Bayesian Network Modelling by Qualitative Patterns

Peter Lucas [1]

Abstract. In designing a Bayesian network for an actual problem, developers need to bridge the gap between the mathematical abstractions offered by the Bayesian-network formalism and the features of the problem to be modelled. A notion that has been suggested in the literature to facilitate Bayesian-network development is causal independence. It allows exploiting compact representations of probabilistic interactions among variables in a network. However, only very few types of causal independence are in use today, as only the most obvious ones are really understood. We believe that qualitative probabilistic networks (QPNs) may be useful in helping understand causal independence. Originally, QPNs have been put forward as qualitative analogues to Bayesian networks. In this paper, we deploy QPNs in developing and analysing a collection of qualitative, causal interaction patterns, called *QC patterns*. These are endowed with a fixed qualitative semantics, and are intended to offer developers a high-level starting point when developing Bayesian networks.

1 INTRODUCTION

The Bayesian network formalism offers a powerful framework for the modelling of uncertain interactions among variables in a domain. Such interactions are represented in two different manners. Firstly, in a qualitative manner, by means of a directed acyclic graph. Secondly, in a quantitative manner, by specifying a set of conditional probability distributions for every variable in the network. These probability distributions allow for expressing various logical, functional and probabilistic relationships among variables; much of the power of the Bayesian network formalism derives from this feature [6].

It is well known that ensuring that the topology of a Bayesian network is sparse eases the assessment of its underlying joint probability distribution, as the required probability tables will then be relatively small. Unfortunately, designing a network with a topology that is sparse is neither easy nor always possible. Researchers have therefore proposed special types of independence relationship in order to facilitate probability assessment. In particular the theory of *causal independence* fulfils this purpose [3]. The theory allows for the specification of the interactions among variables in terms of cause-effect relationships and functions, adopting particular statistical independence assumptions. Causal independence is frequently used in practical networks. However, a limitation of the theory of causal independence is that it is usually unclear with what sort of qualitative behaviour a network will be endowed when choosing for a particular interaction type. As a consequence, only two types of interaction are in frequent use: the noisy-OR and the noisy-MAX [1, 4, 6].

Qualitative probabilistic networks (QPNs) offer a qualitative analogue to the formalism of Bayesian networks. They allow describing the dynamics of interaction among variables in a purely qualitative fashion by means of the specification and propagation of qualitative signs [7, 8]. Hence, QPNs abstract from the numerical detail.

The aim of the present work was to develop a theory of qualitative, causal interaction patterns, *QC patterns* for short, in the context of Bayesian networks. Interaction types are proposed, and QPNs are then used to provide a qualitative semantic foundation for these interactions. The Bayesian-network developer is supposed to utilise the theory by selecting appropriate interaction patterns based on domain properties, which thus can guide Bayesian-network development.

In the following section, the basic properties of Bayesian networks are introduced, as are the notion of causal independence and qualitative probabilistic networks. We start the analysis by considering various causal-independence models, unravelling the qualitative behaviour of these causal models using QPNs in Section 3. Finally, in Section 4, it is summarised what has been achieved.

2 PRELIMINARIES

2.1 Bayesian networks

A *Bayesian network* is a concise representation of a joint probability distribution on a set of statistical variables [6]. It consists of a qualitative part and an associated quantitative part. The *qualitative part* takes the form of an acyclic directed graph, or digraph for short, $G = (V(G), A(G))$. Here, $V(G)$ is a set of nodes standing for statistical variables; all variables $V \in V(G)$ are assumed to be binary. For abbreviation, we will often use v to denote $V = \top$ (true) and $\bar{v}$ to denote $V = \bot$ (false). Sometimes, a variable's values are kept unspecified, i.e. it is utilise as a free variable. Arcs $V \to V' \in A(G)$ are used to model statistical (in)dependence.

We use the notation $V_1, \ldots, V_n \backslash V_i, \ldots, V_j$ to stands for the set of variables $\{V_1, V_2, \ldots, V_{i-1}, V_{i+1}, \ldots, V_{j-1}, V_{j+1}, \ldots, V_n\}$. Furthermore, an expression such as

$$\sum_{\substack{I_1, I_2 \\ \psi(I_1, \ldots, I_n) = e}} g(I_1, \ldots, I_n)$$

stands for summing over the firstly mentioned collection of variables, here the variables I_1, I_2, with the equality acting as a constraint. If these variables are not mentioned separately, the expression $\psi(I_1, \ldots, I_n) = e$ is interpreted as varying over all variables I_k in the equality for which the constraint holds.

The *quantitative part* of a Bayesian network consists of a set of *conditional probabilities* $\Pr(V \mid \pi(V))$, for each $V \in V(G)$, describing the joint influence of values for the parents $\pi(V)$ of V on the probabilities of variable V's values. A Bayesian network $\mathcal{B} = (G, \Pr)$ provides for computing any probability of interest.

A real-life example is shown in Figure 1. In this network, it is modelled that patients may become colonised by specific bacteria, for

[1] Institute for Informatics and Information Science, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands, E-mail: peterl@cs.kun.nl, lucas@cs.uu.nl

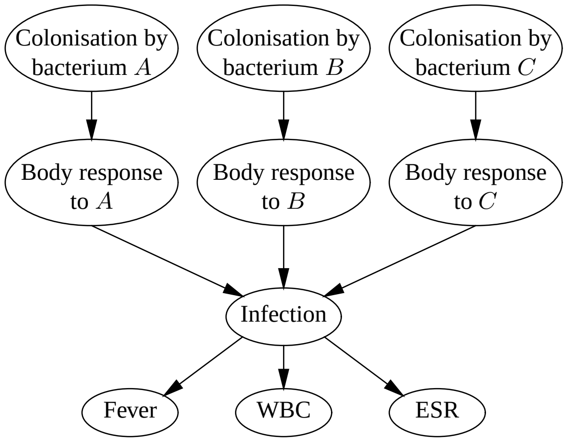

Figure 1. Example Bayesian networks, modelling the interaction among bacteria possibly causing an infection in a patient after colonisation.

example *P. aeruginosa*, after admission to a hospital. As the actual names of the bacteria do not matter here, they are simply called A, B and C. After having been colonised, the patient's body responds to the bacteria in various ways; for example, an infection may develop. An infection is clinically recognised by symptoms and signs such as fever, high white blood cell count (WBC), and increased sedimentation rate of the blood (ESR). Clearly, the probability distribution $\Pr(\text{Infection} \mid \text{BR}_A, \text{BR}_B, \text{BR}_C)$ specified for the network, where BR_X stands for 'Body response to X', is of great importance in modelling interactions among the various mechanisms causing infection.

2.2 Causal independence

One way to specify interactions among statistical variables in a compact fashion is offered by the notion of *causal independence* [3]. This theory offers one half of our method of QC patterns.

2.2.1 Probabilistic representation

The general structure of a causal-independence model is shown in Figure 2; it expresses the idea that causes $C_1, \ldots, C_n$ influence a given common effect E through intermediate variables $I_1, \ldots, I_n$. The *interaction function* f represents in which way the intermediate effects I_k, and indirectly also the causes C_k, interact. Hence, this function f is defined in such way that when a relationship between the I_k's and $E = \top$ is satisfied, then it holds that $e = f(I_1, \ldots, I_n)$. Under this condition, it is assumed that $\Pr(e \mid I_1, \ldots, I_n) = 1$; otherwise, when $f(I_1, \ldots, I_n) = \bar{e}$, it holds that $\Pr(e \mid I_1, \ldots, I_n) = 0$. Using information from the topology of the network, the notion of causal independence can be formalised for the occurrence of effect E, i.e. $E = \top$, in terms of probability theory as follows:

$$\Pr(e \mid C_1, \ldots, C_n) = \sum_{f(I_1,\ldots,I_n)=e} \prod_{k=1}^{n} \Pr(I_k \mid C_k) \tag{1}$$

Based on the assumptions above, it also holds that

$$\Pr(e \mid C_1, \ldots, C_n) = \sum_{I_1,\ldots,I_n} \Pr(e \mid I_1, \ldots, I_n) \prod_{k=1}^{n} \Pr(I_k \mid C_k) \tag{2}$$

Finally, it is assumed that $\Pr(i_k \mid \bar{c}_k) = 0$ (absent causes do not contribute to the effect); otherwise, $\Pr(I_k \mid C_k) > 0$.

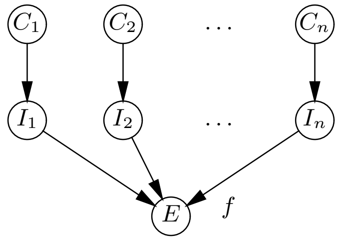

Figure 2. Causal-independence model.

An important subclass of causal-independence models is obtained if the deterministic function f is defined in terms of separate binary functions g_k; it is then called a *decomposable* causal-independence model [3]. Usually, all functions $g_k(I_k, I_{k+1})$ are identical for each k. Typical examples of decomposable causal-independence models are the noisy-OR [1, 4, 6] and noisy-MAX [1, 3] models, where the function g represents a logical OR and a MAX function, respectively.

2.2.2 Boolean functions

The function f in equation (1) is actually a Boolean function, which can also be represented by the probabilities $\Pr(e \mid I_1, \ldots, I_n)$ in equation (2). Recall that there are 2^{2^n} different n-ary Boolean functions [2]; hence, the number of possible causal-independence models is huge.

As mentioned above, in the case of causal independence it is usually assumed that the function f is decomposable, and that all binary functions g_k of which f is composed are identical. As there are 16 different binary Boolean functions, and a causal-independence model contains at least two causes, there are at least 16 n-ary Boolean functions, with $n \geq 2$, in that case. These Boolean functions can be interpreted as Boolean expressions of the form $\mathcal{I}_n = I_1 \odot \cdots \odot I_n = E$, if $\odot$ is a binary, associative Boolean operator. Not every binary Boolean operator is associative; also note that a Boolean operator $\odot$ need not be commutative either, and hence $i_1 \odot \mathcal{I}_{n-1} = \mathcal{I}_{n-1} \odot i_1$, where $\mathcal{I}_{n-1} = I_2 \odot \cdots \odot I_n$, need not hold. Table 1 indicates which of the operators are commutative and associative, and which are not.

We return to our example Bayesian-network model shown in Figure 1. If we assume that the bacteria A, B and C are all pathogenic, and thus give rise to an infectious response if the patient becomes colonised by them, the interaction among the 'Body response' variables could be modelled by a logical OR. This expresses the idea that an infection must be caused by one or more pathogenic bacteria.

2.3 Qualitative probabilistic networks

Qualitative probabilistic networks (QPNs) are qualitative abstractions of Bayesian networks [7, 8]. Instead of conditional probabilities, a QPN associates signs with its digraph, which serve to capture the probabilistic influences and synergies between its variables.

A *qualitative probabilistic influence* between two variables expresses how the values of one variable influence the probabilities of the values of the other variable. For example, a *positive qualitative influence* of a variable A on its effect B, denoted $S^+(A, B)$, expresses that observing the value $\top$ for A makes the value $\top$ for B more likely, regardless of any other direct influences on B, that is,

$$\Pr(b \mid ax) \geq \Pr(b \mid \bar{a}x) \tag{3}$$

Table 1. The binary Boolean operators.

Commutative, associative operators	
$\wedge$	and
$\vee$	or
$\leftrightarrow$	bi-implication
$\oslash$	xor
$\top$	always true
$\bot$	always false
Commutative, non-associative operators	
$\downarrow$	nor
$\mid$	nand
Non-commutative, associative operators	
p_1	projection to the first argument
p_2	projection to the second argument
n_1	negation of first argument
n_2	negation of second argument
Non-commutative, non-associative operators	
$\rightarrow$	implication
$\leftarrow$	reverse implication
$<$	increasing order
$>$	decreasing order

for any combination of values x for the set $\pi(B) \setminus \{A\}$ of causes of B other than A. A *negative qualitative influence*, denoted $S^-(A, B)$, and a *zero qualitative influence*, denoted $S^0(A, B)$, are defined analogously, replacing $\geq$ in the above formula by $\leq$ and $=$, respectively. If the influence of A on B is non-monotonic, that is, the sign of the influence depends upon the values of other causes of B, or unknown, we say that the influence is *ambiguous*, denoted $S^?(A, B)$. With each arc in a qualitative network's digraph an influence is associated.

In addition to influences, a QPN includes *synergies* modelling interactions between influences. An *additive synergy* between three variables expresses how the values of two variables jointly influence the probabilities of the values of the third variable. For example, a *positive additive synergy* of the variables A and B on their common effect C, denoted $Y^+(\{A, B\}, C)$, expresses that the joint influence of A and B on C is greater than the sum of their separate influences, regardless of any other influences on C, that is,

$$\Pr(c \mid abx) + \Pr(c \mid \bar{a}\bar{b}x) \geq \Pr(c \mid a\bar{b}x) + \Pr(c \mid \bar{a}bx) \qquad (4)$$

for any combination of values x for the set of causes of C other than A and B. *Negative*, *zero*, and *ambiguous additive synergy* are defined analogously. A qualitative network specifies an additive synergy for each pair of causes and their common effect in its digraph.

A *product synergy* between three variables expresses how the value of one variable influences the probabilities of the values of another variable in view of an observed value for the third variable [5]. For example, a *negative product synergy* of a variable A on a variable B given the value $\top$ for their common effect C, denoted $X^-(\{A, B\}, c)$, expresses that, given c, the value $\top$ for A renders the value $\top$ for B less likely, that is,

$$\Pr(c \mid abx) \cdot \Pr(c \mid \bar{a}\bar{b}x) \leq \Pr(c \mid a\bar{b}x) \cdot \Pr(c \mid \bar{a}bx) \qquad (5)$$

for any combination of values x for the set of causes of C other than A and B. *Positive*, *zero*, and *ambiguous product synergy* again are defined analogously. A QPN also specifies a product synergy for when the effect is false. Upon observation of a specific value for a common effect of two causes, the associated product synergy induces an influence between the two causes; the sign of this influence equals the sign of the synergy. A qualitative influence that is thus induced by a product synergy is termed an *intercausal influence*.

3 QUALITATIVE DESCRIPTION OF CAUSAL INTERACTION

We next use QPNs to analyse and describe the interactions obtained by various interaction functions f. We start by considering the qualitative influences among cause and effect variables, which is followed by an analysis of synergies. Together, qualitative influence and synergies constitute a QC pattern for an interaction function.

3.1 Analysis of qualitative influences

Qualitative influences are investigated by considering the expression

$$\Pr(e \mid C_1, \ldots, c_j, \ldots, C_n) - \Pr(e \mid C_1, \ldots, \bar{c}_j, \ldots, C_n) \qquad (6)$$

which results from expression (3), and is denoted by $\delta_j(C_1, \ldots, C_{j-1}, C_{j+1}, \ldots, C_n)$. The sign of the latter function determines the sign σ of the qualitative influence $S^\sigma(C_j, E)$.

The following equation, obtained by using equations (2), enables us to investigate qualitative influences in detail:

$$\delta_j(C_1, \ldots, C_{j-1}, C_{j+1}, \ldots, C_n) =$$
$$\Pr(i_j \mid c_j) \left[\sum_{I_1, \ldots, I_n \setminus I_j} d(\mathcal{I}_n \setminus I_j) \prod_{\substack{k = 1 \\ k \neq j}}^{n} \Pr(I_k \mid C_k) \right]$$

where $\mathcal{I}_n = I_1, \ldots, I_n$, and

$$d(\mathcal{I}_n \setminus I_j) = \Pr(e \mid I_1, \ldots, i_j, \ldots, I_n) - \Pr(e \mid I_1, \ldots, \bar{\imath}_j, \ldots, I_n) \qquad (7)$$

Recall that a probability distribution $\Pr(E \mid I_1, \ldots, I_n)$ represents a Boolean function. The multipliers $\prod_{k=1, k\neq j}^{n} \Pr(I_k \mid C_k)$ are responsible for possible variation among signs of the difference (6), as the difference $d(\mathcal{I}_n \setminus I_j)$ is not dependent of the cause variables C_k.

The analysis starts by considering commutative, associative operators. Note that in that case it is permitted, without loss of generality, to focus the analysis on an arbitrary (cause) variable C_j. Due to lack of space, we will only give one proof to illustrate the approach taken.

Proposition 1 *Let $\mathcal{B} = (G, \Pr)$ be a Bayesian network representing a causal-independence model with decomposable interaction function f that is equal to the bi-implication. Then, $S^?(C_j, E)$ holds for any cause variable C_j and the given effect variable E.*

Proof: Let $\mathcal{I}_{n-1} = \mathcal{I}_n \setminus I_j$, then $(i_j \leftrightarrow \mathcal{I}_{n-1}) \wedge (\bar{\imath}_j \leftrightarrow \mathcal{I}_{n-1}) \equiv \bot$. Therefore, difference (7) is either equal to 1 or to -1, and hence the sign of (6) is ambiguous in general. □

Table 2 summarises the results for all the operators which are commutative and associative. The results for the commutative, non-associative operators are omitted. However, note that as some arguments of the $\downarrow$ and $\mid$ operators will be negated when using their definitions in terms of $\vee$ and $\wedge$, it can be predicted that their signs will be either positive or negative, depending on argument position.

For the Boolean operators which are associative but non-commutative, a distinction must be made between the situation where the cause variable C_j is at the first, last or any other argument position. For the operators which are neither commutative nor associative, a distinction must be made between whether the operator is assumed to be left or right associative. The results are summarised in Tables 3 and 4. Note that argument position may affect the sign of the resulting influence. With the proviso that the first and last argument,

Table 2. Signs of qualitative influences for the commutative, associative operators.

Operator	Sign
$\wedge$	$+$
$\vee$	$+$
$\leftrightarrow$	?
$\oslash$	?
$\top$	0
$\bot$	0

Table 3. Signs of qualitative influences for the non-commutative, associative operators.

	Sign	
Operator	First	Non-first
p_1	$+$	0
n_1	$-$	0
Operator	Last	Non-last
p_2	$+$	0
n_2	$-$	0

and RA and LA need to be swapped, the results for the operators $\leftarrow$ and $>$ are identical to those of $\rightarrow$ and $<$. This also holds for the additive and product synergies discussed in the remainder of the paper.

We return to our example in Figure 1. It is known that some bacteria may protect a host against infection. Suppose that this holds for bacteria A and B, then each of these would make the development of infection less likely, even though there could be circumstances where these bacteria turn pathogenic. Now, let C be a bacterium with only pathogenic strains, then the right-associative version of implication (Table 4) would model this situation appropriately.

3.2 Analysis of additive synergies

Recall that in the case of causal independence, additive synergies describe how two causes jointly influence the probability of the effect variable. Using definition (4) of an additive synergy—moving the right-hand side to the left—and equation (1), and considering interactions between the causes C_{j-1} and C_j, we obtain:

$$\delta_{j-1,j}(C_1, \ldots, C_{j-2}, C_{j+1}, \ldots, C_n) = \sum_{f(I_1,\ldots,I_n)=e} d(I_{j-1}, I_j) \prod_{k=1}^{j-2} \Pr(I_k \mid C_k) \prod_{k=j+1}^{n} \Pr(I_k \mid C_k)$$

where

$$d(I_{j-1}, I_j) = \Pr(I_{j-1} \mid c_{j-1}) \Pr(I_j \mid c_j) + \Pr(I_{j-1} \mid \bar{c}_{j-1}) \Pr(I_j \mid \bar{c}_j) - \Pr(I_{j-1} \mid c_{j-1}) \Pr(I_j \mid \bar{c}_j) - \Pr(I_{j-1} \mid \bar{c}_{j-1}) \Pr(I_j \mid c_j)$$

Let $\Pr(i_{j-1} \mid c_{j-1}) = p$ and $\Pr(i_j \mid c_j) = q$, then the value $d(I_{j-1}, I_j)$ is given in Table 5 for different values of I_{j-1} and I_j. The following equation is then obtained:

$$\delta_{j-1,j}(C_1, \ldots, C_{j-2}, C_{j+1}, \ldots, C_n) =$$

Table 4. Signs of qualitative influences for the non-commutative, non-associative operators; RA: right associative; LA: left associative.

	Sign for RA		Sign for LA	
Operator	Last	Non-last	Last	Non-last
$\rightarrow$	$+$	$-$	$+$	?
$<$	$+$	$-$	$+$	?

Table 5. Difference $d(I_1, I_2)$ for various values of the variables I_1 and I_2.

I_1	I_2	$d(I_1, I_2)$
i_1	i_2	pq
$\bar{\imath}_1$	i_2	$-pq$
i_1	$\bar{\imath}_2$	$-pq$
$\bar{\imath}_1$	$\bar{\imath}_2$	pq

$$\sum_{I_1,\ldots,I_n \setminus I_{j-1}, I_j} \sum_{\substack{I_{j-1}, I_j \\ f(I_1, \ldots, I_n) = e}} \sigma(I_{j-1} \oslash I_j)\, pq \cdot \prod_{k=1}^{j-2} \Pr(I_k \mid C_k) \prod_{k=j+1}^{n} \Pr(I_k \mid C_k) \quad (8)$$

where $\oslash$ represents the exclusive OR, and

$$\sigma(Q) = \begin{cases} -1 & \text{if } Q \equiv \top \\ 1 & \text{otherwise} \end{cases}$$

The multipliers $\prod_{k=1}^{j-2} \Pr(I_k \mid C_k) \prod_{k=j+1}^{n} \Pr(I_k \mid C_k) \geq 0$, will generally differ for different $\delta_{j-1,j}(C_1, \ldots, C_{j-2}, C_{j+1}, \ldots, C_n)$. The sum of terms $\sigma(I_{j-1} \oslash I_j)\, pq$ will not; which of those terms will actually be included in the final sum is determined by the function f.

As before, a distinction has to be made between operators that are associative and commutative, or not. Again, the proof for only one of the Boolean operators is given; the results for the two commutative, non-associative operators are again omitted.

Table 6. Signs of additive synergies for the commutative, associative operators.

Operator	Sign
$\wedge$	$+$
$\vee$	$-$
$\leftrightarrow$	?
$\oslash$	$-$
$\top$	0
$\bot$	0

Proposition 2 *Let $\mathcal{B} = (G, \Pr)$ be a Bayesian network representing a causal-independence model with decomposable interaction function f that is equal to projection to the first argument. Then, it holds that $Y^0(\{C_{j-1}, C_j\}, E)$ for any two cause variables C_{j-1}, C_j and the given effect variable E.*

Proof: The interaction function f is equivalent to $(I_1\ p_1\ I_2\ p_1\ \cdots\ p_1\ I_n) \equiv I_1$. Now, if $j > 2$, then $d(I_{j-1}, I_j)$ will be computed for every value of I_{j-1} and I_j, and hence, summing these results yields 0. If $j = 2$, the sum is taken only over $d(i_1, i_2)$ and $d(i_1, \bar{\imath}_2)$, which also yields 0. □

The proofs for the other non-commutative, associative operators are similar, with results that are in fact identical. The reason for this is that the operators select at most one argument, and hence, either all 4 possible Boolean combinations of the two interaction variables if the selected variable is not among them, or two combinations of Boolean values, with one of them fixed, need to be considered. In both cases, there are an equal number of terms pq and $-pq$, which cancel out each other.

The analysis of the non-commutative and non-associative operators is more difficult, as again a distinction must be made between assuming the operators to be right associative or left associative. Due to lack of space the results are summarised in Table 7 without proof.

Table 7. Signs of additive synergies for the non-commutative, non-associative operators; RA: right-associative; LA: left-associative.

	Sign for RA		Sign for LA	
Operator	Last	Non-last	Last	Non-last
$\rightarrow$	$+$	$-$	$+$	$-$
$<$	$-$	$+$	$-$	$+$

We return to our example in Figure 1. In the previous section, the individual effects, but not the synergies, of the colonisation by bacteria A, B and C on the patient's body response were modelled. It appears that the right-associative version of implication also rightly expresses that colonisation by both bacterium A and B makes development of infection less likely, whereas bacterium C is so pathogenic that it overrides the preventive effects of bacteria A and B.

3.3 Analysis of product synergies

We basically use here for product synergies an approach similar to the one employed in the previous section for additive synergies. For the analysis of product synergies, the equation of interest is:

$$\delta^E_{j-1,j}(C_1,\ldots,C_{j-2},C_{j+1},\ldots,C_n) = \sum_{I_1,\ldots,I_n\backslash I_{j-1},I_j} \{t(c_{j-1},c_j;\mathcal{I}_n\backslash I_{j-1},I_j)\cdot t(\bar{c}_{j-1},\bar{c}_j;\mathcal{I}_n\backslash I_{j-1},I_j) - t(c_{j-1},\bar{c}_j;\mathcal{I}_n\backslash I_{j-1},I_j)\cdot t(\bar{c}_{j-1},c_j;\mathcal{I}_n\backslash I_{j-1},I_j)\}\cdot \prod_{k=1}^{j-2}\Pr(I_k \mid C_k)\prod_{k=j+1}^{n}\Pr(I_k \mid C_k) \qquad (9)$$

where

$$t(C_{j-1},C_j;\mathcal{I}_n\backslash I_{j-1},I_j) = \sum_{\substack{I_{j-1},I_j \\ f(I_1,\ldots,I_n)=E}} \Pr(I_{j-1}\mid C_{j-1})\Pr(I_j\mid C_j)$$

The arithmetic expression between the braces is the essential element in the analysis below; it will be denoted by β. Furthermore, we will once more use the abbreviations $p = \Pr(i_{j-1} \mid c_{j-1})$ and $q = \Pr(i_j \mid c_j)$. As before, for the operators which are commutative and associative, we will focus the analysis on two arbitrary cause variables C_{j-1} and C_j, here, for simplicity's sake, the interaction of the two equally arbitrary variables C_1 and C_2. We present the proof for the logical OR.

Proposition 3 *Let $\mathcal{B} = (G,\Pr)$ be a Bayesian network representing a causal-independence model with decomposable interaction function f that is equal to the logical OR. Then, it holds that $X^-(\{C_{j-1},C_j\},e)$ for any two cause variables C_{j-1},C_j given that the effect is true; and $X^0(\{C_{j-1},C_j\},\bar{e})$ when the effect is assumed to be false.*

Proof: Let the interaction function be represented by the Boolean expression $\mathcal{I}_n = I_1 \vee I_2 \vee \mathcal{I}_{n-2}$. First, we consider the situation where $E=\top$. There are two cases to consider. Let $\mathcal{I}_{n-2}\equiv\bot$, then $\beta = 0 - pq = -pq$. For $\mathcal{I}_{n-2}\equiv\top$, we get $\beta = 1-1 = 0$. So, summing over $I_3,\ldots,I_n$ yields $\sum -pq\cdot\prod_{k=3}^{n}\Pr(I_k \mid C_k)\leq 0$. We conclude that $X^-(\{C_{j-1},C_j\},e)$ holds.

Next, consider $E=\bot$. This implies that both I_1 and I_2 must be false. We get $\beta=(1-p)(1-q)-(1-p)(1-q)=0$; this means that $\delta^{\bar{e}}_{1,2}(C_3,\ldots,C_n)=0$, and thus $X^0(\{C_{j-1},C_j\},\bar{e})$ holds. □

Table 8. Signs of product synergies for the non-commutative, non-associative operators for e; RA: right-associative; LA: left-associative.

	Sign for RA		Sign for LA	
Operator	Last	Non-last	Last	Non-last
$\rightarrow$	$+$	$-$	$+$	$+$
$<$	0	0	0	$+$

Due to lack of space, we have omitted the (other) results for the commutative, (non)associative operators, and the tables for $E=\bot$. For all the non-commutative, associative operators we obtain a zero product synergy. The remaining results are given in Table 8.

We return to the example in Figure 1. Now, assume that there is a patient in hospital having an infectious disease. Recall that bacteria A and B are known to be not particularly pathogenic, whereas bacterium C is. Assuming that the patient is colonised with bacterium C makes it more likely that the patient is colonised with A or B, as the infection is strong evidence that the conditions for colonisation have been met. On the other hand, when we assume that the patient is being colonised by bacterium A (or B), and we use these to explain the infection in the patient, it is *less* likely that the patient is colonised by the other bacteria. This is because we are dealing here with a pathogenic strain of bacterium A (or B), causing the infection. This probabilistic behaviour is again appropriately modelled by the right-associative version of implication.

4 DISCUSSION

The qualitative characteristics of interactions in Bayesian networks have been analysed and described in this paper, taking causal independence and QPNs as a basis. The integration of these two approaches into a coherent theory is the major, novel scientific contribution of this paper. By determining the signs of the relations S, Y and X for a specific interaction function f, we obtain the qualitative, causal pattern or QC pattern for the function. The number of different QC patterns that emerged was limited, as Boolean functions were defined in terms of single binary Boolean operators. As a consequence, only a fraction of the possible QC patterns may have been identified. As different Boolean functions may yield the same QC pattern, it is as yet unknown whether all possible QC patterns can be realised. This is something that requires further research. Another important topic of future research is to associate clear conceptual meanings with the various QC patterns, such that they can be easily understood and used by Bayesian-network developers.

REFERENCES

[1] F.J. Díez, Parameter adjustment in Bayes networks: the generalized noisy OR-gate, in: *UAI-9*, 1993, pp. 99–105.

[2] H.B. Enderton, *A Mathematical Introduction to Logic*. Academic Press, San Diego, 1972.

[3] D. Heckerman and J.S. Breese, Causal independence for probabilistic assessment and inference using Bayesian networks, *IEEE Transactions on Systems, Man and Cybernetics* 26 (6), 1996, pp. 826–831.

[4] M. Henrion, Some practical issues in constructing belief networks, in: *UAI-3*, Elsevier, Amsterdam, 1989, pp. 161–173.

[5] M. Henrion and M.J. Druzdzel, Qualitative propagation and scenario-based approaches to explanation in probabilistic reasoning, in: *UAI-6*, 1991, pp. 17–32.

[6] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, Palo Alto, 1988.

[7] S. Renooij, *Qualitative Approaches to Quantifying Probabilistic Networks*. PhD Thesis, Utrecht University, 2001.

[8] M.P. Wellman, Fundamental concepts of qualitative probabilistic networks. *Artificial Intelligence*, 44, 1990, pp. 257 – 303.