr(x_{C_1} \mid c_1, A=3) = \Pr(e \mid c_1, \bar{c}_2, \bar{c}_3) - \alpha_{1|12} \cdot \Pr(e \mid c_1, \bar{c}_2, \bar{c}_3)\\
&\Pr(x_{C_1} \mid c_1, A=4) = \Pr(e \mid c_1, \bar{c}_2, \bar{c}_3) - \alpha_{1|123} \cdot \Pr(e \mid c_1, \bar{c}_2, \bar{c}_3)
\end{aligned}$$

where the parameter $\alpha_{1|12}$ captures the degree of cancellation of the process x_{C_1} in the presence of the causes c_1 and c_2; $\alpha_{1|13}$ has a similar meaning. Analogously, the parameter $\alpha_{1|123}$ captures the degree of cancellation of x_{C_1} when all causes c_1, c_2 and c_3 are present simultaneously. To arrive at a fully specified probability table, we further set the probability $\Pr(x_{C_1} \mid c_1, A=1)$ to 0; note that any probability value would have suited our purpose since by definition the value combination c_1, $A=1$ is impossible. The conditional probability table for the effect variable E is defined as before: the effect e will arise with probability 1 if at least one of the intermediate processes is activated.

From the underlying conceptual mechanism we now find that application of our cancellation model to n cause variables for a single effect variable results in

$$\Pr(e \mid \mathbf{c}) = 1 - \prod_{j \in J} \left(1 - \Pr(e \mid \bar{c}_1, .., c_j, .., \bar{c}_n) + \alpha_{j|J} \cdot \Pr(e \mid \bar{c}_1, .., c_j, .., \bar{c}_n)\right)$$

where J again is the set of indices of the cause variables C_j which are marked as being present in the value combination $\mathbf{c}$.

To conclude, we observe that our intercausal cancellation model for mechanisms with three cause variables includes nine cancellation parameters. In general, a causal mechanism with n cause variables requires as many as $n \cdot 2^{n-1} - n$ cancellation parameters to model all possible cancellation effects among the processes involved. A present cause can have $1,\ldots,n-1$ simultaneously present causes with which it may show a joint cancellation effect. For this cause therefore, a total of $\sum_{i=1}^{n-1} \binom{n-1}{i} = 2^{n-1} - 1$ cancellation parameters are required. For all causes together, the total number of parameters thus indeed amounts to $n \cdot 2^{n-1} - n$. We note that for each combination of k present causes, k parameter values need to

be specified. Since for all combinations of present causes the patterns of cancellation involved can be quite different, this number of parameters cannot be reduced without losing generality of the model. We note that the number of cancellation parameters is even larger than the 2^n probabilities that would be necessary to directly specify the conditional probability table for the effect variable E. As argued in Section 4.2, values for these cancellation parameters are not elicited directly from domain experts. When during the elicitation process intercausal cancellations are uncovered, only the cancellation parameters involved need be elicited; all other cancellation parameters are set to zero. Based upon our experience moreover, we feel that values for the parameters are generally easier to obtain than the compound probabilities including the cancellation effects. Because the cancellation parameters refer to underlying processes, they are easier to envision for the experts than the direct probabilities or the regulatory parameters themselves are. When an expert still feels uncomfortable with attaching a numeral value to a specific cancellation effect, the maximum-entropy value introduced before can be used. Even with this crude choice of value is our cancellation model better able to describe the intercausal interactions involved than the noisy-OR model. We would further like to emphasise once more that the conditional probability tables detailed in this section belong to the underlying conceptual mechanism and not to the intercausal cancellation model itself. As a consequence, the sizes of these tables are not directly relevant; in fact, use of the intercausal cancellation model does not increase the number of variables and/or probabilities in the Bayesian network being developed when compared to the noisy-OR model.

6. Intercausal cancellation in real life

To investigate the practicability of the intercausal cancellation model developed above, we studied several examples from the domain of pharmacology with the help of a domain expert. We asked her to assess conditional probabilities for particular medical conditions being present given possible treatment regimes. For application of the intercausal cancellation model to our domain, we employed the inverted version, since initially a specific condition (the effect) is present and medication is being administered (the cause) to remove it. Cancellation occurs regularly among medications, where one medication can cancel the intended working of another medication when administered simultaneously. We give a basic example and a more elaborate one.

Example 6.1. We consider two medications for patients with primary type-1 osteoporosis. In healthy persons, the amount of bone mass in the skeleton is controlled by two types of cell: the osteoclasts break down, or resorb, bone material and the osteoblasts form bone tissue from calcium. With osteoporosis, the rate of bone resorption exceeds that of bone formation, resulting in a decrease in bone mass. The risks of bone fracture associated with osteoporosis, can be reduced to some small extent by treatment. Two common treatments for osteoporosis are calcium supplementation and medication with bisphosphonates.

Calcium supplementation is aimed at providing the osteoblasts with sufficient material for bone formation. Bisphosphonates on the other hand inhibit the resorption of bone by binding to the calcium in the bone tissue to increase osteoclast death rate. Bisphosponates are provided in small dosage but are much more effective than calcium supplementation. Upon concurrent intake however, the bisphosphonates will bind to the calcium supplements in the stomach rather than to the calcium in the bone tissue. As a result effects of both treatments are decreased. Because of their small dosage the effect of the bisphosphonates is likely to disappear altogether; since the calcium is administered in larger quantities, some effect of the supplementation is expected to remain. Based upon this knowledge, we conclude that concurrent intake of both medications will fully cancel out the effect of the bisphosphonates; the effect of the calcium supplementation is cancelled out only partially. The domain expert provided the following assessments for the conditional probability table of the effect variable O modelling primary type-1 osteoporosis:

$$\Pr(\bar{o} \mid b, \bar{c}) = 0.85$$
$$\Pr(\bar{o} \mid \bar{b}, c) = 0.15$$

where b captures medication by bisphosphonates and c models the administration of calcium supplements. These probabilities show that after administration of bisphosphonates to a patient with type-1 osteoporosis, the probability of the patient no longer suffering from osteoporosis is 0.85, and when calcium supplements are administered, this probability is 0.15.

When applying the (inverted) noisy-OR model with these two probabilities as its parameters, the probability of osteoporosis being absent with the administration of both bisphosphonates and calcium supplements is

$$\begin{aligned}\Pr(\bar{o} \mid b, c) &= 1 - (1 - 0.85) \cdot (1 - 0.15) \\ &= 0.8725\end{aligned}$$

which is higher compared to administering either bisphosphonates or calcium supplements. When applying the inverted cancellation model with the additional parameters $\alpha_{b|bc} = 1$, modelling full cancellation of the bisphosphonates, and $\alpha_{c|bc} = \frac{1}{2}$, modelling partial cancellation of the calcium supplements, the probability of osteoporosis being absent with the administration of both bisphosphonates and calcium supplements is

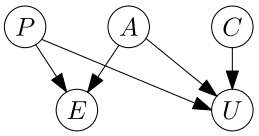


Fig. 8. The graphical structure of the joint causal mechanism with the cause variables P (proton pump inhibitors), A (antacids) and C (cefuroxime), and the effect variables E (epigastric pains) and U (urinary tract infection).

$$\begin{aligned}\Pr(\bar{o} \mid b, c) &= 1 - (1 - 0.85 + \alpha_{b|bc} \cdot 0.85) \cdot (1 - 0.15 + \alpha_{c|bc} \cdot 0.15) \\ &= 1 - (1 - 0.85 + 1 \cdot 0.85) \cdot (1 - 0.15 + \tfrac{1}{2} \cdot 0.15) \\ &= 1 - 1 \cdot 0.925 \\ &= 0.075\end{aligned}$$

We had asked our expert to assess not just the parameter probabilities $\Pr(\bar{o} \mid b, \bar{c})$ and $\Pr(\bar{o} \mid \bar{b}, c)$, but also the probability $\Pr(\bar{o} \mid b, c)$. She had estimated this latter probability of the absence of type-1 osteoporosis when both bisphosphonates and calcium supplements are administered at 0.05. We observe that her assessment is much better approximated by the cancellation model than by the noisy-OR model. □

In the example above, a causal mechanism with just two cause variables was considered. Our model however, is also applicable to more involved joint mechanisms. The following example includes two conditions and their corresponding medications, with interaction effects among the two groups of medications.

Example 6.2. We consider two different medications which are commonly administered to patients suffering from epigastric pains as a result of pyrosis, that is, from heartburn associated with regurgitation of gastric acid. Pyrosis-associated pains are relieved by the administration of either antacids, to reduce the acidity of the stomach contents, or proton pump inhibitors, which reduce the production of acid by the cells in the wall of the stomach. Upon concurrent intake the proton pump inhibitors bar the production of acids, with a reasonably pH-neutral stomach contents for a result. The antacids then no longer contribute to pain relief. Based upon knowledge of the underlying processes, the interaction among the two substances constitutes a single-sided, possibly full cancellation. The cancellation effects for antacids and proton pump inhibitors are modelled with $\alpha_{a|ap} = 1$ and $\alpha_{p|ap} = 0$ respectively, where a captures medication by antacids and p models the administration of proton pump inhibitors.

We now consider a patient with epigastric pains, who in addition suffers from a (common) urinary tract infection. Patients with such an infection are usually administered the cefuroxime antibiotic. The graphical structure of the joint causal mechanism describing the effects of the three medications on the two conditions, is depicted in Fig. 8. Now, an intrinsic property of the cefuroxime molecule is its inability of being absorbed from insufficiently acid stomach contents. On concurrent intake of proton pump inhibitors and cefuroxime in fact, the acidity in the stomach contents will be reduced by the proton pump inhibitors and as a result only some 50% of the cefuroxime substance will be absorbed. A similar observation holds for the concurrent intake of antacids and cefuroxime: some 43% of the cefuroxime substance will then be absorbed. The cancellation effect of cefuroxime on either the proton pump inhibitors or the antacids is non-existent, and with the concurrent intake of all three medications the cancellation effects among the proton pump inhibitors and the antacids remain unaltered. For the intercausal cancellation effects encountered during the elicitation process, the values for the associated cancellation parameters were readily obtained from the expert; the other cancellation parameters were set to zero to describe the lack of cancellation effects. The parameter values thus obtained are as follows:

$$\begin{array}{lll} \alpha_{a|ac} = 0.00 & \alpha_{p|ap} = 0.00 & \alpha_{a|acp} = 1.00 \\ \alpha_{c|ac} = 0.57 & \alpha_{c|cp} = 0.50 & \alpha_{c|acp} = 0.50 \\ \alpha_{a|ap} = 1.00 & \alpha_{p|cp} = 0.00 & \alpha_{p|acp} = 0.00 \end{array}$$

When using our intercausal cancellation model for the variable U modelling the urinary tract infection, the domain expert needs to provide the probabilities $\Pr(u \mid a, \bar{c}, \bar{p})$, $\Pr(u \mid \bar{a}, c, \bar{p})$ and $\Pr(u \mid \bar{a}, \bar{c}, p)$; for the variable E modelling epigastric pains, the parameter probabilities $\Pr(e \mid a, \bar{p})$ and $\Pr(e \mid \bar{a}, p)$ are to be assessed. Our domain experts gave the following assessments for these parameter probabilities:

$$\begin{array}{ll} \Pr(\bar{e} \mid a, \bar{p}) = 0.30 & \Pr(\bar{u} \mid a, \bar{c}, \bar{p}) = 0.00 \\ \Pr(\bar{e} \mid \bar{a}, p) = 0.95 & \Pr(\bar{u} \mid \bar{a}, c, \bar{p}) = 0.97 \\ & \Pr(\bar{u} \mid \bar{a}, \bar{c}, p) = 0.00 \end{array}$$

Using the values of the cancellation parameters and the provided probability assessments, all other conditional probabilities required for the probability tables of the variables E and U are computed by our model. As an example, we show the calculation of one of these probabilities:

Table 1
The probability values calculated by the noisy-OR model and by the intercausal cancellation model respectively, for epigastric pains (e) and a urinary tract infection (u), when proton pump inhibitors (p), antacids (a) and/or cefuroxime (c) are administered.

	Noisy-OR model	Cancellation model
$\Pr(\bar{e} \mid a, p)$	0.97	0.95
$\Pr(\bar{u} \mid a, c, \bar{p})$	0.97	0.42
$\Pr(\bar{u} \mid a, \bar{c}, p)$	0.00	0.00
$\Pr(\bar{u} \mid \bar{a}, c, p)$	0.97	0.49
$\Pr(\bar{u} \mid a, c, p)$	0.97	0.49

$$\begin{aligned}\Pr(\bar{u} \mid a, c, p) = 1 - \big((1 - \Pr(\bar{u} \mid a, \bar{c}, \bar{p}) + \alpha_{a|acp} \cdot \Pr(\bar{u} \mid a, \bar{c}, \bar{p})) \\ \cdot (1 - \Pr(\bar{u} \mid \bar{a}, c, \bar{p}) + \alpha_{c|acp} \cdot \Pr(\bar{u} \mid \bar{a}, c, \bar{p})) \\ \cdot (1 - \Pr(\bar{u} \mid \bar{a}, \bar{c}, p) + \alpha_{p|acp} \cdot \Pr(\bar{u} \mid \bar{a}, \bar{c}, p))\big) = 0.49\end{aligned}$$

Table 1 specifies the calculated probabilities for the variables E and U. For comparison purposes, the table further shows the same probabilities yet now calculated from the noisy-OR model. We observe that the probabilities calculated by the two models for the variable U show large differences as a result of the modelled intercausal cancellation effects. □

For the example above, we spent considerable effort eliciting from our expert all conditional probabilities for the two tables involved; the basic idea was again to compare the expert-assessed probabilities with the calculated ones. During multiple elicitation sessions spread over several meetings, and despite the use of various different elicitation techniques, it proved too difficult to obtain estimates for the probabilities involved with which the expert would feel comfortable: since the three medications are hardly ever administered simultaneously, she had virtually no evidence on which to base her assessments. The expert could more readily describe and quantify the interaction among the various substances however, based on reports of pre-clinical tests and case reports and on her knowledge about the workings of the medications. In fact, she was able, with confidence, to provide us with sufficient information to let us choose appropriate values for the cancellation parameters, thereby providing us with enough knowledge to construct the conditional probability tables for the variables E and U. Even this small example serves to show that without applying the intercausal cancellation model, we would have not been able to arrive at a fully specified network without forcing the expert to provide the yet missing probabilities.

7. Related work

Over the years, numerous variants and generalisations of the noisy-OR model have been developed, some of which allow capturing intercausal cancellation to at least some degree; we would like to argue that the mere existence of these variants suggests a demand for models for accurately describing patterns of cancellation among causal effects. In this section, we compare our newly designed intercausal cancellation model against two existing related models. We will briefly review the recursive noisy-OR and non-impeding noisy-AND tree models, and demonstrate the different modelling abilities and engineering advantages of our model.

7.1. The recursive noisy-OR model

The recursive noisy-OR model, or RNOR model for short [12], was developed as a variant of the noisy-OR model. One of the goals for its design was to provide for more accurate conditional probability tables in the presence of dependent causal interactions. The basic idea of the model is to recursively define a conditional probability table and thereby explicitly allow the possibility of inserting combinations of causes with dependent interactions in any step of the table's construction.

The recursive noisy-OR model is a parameterised conditional probability table, just like the noisy-OR model. Initially, its parameter probabilities are the conditional probabilities of the effect arising given each cause separately. If upon constructing the probability table no combinations of causes with dependent interactions are inserted, the recursive noisy-OR model returns the same probabilities for the effect variable as the noisy-OR model. We now consider the construction of a conditional probability table for the effect variable E of a mechanism with the four cause variables $C_1, \ldots, C_4$. In the first step, given the values obtained for the parameter probabilities, the model calculates the conditional probabilities of the effect arising in the presence of all possible pairs of causes. The calculations in this first step are equal to those performed by the noisy-OR model:

$$\Pr(e \mid c_i, c_j, \bar{\mathbf{c}}) = 1 - (1 - \Pr(e \mid c_i, \bar{c}_j, \bar{\mathbf{c}})) \cdot (1 - \Pr(e \mid \bar{c}_i, c_j, \bar{\mathbf{c}}))$$

for all $i, j = 1, \ldots, 4$, $i \neq j$, where the value combination $\bar{\mathbf{c}}$ model the absence of all causes other than c_i and c_j. If no dependent causal interaction occurs within such a pair of causes, the calculated probability is inserted into the probability table under construction and used in the next step of the recursion. Otherwise, the calculated probability is superseded by

an expert-provided one. In each subsequent step of the recursion the probabilities of the effect arising in the presence of exactly k causes are calculated with the probabilities resulting from the previous two steps. For the four-cause mechanism, the probability $\Pr(e \mid c_1, c_2, c_3, c_4)$ with $k = 4$ for example, is calculated as:

$$1 - \left(\frac{1 - \Pr(e \mid \bar{c}_1, c_2, c_3, c_4)}{1 - \Pr(e \mid \bar{c}_1, c_2, c_3, \bar{c}_4)} \cdot \frac{1 - \Pr(e \mid c_1, \bar{c}_2, c_3, c_4)}{1 - \Pr(e \mid c_1, \bar{c}_2, \bar{c}_3, c_4)} \cdot \frac{1 - \Pr(e \mid c_1, c_2, \bar{c}_3, c_4)}{1 - \Pr(e \mid c_1, c_2, \bar{c}_3, \bar{c}_4)} \cdot \frac{1 - \Pr(e \mid c_1, c_2, c_3, \bar{c}_4)}{1 - \Pr(e \mid c_1, \bar{c}_2, \bar{c}_3, c_4)} \right)$$

The probabilities used in the calculations may have been estimated by an expert or calculated from other probabilities themselves. The recursiveness of the construction of the probability table for the effect variable thus allows the possibility of explicitly inserting expert-provided probabilities and using these upon further calculation. Unfortunately, the probabilities used in such further calculations are not uniquely identified by the model [4]. In the above calculation of the probability $\Pr(e \mid c_1, c_2, c_3, c_4)$, for example, four probabilities of the effect arising in the presence of two out of the four cause variables present have been chosen to constitute the denominator. Choosing another combination of four such probabilities from among the six possibilities can result in a different value for the probability $\Pr(e \mid c_1, c_2, c_3, c_4)$ [4].

When expert-provided probabilities are included in the recursive construction of a probability table, these are restricted to be larger than the maximum of the probabilities used for the model-calculated value. If an expert-provided value is smaller than the probability values used, the probability values calculated upon further construction are not guaranteed to be valid [4], that is, to be within the range $[0, 1]$. Informally spoken, the numerator of the fraction in the definition of a recursive probability needs to be larger than the denominator since otherwise a value larger than 1 and a calculated probability value smaller than 0 can result.

As a further generalisation of the recursive noisy-OR model, the inhibited recursive noisy-OR model [8] was developed to allow a combination of positive and negative influences within a causal mechanism. The basic idea is to separate the negative and positive influences into two groups and apply the (inverted) recursive noisy-OR model to each group separately. For the conditional probability table of the effect variable, the probabilities established from the two groups are combined through an assumption of independence. As it applies the recursive noisy-OR model however, the inhibited recursive noisy-OR model inherits the same limited ability to accurately describe intercausal cancellation.

We conclude that the recursive noisy-OR model and its extension offer limited possibilities for capturing intercausal cancellation, that is, in addition to the inherited shortcomings from the recursive noisy-OR model, modelling two causes which individually have a positive influence on the effect, but when combined they exert a negative influence, can only be achieved by forcing the expert to provide a compound probability value for that given combination. This results in a major drawback for the support offered to the network engineer for describing patterns of cancellation. As argued in Sections 3 and 6, our experiences are showing that obtaining assessments for such probabilities can be especially difficult, if not impossible.

7.2. The NIN-AND tree model

The non-impeding noisy-AND tree model, or NIN-AND tree model for short [22,23], was developed as a causal interaction model allowing the representation of cancelling, or undermining, causal interactions in addition to reinforcing ones. While the noisy-OR model and its variants build on the assumption that individual causes serve to reinforce one another's effect within the bounds defined by independence, the NIN-AND tree model is founded upon two separate logical principles describing reinforcement and cancellation instead [23]. The model uses an explicit representation of the dependent interactions among the causal influences on a common effect before constructing the conditional probability table for the associated effect variable. This representation essentially is a tree structure with noisy-AND gates for its building blocks, each of which models either reinforcement or cancellation among a designated subset of causes. The conditional probability table for the effect variable of a causal mechanism then is constructed by evaluating the NIN-AND tree model of the intercausal interactions in the mechanism, with in essence the same input parameter probabilities as the noisy-OR model.

Evaluation of a reinforcement gate from the constructed NIN-AND tree results in the following conditional probabilities of the effect arising in the presence of the combination $\mathbf{c}$ of causes:

$$\Pr(e \mid \mathbf{c}) = 1 - \prod_{j \in J} (1 - \Pr(e \mid c_j))$$

where J is the set of indices of the causes in the gate at hand. We note that with a NIN-AND tree composed of a single noisy-AND gate modelling reinforcement, the same conditional probabilities are found as with the noisy-OR model. Evaluation of a cancellation gate from the NIN-AND tree results in the following conditional probabilities of the effect arising:

$$\Pr(e \mid \mathbf{c}) = \prod_{j \in J} \Pr(e \mid c_j)$$

Table 2
Probability values calculated by the noisy-OR model, the intercausal cancellation model, the inhibited recursive noisy-OR model and the NIN-AND tree model, and provided by the expert, respectively, for type-1 osteoporosis (o) being absent after both bisphosphonates (b) and calcium supplements (c) have been administered.

	Noisy-OR model	Cancellation model	RNOR model	NIN-AND tree model	Expert provided
$\Pr(\bar{o} \mid b, c)$	0.8725	0.075	0.05	0.1275	0.05

Table 3
Probability values calculated by the noisy-OR model, the intercausal cancellation model, the inhibited recursive noisy-OR model and the NIN-AND model, respectively, for epigastric pains (e) and a urinary tract infection (u), when proton pump inhibitors (p), antacids (a) and/or cefuroxime (c) are administered.

	Noisy-OR model	Cancellation model	RNOR model	NIN-AND model	Expert provided
$\Pr(\bar{e} \mid a, p)$	0.97	0.95	0.97	0.285	–
$\Pr(\bar{u} \mid a, c, \bar{p})$	0.97	0.42	0.97	0.97	–
$\Pr(\bar{u} \mid a, \bar{c}, p)$	0.00	0.00	0.00	0.00	–
$\Pr(\bar{u} \mid \bar{a}, c, p)$	0.97	0.49	0.97	0.97	–
$\Pr(\bar{u} \mid a, c, p)$	0.97	0.49	0.97	0.97	–

where J again is the set of indices of the causes in the gate at hand. Any probability value which is thus established for the conditional probability table for the effect variable under consideration may be replaced with an expert-provided assessment, similar to the recursive noisy-OR model, for use upon further evaluation of the NIN-AND tree.

Although the NIN-AND tree model allows explicitly capturing cancellation among causal effects, it assumes a single fixed type of cancellation which does not allow expressing partial or one-sided cancellation. Another drawback of the model lies in the efforts required from both the network engineer and the domain experts involved. Constructing a tree of noisy-AND gates modelling all intercausal interactions has proven to be a non-trivial and error-prone task [24]. The developers of the NIN-AND tree model therefore propose an iterative approach in which the tree's construction is interleaved with the direct elicitation of probabilities from domain experts. These probabilities again are the conditional probabilities involving compound conditions which we found extremely difficult, if not impossible, to obtain from the expert in our application domain.

7.3. The real-life examples revisited

In Section 6 we studied two real-life examples from the domain of pharmacology and illustrated the use of our intercausal cancellation model for these two fragments of expert knowledge. Now we use the same examples and calculate the conditional probability tables for the effect variables involved by means of the models reviewed above.

Table 2 pertains to the treatment of type-1 osteoporosis as reviewed in Example 6.1, and reports the values which are established by the various models for the conditional probability of osteoporosis being absent after treatment with both bisphosphonates and calcium supplements. When the values discussed in Section 6 are used for the cancellation parameters involved, the intercausal cancellation model results in the probability 0.075. From Section 6, we recall that the noisy-OR model would result in the much higher probability value of 0.8725. The inhibited recursive noisy-OR model will initially also establish this high probability, but will insert the expert-provided value of 0.05 instead. For application of the NIN-AND tree model, we used a NIN-AND tree consisting of a single noisy-AND gate modelling cancellation, to fit the example as well as possible. With this representation of the interactions among the causal effects, the probability of type-1 osteoporosis being absent after treatment was found to be 0.1275. We would like to note that, while we could accommodate in our cancellation model detailed knowledge about the different cancellation effects between the two pharmaceutic substances, the NIN-AND tree model offered less modelling flexibility.

Table 3 pertains to simultaneous treatment of epigastric pains and a urinary tract infection by administering three pharmaceutical substances as reviewed in Example 6.2, and reports the values established by the various models for the conditional probability tables for the two effect variables. As discussed in Section 6, we used knowledge-based values for the cancellation parameters upon application of our intercausal cancellation model. The probabilities thus found for the absence of the urinary tract infection specifically differ from those established from the noisy-OR model as a result. Since we had not been able to obtain assessments from our domain expert for the conditional probabilities involving compound conditions, application of the inhibited recursive noisy-OR model could not benefit from these assessments and based on the parameter probabilities of the inhibited recursive noisy-OR model, the same probability values were returned as for the noisy-OR model. For application of the NIN-AND tree model, we constructed a NIN-AND tree for each probability of the urinary tract infection being resolved by treatment with a specific combination of pharmaceutic substances. For calculating the probability $\Pr(\bar{u} \mid a, c, p)$ for example, a tree composed of two noisy-AND gates was constructed, in which the output of a gate modelling reinforcement among the antacids and proton pump inhibitors served as input for the second gate modelling cancellation of the effect of the cefuroxim antibiotic. The table shows that our attempts to describe knowledge of the degrees of cancellation among the cefuroxim antibiotic and either the anticids or the proton pump inhibitors resulted in the same conditional probabilities as found with the noisy-OR model. Although this

finding originates specifically from the latter two substances having no direct influence on the urinary tract infection, it illustrates the lack of possibilities for capturing the specific patterns of partial cancellation involved in the NIN-AND tree model.

8. Conclusions

While a variety of causal interaction models are available to Bayesian-network engineers, only few models are concerned with cancellation effects among causes. These models moreover offer modelling possibilities for a limited range of cancellation patters and offer little support to the network engineer for eliciting such patterns from domain experts. Motivated by this observation and by the requirements of our application domain, we developed a new interaction model to describe intercausal cancellation. Our model was designed from first principles, along the same lines as the popular noisy-OR model. Just like the noisy-OR model, our intercausal cancellation model serves to ameliorate the burden of probability elicitation upon constructing a Bayesian network with the help of domain experts. The model even requires the same parameters as the noisy-OR model, albeit with additional regulatory ones. The main advantage of our new model lies not so much in reducing the number of numerical values required however, but in forestalling the need to elicit assessments for those probabilities which often are considered hardest to provide, that is, for probabilities given the presence of multiple simultaneous causes. We experienced this advantage in our application in the field of pharmacology, where our intercausal cancellation model resulted in satisfactory probability tables. We found more specifically that the additional regulatory parameters required by our model were considerably easier to obtain than the more involving probabilities with compound conditions. Upon comparing the intercausal cancellation model against existing causal interaction models we further found that our model offered more flexibility for describing the patterns of cancellation found in our application domain of pharmacology. Based on these experiences, we feel that the basic ideas presented in this paper constitute a practical intercausal cancellation model for ready use by network engineers in a wide range of real-world applications.

Acknowledgements

This paper is an extended version of the BNAIC 2013 paper *Intercausal Cancellation in Bayesian Networks* [20]. We would like to thank all reviewers associated with BNAIC and the IJAR, for their helpful suggestions for extending and improving the paper. The research is partly funded with financial support from the Prevention of and Fight against Crime Programme of the EU European Commission – Directorate-General Home Affairs.

References

[1] J. Bolt, L.C. van der Gaag, An empirical study of the use of the noisy-OR model in a real-life Bayesian network, in: E. Hüllermeier, R. Kruse, F. Hoffmann (Eds.), Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Methods, vol. 80, Springer, Berlin, 2010, pp. 11–20.
[2] G.F. Cooper, The computational complexity of probabilistic inference using Bayesian belief networks, Artif. Intell. J. 42 (1990) 393–405.
[3] F.J. Díez, Parameter adjustment in Bayes networks. The generalized noisy-OR gate, in: D. Heckerman (Ed.), Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers, 1993, pp. 99–105.
[4] F.J. Díez, M.J. Druzdzel, Canonical probabilistic models for knowledge engineering, Technical Report CISIAD-06-01, UNED, Madrid, 2006.
[5] A. Oniśko, M.J. Druzdzel, H. Wasyluk, Learning Bayesian network parameters from small data sets: application of noisy-OR gates, Int. J. Approx. Reason. 27 (2001) 165–182.
[6] M.J. Druzdzel, L.C. van der Gaag, Building probabilistic networks: "Where do the numbers come from?" Guest editors introduction, IEEE Trans. Knowl. Data Eng. 12 (2000) 481–486.
[7] L.C. van der Gaag, S. Renooij, C.L.M. Witteman, B.M.P. Aleman, B.G. Taal, Probabilities for a probabilistic network: a case study in oesophageal cancer, Artif. Intell. Med. 25 (2001) 123–148.
[8] D.E. Gossink, M. Shahin, J.F. Lemmer, I. Fuss, A practical approach to dynamic Bayesian network, in: Information, Decision and Control, 2007, pp. 71–77.
[9] D. Heckerman, J. Breese, Causal independence for probabilistic assessment and inference using Bayesian networks, IEEE Trans. Syst. Man Cybern. 26 (1996) 826–831.
[10] M. Henrion, Some practical issues in constructing belief networks, in: J. Lemmer, T. Levitt, L. Kanal (Eds.), Proceedings of the Third Annual Conference on Uncertainty in Artificial Intelligence (UAI-87), Elsevier, 1987, pp. 161–174.
[11] F.V. Jensen, Th.D. Nielsen, Bayesian Networks and Decision Graphs, second edition, Springer-Verlag, New York, 2007.
[12] J.F. Lemmer, D.E. Gossink, Recursive noisy OR – a rule for estimating complex probabilistic interactions, in: Transactions on Systems, Man and Cybernetics, Part B, vol. 34, IEEE Press, Piscataway, NJ, USA, 2004, pp. 2252–2261.
[13] S.L. Lauritzen, D.J. Spiegelhalter, Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems, Morgan Kaufmann Publishers, San Francisco, 1990.
[14] P. Lucas, Bayesian network modelling through qualitative patterns, Artif. Intell. J. 163 (2003) 233–263.
[15] P.P. Maaskant, M.J. Druzdzel, An independence of causal interactions model for opposing influences, in: Proceedings of the 4th European Workshop on Probabilistic Graphical Models, 2008, pp. 185–192.
[16] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann Publishers, San Francisco, 1988.
[17] S. Renooij, Probability elicitation for belief networks: issues to consider, in: The Knowledge Engineering Review, vol. 16, Cambridge University Press, 2001, pp. 255–269.
[18] D. Roth, On the hardness of approximate reasoning, Artif. Intell. 82 (1996) 273–302.
[19] S.P.D. Woudenberg, L.C. van der Gaag, Using the noisy-OR model can be harmful ... but it often is not, in: W. Liu (Ed.), Proceedings of the 11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Springer-Verlag, Berlin, 2011, pp. 122–133.
[20] S.P.D. Woudenberg, L.C. van der Gaag, C.M.A. Rademaker, Intercausal cancellation in Bayesian networks, in: J.W.H.M. Uiterwijk (Ed.), Proceedings of the 24th Benelux Conference on Artificial Intelligence, 2012, pp. 242–249.

[21] S.P.D. Woudenberg, L.C. van der Gaag, The sensitivity of Bayesian networks to noisy-OR calculated probability values, Int. J. Approx. Reason. 61 (2015) 1–15.
[22] Y. Xiang, Non-impeding noisy-AND tree causal models over multi-valued variables, Int. J. Approx. Reason. 53 (2012) 988–1002.
[23] Y. Xiang, N. Jia, Modeling causal reinforcement and undermining for efficient CPT elicitation, IEEE Trans. Knowl. Data Eng. 19 (12) (2007) 1708–1718.
[24] Y. Xiang, M.T. Truong, Z.J. Zhu, Indirect elicitation of NIN-AND trees in causal model acquisition, in: S. Benferhat, J. Grant (Eds.), Proceedings of the Fifth International Conference on Scalable Uncertainty Management, Springer, 2011, pp. 261–274.