Bayesian Network Modelling through Qualitative Patterns*

Peter J.F. Lucas
Institute for Computing and Information Sciences
Radboud University Nijmegen, Toernooiveld 1
6525 ED Nijmegen, The Netherlands
E-mail: peterl@cs.kun.nl

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. Qualitative probabilistic networks (QPNs) have been put forward as qualitative analogues to Bayesian networks, and allow modelling interactions in terms of qualitative signs. They thus have the advantage that developers can abstract from the numerical detail, and therefore the gap may not be as wide as for their quantitative counterparts. 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. In the paper, we deploy both causal independence and 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.

Keywords & Phrases: Bayesian networks, knowledge representation, qualitative reasoning.

1 Introduction

Reasoning with uncertainty is a significant area of research in Artificial Intelligence at least since the early 1970s. Many different methods for representing and reasoning with uncertain knowledge have been developed during the last three decades, including the certainty-factor calculus [2, 22], Dempster-Shafer theory [21], possibilistic logic [8], fuzzy logic [29], and Bayesian networks, also called belief networks and causal probabilistic networks [3, 19, 18]. During the last decade a gradual shift towards the use of probability theory as the foundation of almost all of the work in this area could be observed, mainly due to the impact, both theoretically and practically, of the introduction of Bayesian networks and related graphical probabilistic models into the field.

Bayesian networks offer a powerful framework for the modelling of uncertain interactions among variables in a given domain. Such interactions are represented in two different manners: firstly, in a qualitative manner, by means of a directed acyclic graph, and secondly, in a quantitative manner, by specifying a conditional probability distribution for every variable

*Published in: Artificial Intelligence, 163; 233–263, 2005.

represented in the network. These conditional probability distributions allow for expressing various logical, functional and probabilistic relationships among variables. Much of the appeal of the Bayesian network formalism derives from this feature (cf. [3] for a modern, technical overview).

It is well known that ensuring that the graph 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 identified special types of independence relationships in order to facilitate the process of probability assessment. In particular the theory of *causal independence* fulfils this purpose [15]. The theory allows for the specification of the interactions among variables in terms of cause-effect relationships, adopting particular statistical independence assumptions. Causal independence is frequently used in the construction of practical networks for situations where the underlying probability distributions are complex. The theory has also been exploited to increase the efficiency of probabilistic inference in Bayesian networks [30, 31]. 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; in both cases, interactions among variables are modelled as being disjunctive [4, 16, 19].

Qualitative probabilistic networks offer a qualitative analogue to the formalism of Bayesian networks. They allow describing the dynamics of the interaction among variables in a purely qualitative fashion by means of the specification and propagation of qualitative signs [20, 28]. Hence, qualitative probabilistic networks abstract from the numerical detail, yet retain the qualitative semantics underlying Bayesian networks. The theory of qualitative probabilistic networks, therefore, seems to offer potentially useful tools for the qualitative analysis of Bayesian networks.

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. Such a theory could assist developers of systems based on Bayesian networks in designing such networks, exploiting the qualitative information that is available in the domain concerned as much as possible. In the paper, various interaction types are defined using Boolean algebra; qualitative probabilistic networks 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.

The remainder of this paper is organised as follows. In the following section, the basic properties of Bayesian networks are introduced, as are Boolean functions, and the notions 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 qualitative probabilistic networks in Section 3. Section 4 summarises the various patterns that have been obtained, and discusses these results in the context of all possible patterns. Finally, in Section 5, it is summarised what has been achieved by this research.

2 Preliminaries

To start, the basic theory of Bayesian networks, causal independence and qualitative probabilistic networks are reviewed.

2.1 Bayesian networks

A *Bayesian network* is a concise representation of a joint probability distribution on a set of statistical variables [19]. It consists of a qualitative part and an associated quantitative part. The qualitative part is a graphical representation of the interdependences between the variables in the encoded distribution. It takes the form of an acyclic directed graph (digraph) $G = (V(G), A(G))$, where each node $V \in V(G)$ corresponds to a statistical variable that takes one of a finite set of values, and $A(G) \subseteq V(G) \times V(G)$ is a set of arcs. In this paper, we assume all variables to be binary; for abbreviation, we will often use v to denote $V = \top$ (true) and $\bar{v}$ to denote $V = \bot$ (false). Sometimes, we prefer to leave the specific value of a variable open (i.e. it is taken as a free variable), and then we simply state V. In other cases, we use this notation when a variable is actually bound. The context will make clear which interpretation is intended. Furthermore, for abbreviation, we use the notation $V_1, \ldots, V_n \backslash V_i, \ldots, V_j$ which 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

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

stands for summing over $g(I_1, \ldots, I_n)$ for all possible values of the variables I_k for which the constraint $\psi(I_1, \ldots, I_n) = e$ holds. However, if we refer to variables separate from such constraints, such as in

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

then we only sum over the separately mentioned variables, here the variables I_1, I_2, and the equality only acts as a constraint.

The arcs $A(G)$ in the digraph G model possible dependences between the represented variables. Informally speaking, we take an arc $V \to V'$ between the nodes V and V' to represent an influential relationship between the associated variables V and V'. If this arc is given a causal reading, then the arc's direction marks V' as the *effect* of the *cause* V. Absence of an arc between two nodes means that the corresponding variables do not influence each other directly and, hence, are (conditionally) independent. In the following, causes will often be denoted by C_i and their associated effect variable by E.

Associated with the qualitative part of a Bayesian network are numerical quantities from the encoded probability distribution. With each variable V in the digraph is associated a set of *conditional probabilities* $\Pr(V \mid \pi(V))$, describing the joint influence of values for the parents $\pi(V)$ of V on the probabilities of the variable V's values. These sets of probabilities constitute the quantitative part of the network. A Bayesian network represents a joint probability distribution on its variables and thus provides for computing any probability of interest. Various algorithms for probabilistic inference with a Bayesian network are available [19, 30, 23].

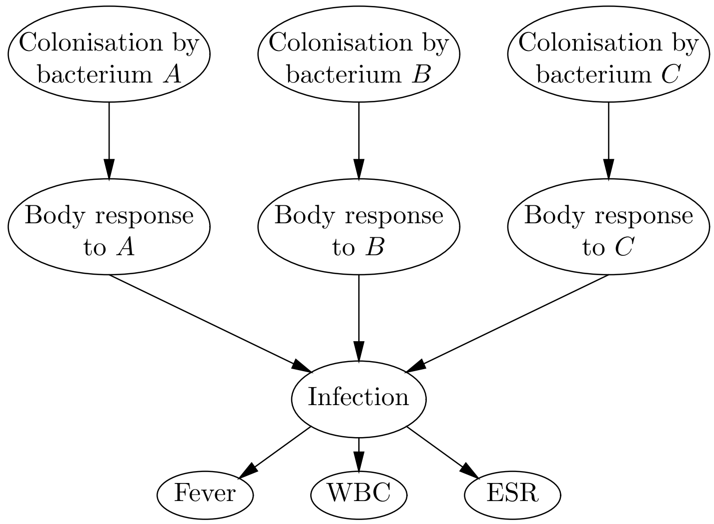

Figure 1: Example Bayesian networks, modelling the interaction among bacteria possibly causing an infection in patients after colonisation.

Bayesian networks are successfully applied in a growing number of fields; biomedical applications in particular have attracted a great deal of research activity (cf. [1, 11, 12, 24, 25]). This may be due to the fact that biological mechanisms can often be described quite naturally in causal terms. Consider, for example, the causal network shown in Figure 1, which models the causal mechanisms by which patients become colonised by bacteria, for example *Pseudomonas 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, depending on the bacteria concerned; in the end an infection may develop. An infection is clinically recognised by signs and symptoms 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; the actual type of interaction depends on the bacteria involved.

As a second example, consider the interaction between bactericidal antimicrobial agents, i.e. drugs that kill bacteria by interference with their metabolism, and bacteriostatic antimicrobial agents, i.e. drugs that inhibit the multiplication of bacteria. Penicillin is an example of a bactericidal drug, whereas chlortetracyclin is an example of a bacteriostatic drug. It is well known among medical doctors that the interaction between bactericidal and bacteriostatic drugs can have antagonistic effects; e.g. the drug combination penicillin and chlortetracyclin may be have as little effect against an infection as prescribing no antimicrobial agent at all, even if the bacteria are susceptible to each of these drugs. Note that here we interpret drugs as statistical variables, not as decision variables as in clinical decision making. The depiction of the causal interaction of the relevant variables is shown in Figure 2; note the similarity in structure of this network in comparison to Figure 1.

As a last example, this time not concerning infectious disease, consider the interaction between natural hormones that have partially related, but possibly opposite, working mechanisms, such as insulin and glucagon: two hormones that are involved in the regulation of

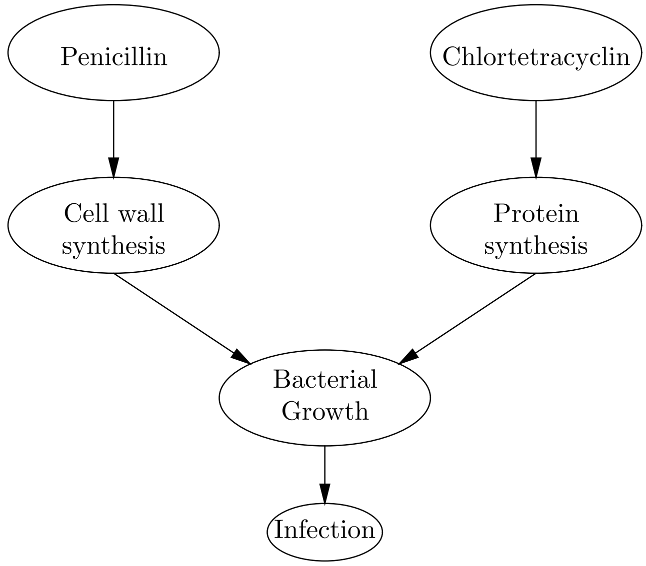

Figure 2: Example Bayesian network, modelling the interaction between the antimicrobial agents penicillin and chlortetracyclin on infection.

glucose levels in the blood. Insulin is needed to let glucose cross the membrane of most of the body cells (exceptions are the brain cells, where, as a protective mechanisms, glucose transfer is not insulin dependent) so that it can be utilised as fuel in the cell metabolism. In this way glucose is transferred from blood to cytoplasm. Glucagon, on the other hand, stimulates the release of glucose from the glycogen deposites, such as the liver, into the blood. In order for glucagon to be effective, it is necessary that insulin is present, as otherwise there will be little glucose stored in the body cells as glycogen. Too high levels of insulin, insulin hypersecretion, as may occur in tumours called insulinomas, may give rise to hypoglycaemia, i.e. abnormally low glucose levels in the blood. In all other cases, levels of glucose in the blood will not be abnormally low (although the levels may be too high, but this is not considered an acute danger). The causal interaction of the relevant variables is shown in Figure 3.

Although the Bayesian networks shown in figures 1, 2 and 3 have a very similar structure, their underlying interaction semantics is very different as we will see below. These networks will be used in the following as running examples to illustrate some of the results.

2.2 Causal independence

In this section, we introduce a type of cause-effect interaction, called causal independence, which essentially is a causal model with rather strong independence assumptions.

2.2.1 Probabilistic representation

One popular way to specify interactions among statistical variables in a compact fashion is offered by the notion of *causal independence* [10, 13, 14, 15]. The global structure of a causal-independence model is shown in Figure 4; it expresses the idea that causes $C_1, \ldots, C_n$ influence a given common effect E through intermediate variables $I_1, \ldots, I_n$ and a deterministic function f, called the *interaction function*. The influence of each cause C_k on the common effect E is independent of each other cause C_j, $j \neq k$. The function f represents in which way

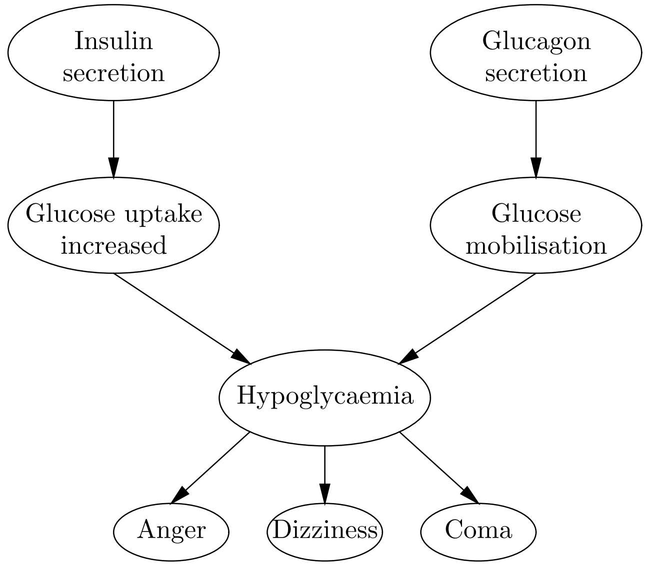

Figure 3: Example Bayesian network, modelling the interaction between insulin and glucagon secretion from hormonal gland tissue.

the intermediate effects I_k, and indirectly also the causes C_k, interact to yield a final effect E. Hence, this function f is defined in such way that when a relationship, as modelled by the function f, between I_k, $k = 1, \ldots, n$, and $E = \top$ is satisfied, then it holds that $e = f(I_1, \ldots, I_n)$.

In terms of probability theory, the notion of causal independence can be formalised for the occurrence of effect E, i.e. $E = \top$, as follows:

$$\Pr(e \mid C_1, \ldots, C_n) = \sum_{f(I_1,\ldots,I_n)=e} \Pr(e \mid I_1, \ldots, I_n) \Pr(I_1, \ldots, I_n \mid C_1, \ldots, C_n) \qquad (1)$$

meaning that the causes $C_1, \ldots, C_n$ influence the common effect E through the intermediate effects $I_1, \ldots, I_n$ only when $e = f(I_1, \ldots, I_n)$ for certain values of I_k, $k = 1, \ldots, n$. 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 $\Pr(e \mid I_1, \ldots, I_n) = 0$. Note that the effect variable E is conditionally independent of $C_1, \ldots, C_n$ given the intermediate variables $I_1, \ldots, I_n$, and that each variable I_k is only dependent on its associated variable C_k; hence, it holds that

$$\Pr(e \mid I_1, \ldots, I_n, C_1, \ldots, C_n) = \Pr(e \mid I_1, \ldots, I_n)$$

and

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

Formula (1) can now be simplified to:

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

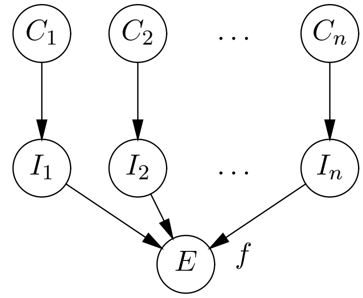

Figure 4: Causal independence model.

Based on the assumptions above, it also holds that

$$\Pr(e \mid C_1, \ldots, C_n) = \sum_{I_1,\ldots,I_n} \Pr(e \mid I_1, \ldots, I_n) \prod_{k=1}^{n} \Pr(I_k \mid C_k) \quad (3)$$

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

Formula (2) is practically speaking not very useful, because the size of the specification of the function f is exponential in the number of its arguments. The resulting probability distribution is therefore in general computationally intractable, both in terms of space and time requirements. An important subclass of causal independence models, however, is formed by models in which the deterministic function f can be defined in terms of separate binary functions g_k, also denoted by $g_k(I_k, I_{k+1})$. Such causal independence models have been called *decomposable* causal independence models [14]; these models are of significant practical importance. Usually, all functions $g_k(I_k, I_{k+1})$ are identical for each k; a function $g_k(I_k, I_{k+1})$ may therefore be simply denoted by $g(I, I')$. Typical examples of decomposable causal independence models are the noisy-OR [4, 10, 16, 19, 26] and noisy-MAX [4, 15, 26] models, where the function g represents a logical OR and a MAX function, respectively.

2.2.2 Boolean functions

The function f in equation (2) is actually a Boolean function; recall that there are 2^{2^n} different n-ary Boolean functions [9, 27]. Hence, the potential number of causal interaction models is huge. The Boolean functions can also be represented by the probabilities

$$\Pr(e \mid I_1, \ldots, I_n)$$

in equation (3), with $\Pr(e \mid I_1, \ldots, I_n) \in \{0, 1\}$.

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 interaction model contains at least two causes, there are at least 16 n-ary Boolean functions, with $n \geq 2$, in that case. Some of these Boolean functions can be interpreted as a Boolean expression of the form

$$I_1 \odot \cdots \odot I_n = E$$

where $\odot$ is a binary, associative Boolean operator. However, not every binary Boolean operator is associative; Table 1 mentions which operators are associative and which are not.

Table 1: The binary Boolean operators.

Commutative, associative operators	
$\wedge$	AND
$\vee$	OR
$\leftrightarrow$	bi-implication
$\oslash$	XOR, exclusive OR
$\top$	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

As a matter of notation, in the following we will frequently make use of the abbreviation:

$$\mathcal{I}_j = (\cdots(I_1 \odot I_2) \odot \cdots) \odot I_{j-1}) \odot I_j) \tag{4}$$

if it is assumed that the Boolean operator $\odot$ is left associative. Similarly, the notation

$$\mathcal{I}^j = (I_j \odot (I_{j+1} \odot (\cdots \odot (I_{n-1} \odot I_n) \cdots) \tag{5}$$

is used if it is assumed that the operator $\odot$ is right associative. Note that $\mathcal{I}_n \equiv \mathcal{I}^1$ if $\odot$ is associative. In that case, we will simply use the notation $\mathcal{I}_{n-1}$ to denote the Boolean expression with one variable less than $\mathcal{I}_n$, as usually will become clear from the context. Finally, note that a Boolean operator $\odot$ need not be commutative, 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 also indicates which of the operators are commutative and which are not.

The commutative and associative binary operators mentioned in Table 1 give rise to Boolean expressions that are special cases of symmetric Boolean functions (the two commutative, non-associative operators are only symmetric for two arguments). A Boolean function f is *symmetric* if

$$f(I_1, \ldots, I_n) = f(I_{j_1}, \ldots, I_{j_n})$$

for any index function $j : \{1, \ldots, n\} \to \{1, \ldots, n\}$ [27]. An example of a symmetric Boolean functions is the *exact* Boolean function e_k, which is defined as:

$$e_k(I_1, \ldots, I_n) = \begin{cases} \top & \text{if } \sum_{j=1}^{n} \nu(I_j) = k \\ \bot & \text{otherwise} \end{cases} \tag{6}$$

with $k \in \mathbb{N}$, and

$$\nu(I) = \begin{cases} 1 & \text{if } I = \top \\ 0 & \text{otherwise} \end{cases}$$

Hence, this function simply checks whether there are k cases where I_j is true. The following basic properties of the exact Boolean function are useful in establishing properties of symmetric Boolean functions explored below.

Lemma 1 *Let $e_k(I_1, \ldots, I_n)$ be the exact Boolean function, then:*

(1) $\forall k \in \mathbb{N}\ \forall I_1, \ldots, I_n$: $e_k(I_1, \ldots, i_j, \ldots, I_n) \wedge e_k(I_1, \ldots, \bar{\imath}_j, \ldots, I_n) \equiv \bot$, and

(2) $\forall I_1, \ldots, I_n\ \exists k \in \mathbb{N}$: $e_k(I_1, \ldots, i_j, \ldots, I_n) \equiv e_{k-1}(I_1, \ldots, \bar{\imath}_j, \ldots, I_n)$, and

(3) $\forall I_1, \ldots, I_n\ \exists k \in \mathbb{N}\ \forall l \in \mathbb{N}, l \neq k, l \neq k-1$: $e_k(I_1, \ldots, i_j, \ldots, I_n) \vDash \neg e_l(I_1, \ldots, I_j, \ldots, I_n)$

Proof: Straight from the definition. □

Another useful symmetric Boolean function is the *threshold* function t_k, which simply checks whether there are at least k trues among the arguments:

$$t_k(I_1, \ldots, I_n) = \begin{cases} \top & \text{if } \sum_{j=1}^{n} \nu(I_j) \geq k \\ \bot & \text{otherwise} \end{cases}$$

Symmetric Boolean functions can be decomposed in terms of the exact functions e_k as follows [27]:

$$f(I_1, \ldots, I_n) = \bigvee_{k=0}^{n} e_k(I_1, \ldots, I_n) \wedge c_k \tag{7}$$

where c_k are Boolean constants only dependent of the function f. For example, for the Boolean function defined in terms of the AND operator we have: $c_0 = \cdots = c_{n-1} = \bot$ and $c_n = \top$, for the Boolean function defined in terms of the OR operator we have $c_0 = \bot$ and $c_1 = \cdots = c_n = \top$, and for the XOR operator we have that $c_k = \bot$ if *even*(k) and $c_k = \top$ if *odd*(k).

Symmetric functions are generally not decomposable in the sense of the previous section, but as the exact function e_k simply checks sums, a symmetric Boolean function can nevertheless be split into parts using equality (7).

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 can be modelled by a logical OR, $\vee$. This expresses the idea that an infection must be caused by one or more pathogenic bacteria. The interaction between penicillin and chlortetracyclin as depicted in Figure 2 can be described my means of an exclusive OR, $\oslash$, as presence of either of these in the patient's body tissues leads to a decrease in bacterial growth, whereas if both are present or absent, there will be little or no effect on bacterial growth. The interaction between insulin and glucagon secretion as shown in Figure 3 can be described my means of the decreasing order operator, $>$, as insulin hypersecretion is a cause of hypoglycaemia, but only if there is no glucagon hypersecretion. If there is only glucagon hypersecretion, we will not have hypoglycaemia, whereas if we have neither insulin hypersecretion nor glucagon hypersecretion, hypoglycaemia does not occur either. In the following we use the interaction between various types of bacteria as examples to illustrate how Boolean functions can be used to model different interactions with their associated meanings. The way penicillin and chlortetracyclin interact, as well as insulin and glucagon, are, however, kept fixed.

2.3 Qualitative probabilistic networks

Qualitative probabilistic networks, or QPNs for short, are qualitative abstractions of Bayesian networks, bearing a strong resemblance to their quantitative counterparts [28]. A qualitative probabilistic network equally comprises a graphical representation of the interdependences between statistical variables, once again taking the form of an acyclic digraph. Instead of conditional probabilities, however, a qualitative probabilistic network associates signs with its digraph. These signs serve to capture the probabilistic influences and synergies between 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,

$$\Pr(b \mid a, x) \geq \Pr(b \mid \bar{a}, x) \tag{8}$$

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.

The set of influences of a qualitative probabilistic network exhibits various convenient properties [28]. The property of *symmetry* guarantees that, if the network includes the qualitative influence $S^+(A, B)$, then it also includes $S^+(B, A)$. The property of *transitivity* asserts that the qualitative influences along a path between two variables, specifying at most one incoming arc for each variable, combine into a single compound influence between these variables with the $\otimes$-operator from Table 2. The property of *composition* further asserts that multiple qualitative influences between two variables along parallel paths combine into a compound influence between these variables with the $\oplus$-operator. In addition to influences,

Table 2: The operators for combining signs.

$\otimes$	$+$	$-$	0	$?$
$+$	$+$	$-$	0	$?$
$-$	$-$	$+$	0	$?$
0	0	0	0	0
$?$	$?$	$?$	0	$?$

$\oplus$	$+$	$-$	0	$?$
$+$	$+$	$?$	$+$	$?$
$-$	$?$	$-$	$-$	$?$
0	$+$	$-$	0	$?$
$?$	$?$	$?$	$?$	$?$

a qualitative probabilistic network 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,

$$\Pr(c \mid a, b, x) + \Pr(c \mid \bar{a}, \bar{b}, x) \geq \Pr(c \mid a, \bar{b}, x) + \Pr(c \mid \bar{a}, b, x) \tag{9}$$

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 [17]. 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,

$$\Pr(c \mid a, b, x) \cdot \Pr(c \mid \bar{a}, \bar{b}, x) \leq \Pr(c \mid a, \bar{b}, x) \cdot \Pr(c \mid \bar{a}, b, x) \tag{10}$$

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. For each pair of causes and their common effect, a qualitative probabilistic network specifies two product synergies, one for each value of the effect. 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 analysis of causal independence

Even though the notion of causal independence is described in a qualitative fashion in Section 2.2, the actual interactions obtained are determined by the interaction function f used in defining it. QPNs offer qualitative abstractions of Bayesian networks, and, thus, could serve in principle as tools for describing and analysing qualitative phenomena in Bayesian networks. This is exactly what is done in this and subsequent sections. In this section, we use QPNs to analyse and describe the interactions for various interaction functions f. We start by considering qualitative influences among cause and effect variables, which is followed by an analysis of synergies. Throughout the paper it is assumed that the number of causes n is greater than or equal to 2.

3.1 Qualitative influences

Qualitative influences are investigated by considering the sign of the expression

$$\Pr(e \mid C_1, \ldots, c_j, \ldots, C_n) - \Pr(e \mid C_1, \ldots, \bar{c}_j, \ldots, C_n) \tag{11}$$

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

The following result, obtained by using equation (3), enables us to investigate qualitative influences in detail:

$$\delta_j(C_1, \ldots, C_{j-1}, C_{j+1}, \ldots, C_n) =$$
$$\Pr(e \mid C_1, \ldots, c_j, \ldots, C_n) - \Pr(e \mid C_1, \ldots, \bar{c}_j, \ldots, C_n) =$$

$$\Pr(i_j \mid c_j) \sum_{I_1,\ldots,I_n \backslash I_j} \Pr(e \mid I_1,\ldots,i_j,\ldots,I_n) \prod_{\substack{k=1\\k\neq j}}^{n} \Pr(I_k \mid C_k) +$$

$$\Pr(\bar{\imath}_j \mid c_j) \sum_{I_1,\ldots,I_n \backslash I_j} \Pr(e \mid I_1,\ldots,\bar{\imath}_j,\ldots,I_n) \prod_{\substack{k=1\\k\neq j}}^{n} \Pr(I_k \mid C_k) -$$

$$\sum_{I_1,\ldots,I_n \backslash I_j} \Pr(e \mid I_1,\ldots,\bar{\imath}_j,\ldots,I_n) \prod_{\substack{k=1\\k\neq j}}^{n} \Pr(I_k \mid C_k) =$$

$$\Pr(i_j \mid c_j) \left[\sum_{I_1,\ldots,I_n \backslash I_j} d_e(\mathcal{I}_n \backslash I_j) \prod_{\substack{k=1\\k\neq j}}^{n} \Pr(I_k \mid C_k) \right]$$

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

$$d_e(\mathcal{I}_n \backslash I_j) = \Pr(e \mid I_1,\ldots,i_j,\ldots,I_n) - \Pr(e \mid I_1,\ldots,\bar{\imath}_j,\ldots,I_n) \tag{12}$$

Recall that it is assumed that $\Pr(i \mid \bar{c}) = 0$. 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 (11) for various values of $C_1,\ldots,C_{j-1},C_{j+1},\ldots,C_n$, as the difference (12) is not influenced by the values of cause variables C_k. As the constituents in the difference (12) represent Boolean functions, this difference can be interpreted as a mapping $\{\bot,\top\} \times \{\bot,\top\} \to \{-1,0,1\}$. Hence, from the combined effect of the multipliers and the difference it appears that any qualitative influence can be represented using causal independence.

Recall that a probability distribution $\Pr(E \mid I_1,\ldots,I_n)$ representing a Boolean function can be interpreted as a Boolean expression or function. We first consider cases where the Boolean function is symmetric; as the exact and threshold Boolean function are fundamental, we examine these two functions first.

Proposition 1 *Let $\mathcal{B} = (G,\Pr)$ be a Bayesian network representing a causal independence model with interaction function f equal to the exact function e_k, then the sign σ in $S^\sigma(C_j,E)$ is equal to '?' for $1 \leq k \leq n-1$, whereas $\sigma = -$ for $k = 0$ and $\sigma = +$ for $k = n, n > 0$.*

Proof: Lemma (1) indicates that $e_k(I_1,\ldots,i_j,\ldots,I_n) \wedge e_k(I_1,\ldots,\bar{\imath}_j,\ldots,I_n)$ is always unsatisfiable. Both expressions are satisfiable for $1 \leq k \leq n-1$, but never both at the same time according to Lemma (1), and thus it holds that $\sigma = ?$. For $k = 0$, it holds that $e_0(I_1,\ldots,i_j,\ldots,I_n) \equiv \bot$ for any truth value for $I_1,\ldots,I_{j-1},I_{j+1},I_n$, whereas $e_0(I_1,\ldots,\bar{\imath}_j,\ldots,I_n)$ is satisfiable. Hence, it holds that $\sigma = -$. For $k = n, n > 0$, it holds that $e_n(I_1,\ldots,\bar{\imath}_j,\ldots,I_n) \equiv \bot$, whereas $e_n(I_1,\ldots,i_j,\ldots,I_n)$ is satisfiable. We conclude that $\sigma = +$. □

For the threshold function, the following result is obtained.

Proposition 2 *Let $\mathcal{B} = (G,\Pr)$ be a Bayesian network representing a causal independence model with interaction function f equal to the threshold function t_k, then the sign σ in $S^\sigma(C_j,E)$ is equal to $+$ for $k \geq 1$, and $\sigma = 0$ for $k = 0$.*

Proof: The threshold function can be defined using equation (7) by taking $c_0 = \cdots = c_{k-1} =$

$\bot$ and $c_k = \cdots = c_n = \top$. As a consequence, $t_k(I_1, \ldots, i_j, \ldots, I_n)$ and $t_k(I_1, \ldots, \bar{\imath}_j, \ldots, I_n)$ can both be satisfied, but it is also possible that $t_k(I_1, \ldots, i_j, \ldots, I_n)$ is satisfied because $e_k(I_1, \ldots, i_j, \ldots, I_n)$ is satisfied, which for $k \geq 1$ implies that $t_k(I_1, \ldots, \bar{\imath}_j, \ldots, I_n)$ is not satisfied. Finally, if $t_k(I_1, \ldots, i_j, \ldots, I_n)$ is falsified, so is $t_k(I_1, \ldots, \bar{\imath}_j, \ldots, I_n)$. Summarising, for $k \geq 1$ the qualitative influence $\sigma = +$. For $k = 0$, both $t_0(I_1, \ldots, i_j, \ldots, I_n)$ and $t_0(I_1, \ldots, \bar{\imath}_j, \ldots, I_n)$ are always true, and hence $\sigma = 0$. □

Next suppose that the interaction function f is decomposable. We start by considering Boolean expressions built up from the commutative, associative Boolean operators, as discussed above, which we shall study as special cases of symmetric Boolean functions.

Proposition 3 *Let $\mathcal{B} = (G, \Pr)$ be a Bayesian network representing a causal independence model with decomposable interaction function f that is defined in terms of the commutative, associative binary operators shown in Table 1. Then, the sign σ in $S^\sigma(C_j, E)$ as indicated in Table 3 holds for any cause variable C_j and given effect variable E.*

Proof: Let f be the symmetric Boolean function corresponding to $\Pr(e \mid I_1, \ldots, I_n)$ in difference equation (12). Then, the following results are obtained, using equation (7):

- $\wedge$: $c_n = \top$, and $c_k = \bot$, for each $k \neq n$, hence $e_k(I_1, \ldots, i_j, \ldots, I_n) \wedge c_k$ is only satisfiable for $k = n$, and if the expression is satisfied, it follows from Lemma (1) that $e_n(I_1, \ldots, \bar{\imath}_j, \ldots, I_n) \equiv \bot$. Thus, it follows that $\sigma = +$.

- $\vee$: $c_0 = \bot$, and $c_k = \top$, for each $k > 0$, hence, according to Lemma (1) $\exists k, k > 0$: $e_k(I_1, \ldots, i_j, \ldots, I_n) \equiv e_{k-1}(I_1, \ldots, \bar{\imath}_j, \ldots, I_n) \equiv \top$. However, for $k = 1$, it holds that $e_1(I_1, \ldots, i_j, \ldots, I_n) \wedge c_1$ is satisfiable and $e_0(I_1, \ldots, \neg i_j, \ldots, I_n) \wedge c_0 \equiv \bot$. Therefore, $\sigma = +$.

- $\leftrightarrow$: $c_k = \top$ if $n - k$ is even; otherwise $c_k = \bot$. We obtain that if for some k, and the appropriate truth values for the variables $I_1, \ldots, I_n$: $e_k(I_1, \ldots, i_j, \ldots, I_n) \wedge c_k \equiv \top$ then $e_{k-1}(I_1, \ldots, i_j, \ldots, I_n) \wedge c_{k-1} \equiv \bot$ and vice versa. Hence, $\sigma = ?$.

- $\oslash$: for each k, $c_k = \top$ if *odd*(k); $c_k = \bot$ if *even*(k). For k being odd, $e_k(I_1, \ldots, i_j, \ldots, I_n)$ may be satisfied, but this also holds for $e_k(I_1, \ldots, \bar{\imath}_j, \ldots, I_n)$. From Lemma (1) it then follows that $\sigma = ?$.

- $\top, \bot$: here we have that $f_n(I_1, \ldots, I_n) = \top$ or $f_n(I_1, \ldots, I_n) = \bot$ for any I_k, $k = 1, \ldots, n$. In both cases: $\sigma = 0$.

□

Next, the commutative, non-associative operators are studied. Firstly, consider the NOR operator, and assume it to be right associative. It holds that

$$(I_1 \downarrow (I_2 \downarrow (I_3 \downarrow \cdots \downarrow (I_{n-1} \downarrow I_n) \cdots) \equiv \neg I_1 \wedge (I_2 \vee (\neg I_3 \wedge \cdots \vee (\neg I_{n-1} \wedge \neg I_n) \cdots) \quad (13)$$

if n is even, and

$$(I_1 \downarrow (I_2 \downarrow (I_3 \downarrow \cdots \downarrow (I_{n-1} \downarrow I_n) \cdots) \equiv \neg I_1 \wedge (I_2 \vee (\neg I_3 \wedge \cdots \wedge (I_{n-1} \vee I_n) \cdots) \quad (14)$$

if n is odd. Clearly, it matters whether a variable I_j takes an odd or even argument position, as this determines whether or not it will be negated. Table 4 gives the signs assuming the operators to be right associative; one proof is given below.

Table 3: Qualitative influences: commutative, associative operators.

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

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

	Sign for even		Sign for odd	
Operator	**Last**	**Non-last**	**Last**	**Non-last**
$\downarrow$	$-$	$+$	$+$	$-$
$\mid$	$-$	$+$	$+$	$-$

Proposition 4 *Let $\mathcal{B} = (G, \Pr)$ be a Bayesian network representing a causal independence model with decomposable interaction function f that is equal to the NOR operator $\downarrow$. Then, $S^-(C_j, E)$ is satisfied for $j < n$ and odd, or $j = n$ and even; $S^+(C_j, E)$ holds for $j < n$ and even, or $j = n$ and odd.*

Proof: Consider the case that the subscript j of I_j is odd, with $j < n$. Then, Boolean expressions of the form $(\neg I_1 \wedge (I_2 \vee (\cdots \vee (\neg I_j \wedge \mathcal{I}^{j+1}) \cdots)$ have to be considered. Clearly, $\neg I_j \wedge \mathcal{I}^{j+1} \equiv \bot$ for any combination of truth values of I_j and $\mathcal{I}^{j+1}$, with the exception of $\bar{\imath}_j$ if $\mathcal{I}^{j+1} = \top$. Hence, $S^-(C_j, E)$ holds.

Next, suppose that the subscript j of I_j is even, with $j < n$. Then, Boolean expressions of the form $(\neg I_1 \wedge (I_2 \vee (\cdots \wedge (I_j \vee \mathcal{I}^{j+1}) \cdots)$ need consideration. If $\mathcal{I}^{j+1} = \top$, then the value of I_j does not matter, and hence $d_e(\mathcal{I}_n \backslash I_j) = 0$. However, if $\mathcal{I}^{j+1} = \bot$, we obtain $(I_j \vee \mathcal{I}^{j+1}) \equiv \top$ only for i_j. Hence, $d_e(\mathcal{I}_n \backslash I_j) \geq 0$ in that case, i.e. $S^+(C_j, E)$ holds.

Finally, if $j = n$ then the cases considered above still apply, except that odd and even need to be reversed. □

The left-associative case is simply obtained by determining $n - j + 1$ for I_j in the left-associative Boolean expression, and looking up the sign in Table 4, where the result for the first argument becomes the result for the last argument. This is a consequence of commutativity.

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 in the first or any other argument position, and in the last or any other argument position. This is illustrated by the proof below for the p_1 operator.

Proposition 5 *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 p_1. Then, $S^+(C_j, E)$ is satisfied for $j = 1$, otherwise, if $j \neq 1$, $S^0(C_j, E)$ holds.*

Proof: Let $\mathcal{I}_n$ and $\mathcal{I}'_n$ be Boolean expressions corresponding to the constituents of difference (12). In general, it holds that $I_1 \; p_1 \; \mathcal{I}_{n-1} = I_1$. Now, if $j = 1$, then $\mathcal{I}_n \equiv i_1$ and $\mathcal{I}'_n \equiv \bar{\imath}_1$. This

Table 5: Qualitative influences: 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

implies that $S^+(C_j, E)$ holds, as $\Pr(e \mid I_1, \ldots, i_j, \ldots, I_n) = 1$, and $\Pr(e \mid I_1, \ldots, \bar{\imath}_j, \ldots, I_n) = 0$, which follows from the logical analysis above.

Next, assume that $j \neq 1$, then $\mathcal{I}_n = \mathcal{I}'_n = I_1$, and thus the difference (12) is always equal to 0, i.e. $S^0(C_j, E)$ holds. □

Finally, the results in Table 6 for the increasing order operator are proven.

Proposition 6 *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 increasing order operator $<$. Let $V(G)$ contain $n \geq 1$ interaction variables $I_1, \ldots, I_n$ and an effect variable E. If assuming that the operator $<$ is left associative, it holds that $S^?(C_j, E)$ if $j < n$, otherwise, if $j = n$, $S^+(C_j, E)$ holds. Assume that the operator $<$ is right associative, then $S^-(C_j, E)$ holds for $j < n$; otherwise, for $j = n$, $S^+(C_j, E)$ is satisfied.*

Proof: Consider the case where $<$ is assumed to be left associative. Recall the definition of $\mathcal{I}_j = (\cdots(I_1 < I_2) < \cdots) < I_{j-1}) < I_j)$. First, we assume $j < n$. Clearly, if $\bar{\imath}_n$ holds, then $\mathcal{I}_n \equiv \bot$. So, we only consider the case for i_n. Now, assume that $\mathcal{I}_{j-1} \equiv \top$, then for both i_j and $\bar{\imath}_j$ it holds that $\mathcal{I}_j \equiv \bot$. So, there is no difference in the resulting truth values for i_j and $\bar{\imath}_j$, and the difference (12) is therefore equal to 0. Next, consider the case that $\mathcal{I}_{j-1} \equiv \bot$. Then, we obtain: $(\mathcal{I}_{j-1} < i_j) \equiv \top$ and $(\mathcal{I}_{j-1} < \bar{\imath}_j) \equiv \bot$. For i_{j+1}, this would yield $\mathcal{I}_{j+1} \equiv \bot$ and $\mathcal{I}_{j+1} \equiv \top$, respectively, inverting the truth values of $\mathcal{I}_j$. The resulting truth values can also be both $\bot$ when taking $\bar{\imath}_{j+2}$. So, this means that the difference can be 0, $-$, or $+$, thus $S^?(C_j, E)$ holds for $j < n$. Now, assume that $j = n$, then only i_n can satisfy the expression. Hence, $S^+(C_j, E)$ holds.

Next, consider the case that $<$ is right associative. Recall definition (5) of $\mathcal{I}^j = (I_j < (I_{j+1} < (\cdots < (I_{n-1} < I_n) \cdots)$. Assume that $j < n$. Now, $(i_j < \mathcal{I}_{j+1}) \equiv \bot$, whereas $(\bar{\imath}_j < \mathcal{I}_{j+1})$ is satisfiable. Hence, $S^-(C_k, E)$ holds. Next, consider the case that $j = n$. Then, only i_n is able to satisfy $\mathcal{I}^1$, i.e. $S^+(C_j, E)$ holds. □

The proofs for the other non-commutative, non-associative operators are similar; the results are given in Table 6. Tables 3, 5 and 6 clearly indicate that it is possible to model all possible qualitative influences among causes and effects, even if it is assumed that the interaction function is decomposable.

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 the implication (Table 6) would model this situation

Table 6: Qualitative influences: non-commutative, non-associative operators; RA: right associative; LA: left associative.

	Sign			
	RA		**LA**	
Operator	First	Non-first	First	Non-first
$\leftarrow$	$+$	?	$+$	$-$
$>$	$+$	?	$+$	$-$
Operator	Last	Non-last	Last	Non-last
$\rightarrow$	$+$	$-$	$+$	?
$<$	$+$	$-$	$+$	?

appropriately. For the qualitative influence of penicillin or chlortetracyclin on bacterial growth we obtain an ambiguity due to the exclusive OR. This clearly expresses that the effect of penicillin on bacterial growth as depicted in Figure 2 is dependent on the presence or absence of chlortetracyclin and vice versa. For the qualitative influence of insulin hypersecretion on hypoglycaemia, as shown in Figure 3, we obtain a positive sign due to the decreasing order operator, whereas for glucagon hypersecretion we obtain a negative sign. Here we take the left-associative version of the decreasing order operator $>$, as this is the most specific one, and, therefore, expresses the situation when two variables are involved. This formal representation is clearly consistent with what has been described about the glucose metabolism above.

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 (9) of an additive synergy, and considering interactions between the causes C_{j-1} and C_j, we obtain:

$$
\begin{aligned}
&\delta_{j-1,j}(C_1,\ldots,C_{j-2},C_{j+1},\ldots,C_n) = \\
&\quad \Pr(e \mid C_1,\ldots,c_{j-1},c_j,C_{j+1},\ldots,C_n) + \Pr(e \mid C_1,\ldots,\bar{c}_{j-1},\bar{c}_j,C_{j+1},\ldots,C_n) - \\
&\quad \Pr(e \mid C_1,\ldots,c_{j-1},\bar{c}_j,C_{j+1},\ldots,C_n) - \Pr(e \mid C_1,\ldots,\bar{c}_{j-1},c_j,C_{j+1},\ldots,C_n) = \\
&\quad \sum_{f(I_1,\ldots,I_n)=e} d(I_{j-1},I_j) \prod_{k=1}^{j-2} \Pr(I_k \mid C_k) \prod_{k=j+1}^{n} \Pr(I_k \mid C_k)
\end{aligned}
$$

where

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

As the function f renders the variables $I_1,\ldots,I_n \backslash I_{j-1},I_j$ dependent of the variables I_{j-1} and I_j it is not possible to distribute summation over the expression.

Let $\Pr(i_{j-1} \mid c_{j-1}) = p$ and $\Pr(i_j \mid c_j) = q$, then the difference $d(I_{j-1},I_j)$ corresponds for different values of I_{j-1} and I_j, using the assumptions introduced in Section 2.2, to the results given in Table 7. As a consequence $\delta_{j-1,j}(C_1,\ldots,C_{j-2},C_{j+1},\ldots,C_n)$ can be simplified to obtain the following result:

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

Table 7: Difference $d(I_{j-1}, I_j)$ for various values of the variables I_{j-1} and I_j.

I_{j-1}	I_j	$d(I_{j-1}, I_j)$
i_{j-1}	i_j	pq
$\bar{\imath}_{j-1}$	i_j	$-pq$
i_{j-1}	$\bar{\imath}_j$	$-pq$
$\bar{\imath}_{j-1}$	$\bar{\imath}_j$	pq

Table 8: Signs of additive synergies for the commutative, associative operators.

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

$$\sum_{I_1,\ldots,I_n \setminus I_{j-1}, I_j} \sum_{\substack{I_{j-1}, I_j \\ f(I_1,\ldots,I_n) = e}} \sigma(I_{j-1} \oslash I_j)\, pq \prod_{k=1}^{j-2} \Pr(I_k \mid C_k) \prod_{k=j+1}^{n} \Pr(I_k \mid C_k) \tag{15}$$

where $\oslash$ represents the exclusive or, and

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

The multipliers $\prod_{k=1}^{j-2} \Pr(I_k \mid C_k) \prod_{k=j+1}^{n} \Pr(I_k \mid C_k)$, with $\prod_{k=1}^{j-2} \Pr(I_k \mid C_k) \prod_{k=j+1}^{n} \Pr(I_k \mid C_k) \geq 0$, will generally differ for various $\delta_{j-1,j}(C_1, \ldots, C_{j-2}, C_{j+1}, \ldots, C_n)$. 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, those that are non-commutative but associative, and those that are neither commutative nor associative. The results for the associative and commutative operators are given in Table 8. In this case, we can simply assume that $j = 2$ without loss of generality, which simplifies the proofs. Again, the proof for only some of the Boolean operators is given here.

Proposition 7 *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 $Y^-(\{C_{j-1}, C_j\}, E)$ for any two cause variables C_{j-1}, C_j and the given effect variable E.*

Table 9: Signs of additive synergies for the commutative, non-associative operators; right-associative case.

	Sign for even		**Sign for odd**	
Operator	**Last**	**Non-last**	**Last**	**Non-last**
$\downarrow$	$+$	$-$	$-$	$+$
$\mid$	$-$	$+$	$+$	$-$

Table 10: Signs of additive synergies for the non-commutative, associative operators.

Operator	**Sign**
p_1	0
p_2	0
n_1	0
n_2	0

Table 11: Signs of additive synergies for the non-commutative, non-associative operators; RA: right-associative; LA: left-associative.

	Sign			
	RA		**LA**	
Operator	First	Non-first	First	Non-first
$\leftarrow$	$+$	$-$	$+$	$-$
$>$	$-$	$+$	$-$	$+$
Operator	Last	Non-last	Last	Non-last
$\rightarrow$	$+$	$-$	$+$	$-$
$<$	$-$	$+$	$-$	$+$

Proof: If the interaction function is represented by the Boolean expression $I_1 \vee I_2 \vee \cdots \vee I_n$, then it is easily verified that

$$\sum_{\substack{I_1, I_2 \\ I_1 \vee \cdots \vee I_n}} \sigma(I_1 \oslash I_2)\, pq$$

for given values of $I_3, \ldots, I_n$ is either equal to $-pq\,(= pq - 2pq)$ or to $0\,(= 2pq - 2pq)$. Since, $\prod_{k=3}^{n} \Pr(I_k \mid C_k) \geq 0$, a logical OR interaction function clearly results in a negative additive synergy. □

We next present the proof for the case that the Boolean operator is equal to the bi-implication.

Proposition 8 *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 bi-implication. Then, it holds that $Y^?(\{C_{j-1}, C_j\}, E)$ for any two cause variables C_{j-1}, C_j and the given effect variable E.*

Proof: Let the interaction function be represented by the Boolean expression $I_1 \leftrightarrow I_2 \leftrightarrow \cdots \leftrightarrow I_n \equiv I_1 \leftrightarrow I_2 \leftrightarrow \mathcal{I}_{n-2}$. Two general cases for which the Boolean expression is true are distinguished. Firstly, assume that $(I_1 \leftrightarrow I_2) \equiv \top$ and $\mathcal{I}_{n-2} \equiv \top$. It is easily verified that then the inner sum of equation (15) is equal to $2pq$. Secondly, assume that $(I_1 \leftrightarrow I_2) \equiv \bot$, and $\mathcal{I}_{n-2} \equiv \bot$ as well. Then, the inner sum is equal to $-2pq$. As the used multipliers will be different, the result is ambiguous. □

Next, the two commutative, non-associative operators are considered. Here, we only supply a proof for the NAND $\mid$ operator; the results are summarised in Table 9. Note that it is now no longer permitted to only look at the variables I_1 and I_2.

Proposition 9 *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 NAND $|$. Then, it holds that $Y^+(\{C_{j-1}, C_j\}, E)$ for $j < n$ and even or $j = n$ and odd, and $Y^-(\{C_{j-1}, C_j\}, E)$ holds for $j < n$ and odd or $j = n$ and even.*

Proof: Note that we have that:

$$(I_1 \mid (I_2 \mid (I_3 \mid \cdots \mid (I_{n-1} \mid I_n) \cdots) \equiv \neg I_1 \vee (I_2 \wedge (\neg I_3 \vee \cdots \wedge (\neg I_{n-1} \vee \neg I_n) \cdots)$$

if n is even, and

$$(I_1 \mid (I_2 \mid (I_3 \mid \cdots \mid (I_{n-1} \mid I_n) \cdots) \equiv \neg I_1 \vee (I_2 \wedge (\neg I_3 \vee \cdots \vee (I_{n-1} \wedge I_n) \cdots)$$

if n is odd. First, we consider $j < n$ and even. Then, $\neg I_{j-1} \vee (I_j \wedge \mathcal{I}^{j+1}) \equiv \bot$ for the combination i_{j-1} and $\bar{\imath}_j$; for the other combinations of truth values this expression is satisfiable. If $j = n$ and odd, the Boolean expression $I_{n-1} \wedge I_n$ needs to be considered, and this is only true for the combination i_{n-1} and i_n. In both cases, the result is $Y^+(\{C_{j-1}, C_j\}, E)$.

Secondly, consider the case that $j < n$ and odd. Then, the Boolean expression $I_{j-1} \wedge (\neg I_j \vee \mathcal{I}^{j+1})$ must be considered. This is true for i_{j-1} and i_j, satisfiable for i_{j-1} and $\bar{\imath}_j$, and otherwise false. If $j = n$ and even, we need to consider $\neg I_{n-1} \vee \neg I_n$. It holds that $\neg i_{n-1} \vee \neg i_n \equiv \bot$; otherwise, the Boolean expression is true. It is concluded that $Y^-(\{C_{j-1}, C_j\}, E)$ holds in both cases. □

We next move on to consider the non-commutative, associative operators.

Proposition 10 *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: It holds that the Boolean expression representation of the interaction function f is equal 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, the result of summing these results will be 0. If $j = 2$, then I_{j-1} should always be equal to i_{j-1}, and hence the sum is only taken over $d(i_1, i_2)$ and $d(i_1, \bar{\imath}_2)$, which, however, also yields 0. □

Similar results are obtained for the other non-commutative, associative operators, and their proofs are similar. Note that, in contract to the results for the qualitative S relation, there are no differences in results when considering either the first, last of any other pair of causes. The reason for this is that the operators select at most one argument, and hence, either all 4 possible Boolean combinations of Boolean values 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 products pq and $-pq$, which cancel out each other, resulting in a total of 0. The results are summarised in Table 10.

Finally, the non-commutative and non-associative operators have to be considered. This analysis is more difficult, as a distinction must be made between assuming the operators to be right associative or left associative. We present the proof for the increasing order operator $<$. The other proofs are similar.

Proposition 11 *Let $\mathcal{B} = (G, \Pr)$ be a Bayesian network representing a causal independence model with decomposable interaction function f that is equal to the increasing order operator $<$. Let $V(G)$ contain $n \geq 2$ interaction variables $I_1, \ldots, I_n$ and an effect variable E. Assume that the operator is left associative, then it holds that $Y^+(\{C_{j-1}, C_j\}, E)$ for $j < n$; for $j = n$ it holds that $Y^-(\{C_{j-1}, C_j\}, E)$. Next, assume that the operator is right associative, then it holds that $Y^+(\{C_{j-1}, C_j\}, E)$ for $j < n$; for $j = n$ it holds that $Y^-(\{C_{j-1}, C_j\}, E)$.*

Proof: Firstly, consider the case that the operator $<$ is assumed to be left associative. Recall definition (4) of $\mathcal{I}_j$. We take I_{j-1} and I_j as the interaction variables, with $j < n$. Now if $\bar{\imath}_j$ holds, then $\mathcal{I}_j \equiv \bot$, and hence $\mathcal{I}_n$ is satisfiable. We conclude that we need to take into account $d(i_{j-1}, \bar{\imath}_j) + d(\bar{\imath}_{j-1}, \bar{\imath}_j) = 0$. Next, consider the case that i_j holds. Then if $\bar{\imath}_{j-1}$ holds, we have that $\mathcal{I}_j \equiv \top$. As $\mathcal{I}_{n-1}$ must be false in order to make $\mathcal{I}_n \equiv \top$, there must be at least one variable I_k, $k > j$, which falsifies $\mathcal{I}_{n-1}$. Finally, for i_{j-1}, the expression $\mathcal{I}_j$ is again satisfiable. Now, as $d(i_{j-1}, i_j) = pq$, we know that $\delta_{j-1,j}(C_1, \ldots, C_{j-2}, C_{j+1}, \ldots, C_n) \geq 0$. Next, consider the case that $j = n$. Then I_n must always be true in order $\mathcal{I}_n$ to be true. Now, if $\bar{\imath}_{n-1}$ holds, we know that $\mathcal{I}_n \equiv \top$, whereas if i_{n-1} holds, then $\mathcal{I}_n$ is only satisfiable. Hence, it was shown that $\delta_{n-1,n}(C_1, \ldots, C_{n-2}) \leq 0$.

Secondly, consider the case that the operator $<$ is right associative. Recall definition (5) of $\mathcal{I}^j$. We first consider the case that $j < n$. Clearly, in order $\mathcal{I}^j$ to be satisfiable, i_n must holds. However, $\mathcal{I}^j \equiv \bot$ if it holds that there exists an I_k that is true for $j \leq k < n$. Hence, only $d(\bar{\imath}_{j-1}, \bar{\imath}_j) = pq$ needs to be taken into account, resulting in $\delta_{j-1,j}(C_1, \ldots, C_{j-2}, C_{j+1}, \ldots, C_n) \geq 0$. Finally, consider $j = n$. Here, we simply have that only $d(\bar{\imath}_{n-1}, i_n) = -pq$ needs to be taken into account, so we conclude that $\delta_{n-1,n}(C_1, \ldots, C_{n-2}) \leq 0$. □

The proofs for the other operators that are non-commutative and non-associative are along similar lines. The results are summarised in Table 11.

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 an infection less likely, whereas bacterium C is so pathogenic that it overrides the preventive effects of bacteria A and B. For the interaction between penicillin and chlortetracyclin, modelled as an exclusive OR and shown in Figure 2, we have a negative additive synergy. This is as might be expected, as when these potential causes of decreased bacterial growth are both present or absent, there will be no antimicrobial effect, in contrast to when only one of these is present. For the interaction between insulin and glucagon hypersecretion, modelled as the decreasing order operator $>$ and shown in Figure 3, we have a negative additive synergy, as the two hormones have opposite effects on the glucose level of blood.

3.3 Analysis of product synergies

We basically use the same approach as employed in the previous section on additive synergies for product synergies in this section. For the analysis of product synergies, the equation of interest is:

$$\delta^E_{j-1,j}(C_1, \ldots, C_{j-2}, C_{j+1}, \ldots, C_n) =$$
$$\Pr(E \mid C_1, \ldots, c_{j-1}, c_j, C_{j+1}, \ldots, C_n) \cdot \Pr(E \mid C_1, \ldots, \bar{c}_{j-1}, \bar{c}_j, C_{j+1}, \ldots, C_n) -$$

Table 12: Signs of product synergies for the commutative and associative operators.

Operator	Sign for e	Sign for $\bar{e}$
$\wedge$	0	$-$
$\vee$	$-$	0
$\leftrightarrow$	?	?
$\oslash$	?	?
$\top$	0	0
$\bot$	0	0

Table 13: Signs of product synergies for the commutative, non-associative operators assuming that $E = \top$; right-associative case.

	Sign for e			
	Even		**Odd**	
Operator	**Last**	**Non-last**	**Last**	**Non-last**
$\downarrow$	0	0	$-$	$+$
$\mid$	$-$	$+$	0	0

$$\begin{aligned} &\Pr(E \mid C_1, \ldots, c_{j-1}, \bar{c}_j, C_{j+1}, \ldots, C_n) \cdot \Pr(E \mid C_1, \ldots, \bar{c}_{j-1}, c_j, C_{j+1}, \ldots, C_n) = \\ &\sum_{I_1, \ldots, I_n \backslash I_{j-1}, I_j} \{\tau(c_{j-1}, c_j; \mathcal{I}_n \backslash I_{j-1}, I_j) \cdot \tau(\bar{c}_{j-1}, \bar{c}_j; \mathcal{I}_n \backslash I_{j-1}, I_j) - \\ &\quad \tau(c_{j-1}, \bar{c}_j; \mathcal{I}_n \backslash I_{j-1}, I_j) \cdot \tau(\bar{c}_{j-1}, c_j; \mathcal{I}_n \backslash I_{j-1}, I_j)\} \prod_{k=1}^{j-2} \Pr(I_k \mid C_k) \prod_{k=j+1}^{n} \Pr(I_k \mid C_k) \end{aligned} \tag{16}$$

where

$$\tau(C_{j-1}, C_j; \mathcal{I}_n \backslash I_{j-1}, I_j) = \sum_{\substack{I_{j-1},\, I_j \\ f(I_1, \ldots, I_n) = E}} \Pr(I_{j-1} \mid C_{j-1}) \Pr(I_j \mid C_j) \tag{17}$$

The arithmetic expression between the braces in (16), consisting of additions and subtractions of products of instances of $\tau(C_{j-1}, C_j; \mathcal{I}_n \backslash I_{j-1}, I_j)$, 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 $q = \Pr(i_j \mid c_j)$. In the following, the equation above will be studied for all possible Boolean operator definitions of the interaction function f, which is again assumed to be decomposable. As before for the operators which are commutative and associative, instead of focusing the analysis on two arbitrary cause variables C_{j-1} and C_j, for simplicity's

Table 14: Signs of product synergies for the commutative, non-associative operators assuming that $E = \bot$; right-associative case.

	Sign for $\bar{e}$			
	Even		**Odd**	
Operator	**Last**	**Non-last**	**Last**	**Non-last**
$\downarrow$	$-$	$+$	0	0
$\mid$	0	0	$-$	$+$

sake, the interaction of the two equally arbitrary variables C_1 and C_2 is examined, i.e. we take $j = 2$. Again, for only some of the Boolean operators a proof is provided.

Proposition 12 *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 is true. 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 $X^-(\{C_1, C_2\}, e)$ holds.

Let us now consider the case that E is false. 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 $\delta_{1,2}^{\bar{e}}(C_3, \ldots, C_n) = 0$, and thus $X^0(\{C_1, C_2\}, \bar{e})$ holds. □

For the bi-implication, the following result is obtained.

Proposition 13 *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 bi-implication. 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 $X^?(\{C_{j-1}, C_j\}, \bar{e})$ when the effect is assumed to be false.*

Proof: Let the interaction function be represented by the Boolean expression $\mathcal{I}_n = I_1 \leftrightarrow I_2 \leftrightarrow \mathcal{I}_{n-2}$. Firstly, take E to be true. Let us assume that $I_{n-2} \equiv \top$, then I_1 and I_2 must be both true or false. The result is then $\beta = pq$. Next, we assume that $I_{n-2} \equiv \bot$; then, I_1 must be true and I_2 must be false, or I_1 must be false and I_2 must be true. We get: $\beta = [p(1-q) + (1-p)q] \cdot 0 - pq = -pq$. It is concluded that $X^?(\{C_1, C_2\}, e)$ holds.

Assume now that E is false. For $\mathcal{I}_{n-2} \equiv \top$ we get that either I_1 or I_2 is true, but not both. From this, as above we conclude that $\beta = -pq$. Subsequently assuming that $\mathcal{I}_{n-2} \equiv \bot$ yields the same as above for e and $I_{n-2} \equiv \top$; hence, $\beta = pq$. Again, $X^?(\{C_1, C_2\}, e)$ is satisfied. □

The results for the commutative, non-associative operators $\downarrow$ and $|$ are shown in Tables 13 and 14. Again only the right-associative case is covered in the tables, but using commutativity, it is easy to obtain the signs for the left-associative case. Below, the proof for the NOR operator $\downarrow$ is given.

Proposition 14 *Let $\mathcal{B} = (G, \Pr)$ be a Bayesian network representing a causal independence model with decomposable interaction function f that is equal to the NOR operator $\downarrow$. Then, it holds that $X^0(\{C_{j-1}, C_j\}, e)$, resp. $X^0(\{C_{j-1}, C_j\}, \bar{e})$, for any two cause variables C_{j-1}, C_j given that the effect is true, resp. false, with j even, resp. odd. Furthermore, $X^+(\{C_{j-1}, C_j\}, e)$, resp. $X^+(\{C_{j-1}, C_j\}, \bar{e})$, holds for $j < n$ and odd, resp. even, whereas $X^-(\{C_{j-1}, C_j\}, e)$, resp. $X^-(\{C_{j-1}, C_j\}, \bar{e})$, holds for $j = n$ and odd, resp. even.*

Proof: The equivalences (13) and (14) are again used as a basis for the proof.

Case I: e (E is true) holds. Let $j < n$ and odd, then $I_{j-1} \vee (\neg I_j \wedge \mathcal{I}^{j+1})$ needs to be considered. Firstly, suppose that $\mathcal{I}^{j+1} = \top$, then, $\bar{\imath}_{j-1}, i_j$ are not taken into account, yielding $\beta = [1 - (1-p)q] \cdot 1 - 1 \cdot [1 - q] = pq \geq 0$. Secondly, suppose that $\mathcal{I}^{j+1} = \bot$, then only i_{j-1}, i_j and $i_{j-1}, \bar{\imath}_j$ must be considered, yielding: $\beta = [pq + p(1-q)] \cdot 0 - 0 = 0$. Hence, it can be concluded that $X^+(\{C_{j-1}, C_j\}, e)$ holds.

Table 15: Signs of product synergies for the non-commutative, associative operators.

Operator	**Sign for both e and $\bar{e}$**
p_1	0
p_2	0
n_1	0
n_2	0

Next, let $j = n$ and odd, then $I_{n-1} \vee I_n$ needs to be considered. Here, only $\bar{\imath}_{j-1}, \bar{\imath}_j$ is discarded, yielding $\beta = [1-(1-p)(1-q)] \cdot 0 - [p][q] = -pq \leq 0$. Hence, $X^-(\{C_{n-1}, C_n\}, e)$ holds.

For $j < n$ and j even, the Boolean expression $\neg I_{j-1} \wedge (I_j \vee \mathcal{I}^{j+1})$ needs to be analysed. Firstly, suppose that $\mathcal{I}^{j+1} = \top$, then only the combinations $\bar{\imath}_{j-1}, i_j$ and $\bar{\imath}_{j-1}, \bar{\imath}_j$ are able to satisfy this expression. As result, we have that $\beta = [(1-p)q + (1-p)(1-q)] \cdot 1 - [(1-p)] \cdot 1 = 0$. Secondly, suppose that $\mathcal{I}^{j+1} = \bot$, then only $\bar{\imath}_{j-1}, i)j$ needs to be taken into account. This implies that $\beta = 0 - 0 = 0$ holds. Summarised, $X^0(\{C_{j-1}, C_j\}, e)$ holds.

Next, let $j = n$ and even, then the Boolean expression that needs to be considered is $\neg I_{n-1} \wedge \neg I_n$. Hence, only the pair $\bar{\imath}_{n-1}, \bar{\imath}_n$ is able to satisfy this expression. It therefore holds that $\beta = [(1-p)(1-q)] \cdot 1 - [1-p][1-q] = 0$, i.e. $X^0(\{C_{n-1}, C_n\}, e)$ holds.

Case II: $\bar{e}$ (E is false) holds. The proofs are very similar to the ones given above, as the same cases have to be considered, which we shall not fully repeat. If $j < n$ is odd, then for $\mathcal{I}^{j+1} = \top$, only $\bar{\imath}_{j-1}, \bar{\imath}_j$ needs to be taken into account. This yields $\beta = [(1-p)(1-q)] \cdot 1 - [1-p][1-q] = 0$. Secondly, suppose that $\mathcal{I}^{j+1} = \bot$, then $\bar{\imath}_{j-1}, i_j$ and $\bar{\imath}_{j-1}, \bar{\imath}_j$ must be considered, yielding: $\beta = [(1-p)q + (1-p)(1-q)] \cdot 1 - [1-p] \cdot 1 = 0$. Hence, it can be concluded that $X^0(\{C_{j-1}, C_j\}, e)$ holds.

Next, let $j = n$ and odd, then only $\bar{\imath}_{j-1}, \bar{\imath}_j$ are considered, yielding $\beta = [(1-p)(1-q)] \cdot 1 - [1-p][1-q] = 0$. Hence, $X^0(\{C_{n-1}, C_n\}, e)$ holds.

For $j < n$ and j even, let $\mathcal{I}^{j+1} = \top$, then only the combinations i_{j-1}, i_j and $i_{j-1}, \bar{\imath}_j$ falsify the Boolean expression above. As result, we have that $\beta = 0$. Secondly, suppose that $\mathcal{I}^{j+1} = \bot$, then all combinations of values for I_{j-1} and I_j, with the exception of $\bar{\imath}_{j-1}, i_j$, need to be taken into account. This implies that $\beta = [1-(1-p)q] \cdot 1 - 1 \cdot [1-q] = pq \geq 0$ holds. Summarised, $X^+(\{C_{j-1}, C_j\}, e)$ holds.

Finally, let $j = n$ and even, then all combinations of values for I_{j-1} and I_j with the exception of $\bar{\imath}_{n-1}, \bar{\imath}_n$ need to be taken into account. It therefore holds that $\beta = [1-(1-p)(1-q)] \cdot 0 - [p][q] = -pq \leq 0$, i.e. $X^-(\{C_{n-1}, C_n\}, e)$ holds. □

Next, we consider the operators which are non-commutative, but associative. The results are summarised in Table 15, and correspond to those for the additive synergies, discussed in the previous section. The explanation why only zero product synergies are obtained is analogous to that for the additive synergies as well.

Finally, the operators which are neither commutative nor associative are considered. The proof for the implication is given below. Again we make a distinction between a right-associative and a left-associative reading of Boolean expressions.

Proposition 15 *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 implication $\rightarrow$. Let $V(G)$ contain $n \geq 2$ interaction variables $I_1, \ldots, I_n$ and an effect variable E. For any two*

Table 16: Signs of product synergies for the non-commutative, non-associative operators assuming that $E = \top$; RA: right-associative; LA: left-associative.

	Sign for e			
	RA		**LA**	
Operator	First	Non-first	First	Non-first
$\leftarrow$	$+$	$+$	$+$	$-$
$>$	0	$+$	0	0
Operator	Last	Non-last	Last	Non-last
$\rightarrow$	$+$	$-$	$+$	$+$
$<$	0	0	0	$+$

Table 17: Signs of product synergies for the non-commutative, non-associative operators assuming that $E = \bot$; RA: right-associative; LA: left-associative.

	Sign for $\bar{e}$			
	RA		**LA**	
Operator	First	Non-first	First	Non-first
$\leftarrow$	0	$+$	0	0
$>$	0	0	$+$	$-$
Operator	Last	Non-last	Last	Non-last
$\rightarrow$	0	0	0	$+$
$<$	$+$	$-$	0	0

cause variables C_{j-1}, C_j given that the effect is true and assuming $\rightarrow$ to be left associative, it holds that $X^+(\{C_{j-1}, C_j\}, e)$ for $j < 1 \leq n$. Assuming that $\rightarrow$ is right associative, it holds that $X^-(\{C_{j-1}, C_j\}, e)$ for $j < n$, and $X^+(\{C_{j-1}, C_j\}, e)$ for $j = n$. If the effect is taken to be false, the following holds. Assuming that $\rightarrow$ is left associative, we obtain $X^+(\{C_{j-1}, C_j\}, \bar{e})$ for $j < n$, and $X^0(\{C_{j-1}, C_j\}, \bar{e})$ for $j = n$. Finally, assuming that $\rightarrow$ is right associative, the product synergy is equal to $X^0(\{C_{j-1}, C_j\}, \bar{e})$ for $1 < j \leq n$.

Proof: Case I: e (E is true) holds. Assume that the operator $\rightarrow$ is left associative. Recall definition (4) of $\mathcal{I}_j$ and take $j < n$. If i_n is assumed to hold, then we have to take into account all four Boolean combinations of I_{j-1} and I_j, resulting in $\beta = 0$. Now, assume that $\bar{\imath}_n$ holds. Suppose that $\mathcal{I}_{j-2} \equiv \top$, then we obtain $\mathcal{I}_j \equiv \bot$ for i_{j-1}, $\bar{\imath}_j$, and $\mathcal{I}_j \equiv \top$ for the other three combinations. For $\mathcal{I}_{j-2} \equiv \bot$, it holds that $\mathcal{I}_j \equiv \bot$ for i_{j-1}, $\bar{\imath}_j$ and $\bar{\imath}_{j-1}$, $\bar{\imath}_j$. As the truth value of $\mathcal{I}_k$ may change from $\top$ to $\bot$ for $k > j$, we need to take into account i_{j-1}, $\bar{\imath}_j$ with $\beta = 0$, or i_{j-1}, $\bar{\imath}_j$ and $\bar{\imath}_{j-1}$, $\bar{\imath}_j$ with $\beta = 0$, or i_{j-1}, i_j and $\bar{\imath}_{j-1}$, i_j with $\beta = 0$, or all four combinations with the exception of i_{j-1}, $\bar{\imath}_j$ yielding $\beta = [1 - p(1-q)] \cdot 1 - (1-p) \cdot 1 = pq \geq 0$. Hence, $X^+(\{C_{j-1}, C_j\}, e)$ is satisfied. Finally, assume that $j = n$. For i_n it holds that $\mathcal{I}_n \equiv \top$, whereas for $\bar{\imath}_n$, we obtain $\mathcal{I}_n \equiv \bot$ for i_{n-1}, whereas for $\bar{\imath}_{n-1}$ the truth value depends on the truth value of $\mathcal{I}_{n-2}$. The result is equal to $\beta = [pq + (1-p)q + (1-p)(1-q)] \cdot q - (1-p) \cdot 1 = pq \geq 0$; hence again $X^+(\{C_{j-1}, C_j\}, e)$ is shown to hold.

Next, the operator $\rightarrow$ is assumed to be right associative. Recall definition (5) of $\mathcal{I}^j$. Assume that $j < n - 1$. Again, there are two cases to consider: $\mathcal{I}^{j+1} \equiv \top$ and $\mathcal{I}^{j+1} \equiv \bot$. For $\mathcal{I}^{j+1} \equiv \top$, we have that $\beta = 0$, as we sum over all possible values of both I_{j-1} and I_j, yielding $\beta = 0$. For $\mathcal{I}^{j+1} \equiv \bot$, we sum over all values of I_{j-1} and I_j, with the exception of the combination i_{j-1} and i_j (as $i_{j-1} \to (i_j \to \bot) \equiv \bot$). We obtain $\beta = (1-pq)\cdot 1 - 1\cdot 1 = -pq \leq 0$, and thus it can be concluded that $X^-(\{C_{j-1}, C_j\}, e)$ holds. Now, assume that $j = n$. If i_n is assumed to hold, then $\mathcal{I}^1 \equiv \top$. Similarly, it holds that $\mathcal{I}^1 \equiv \top$ if $I_{n-1} \equiv I_n \equiv \top$. The expression $\mathcal{I}^1$ can only be false for the combination i_{n-1} and $\bar{\imath}_n$. It is concluded that $\beta = [1 - p(1-q)] \cdot 1 - (1-p) \cdot 1 \geq 0$, i.e. $X^+(\{C_{j-1}, C_j\}, e)$ was shown to hold.

Case II: $\bar{e}$ (E is false) holds. The notational conventions as introduced above will again be adopted. It is clearly possible to reuse most of the results obtained for case I above. First, assume that $\to$ is left associative. Take $j < n$. For i_n, the summations over I_{j-1} and I_j are empty. Now, assume that $\bar{\imath}_n$ holds. Then we sum over the same values of the interaction variables as for $E = \top$. So, the results are exactly the same. Now take $j = n$; only $\bar{\imath}_n$ is of interest then. For i_{n-1}, it holds that $\mathcal{I}_n \equiv \bot$, whereas for $\bar{\imath}_{n-1}$ it holds that $\mathcal{I}_n$ is satisfiable. Hence, we obtain $\beta = 0$. Next, assume that $\to$ is right associative. Take $j < n$. Then, if $\mathcal{I}^{j+1} \equiv \top$, it holds that $\mathcal{I}^1 \equiv \top$, and the sums over I_{j-1} and I_j are all empty. Assume now that $\mathcal{I}^{j+1} \equiv \bot$, then only the combination $\bar{\imath}_{j-1}$ and $\bar{\imath}_j$ yields $\mathcal{I}^{j-1} \equiv \bot$. The corresponding β is equal to $\beta = 0$. Finally, assume that $j = n$. The only possibility of falsifying $\mathcal{I}^1$ is by the combination i_{j-1} and $\bar{\imath}_j$. This yields $\beta = 0$. $\Box$

The results of the analysis are given in Table 16 assuming that the effect is positive and in Table 17 assuming a negative effect.

We return to the example in Figure 1, where, as in the previous section, we assume that logical implication provides a suitable formalisation of the interactions between the bacteria involved in infection. 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 if A or B are present, then C is also present. 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, as otherwise A or B would not have caused the infection. This probabilistic behaviour is appropriately modelled by the right-associative version of implication. The causal mechanisms involved in the interaction between bacterial growth and antimicrobial agents, as shown in Figure 2, was modelled by means of the exclusive OR. According to Table 12 the intercausal influences modelled by the exclusive OR with an arbitrary number of causes are ambiguous. However, if we assume that there are no other factors involved but these two antimicrobial agents, it can be shown, along the lines of Proposition 13, that assuming that there is decreased bacterial growth due to penicillin (chlortetracyclin) it holds that use of chlortetracyclin (penicillin) becomes less likely (negative product synergy); assuming that there is no decrease in bacterial growth while using penicillin (chlortetracyclin) it holds that use of chlortetracyclin (penicillin) becomes more likely (positive product synergy). This is consistent with our knowledge of the underlying mechanisms as described in Section 2.1. Recall that the modelling of the causal mechanisms involved in hypoglycaemia, as shown in Figure 3, was done by means of the decreasing order operator. Assuming that hypoglycaemia is present in a patient forces activation of the insulin hypersecretion mechanism and inactivation of glucagon hypersecretion mechanism. As a consequence, assumptions about insulin and glucagon hypersecretion can no longer influence each other (zero product synergy). On the other hand, assuming absence of hypoglycaemia in a patient with insulin hypersecretion (glucagon hypersecretion) renders glucagon hypersecretion (insulin hypersecretion) more likely, as in that case the two complementary mechanisms have to compensate for each other (positive product synergy). Again we use the most specific result, which in this case concerns the first argument for the left-associative version of the interaction as described in Table 17. This description is again consistent with our knowledge about the underlying physiological mechanisms as described in Section 2.1.

4 The qualitative patterns

From the results obtained in the previous section, it follows that it is possible to exploit the semantics of causal independence models using Boolean operators in developing a Bayesian network fulfilling particular qualitative requirements. In this paper, we have considered 26 of the most common ones. Three different qualitative relationships were studied, with one of them, the product synergy, consisting of two relationships: one for a positive effect e and one for a negative effect $\bar{e}$. For each qualitative relationship there are 4 different possible signs. As a consequence, the maximum number of different possible interaction models, which we have called *QC patterns*, is $4^4 = 256$. The number of patterns that can be realised is determined by relationships between the relations S, Y and X.

Note that $S^\sigma(C_j, E)$, with $\sigma \in \{-, +, ?\}$, leaves $Y^{\sigma'}(\{C_i, C_j\}, E)$, $i \neq j$, undetermined. For example, assume that $\sigma = +$ it holds that

$$\Pr(e \mid C_1, \ldots, c_j, \ldots, C_n) - \Pr(e \mid C_1, \ldots, \bar{c}_j, \ldots, C_n) \geq 0 \tag{18}$$

This implied that σ' can still be anything, as (18) simply says that in the expression

$$\Pr(e \mid c_1, c_2, C_3, \ldots, C_n) + \Pr(e \mid \bar{c}_1, \bar{c}_2, C_3, \ldots, C_n) - \Pr(e \mid c_1, \bar{c}_2, C_3, \ldots, C_n) - \Pr(e \mid \bar{c}_1, c_2, C_3, \ldots, C_n)$$

we have that

$$\Pr(e \mid c_1, c_2, C_3, \ldots, C_n) \geq \Pr(e \mid \bar{c}_1, c_2, C_3, \ldots, C_n)$$

and

$$\Pr(e \mid \bar{c}_1, \bar{c}_2, C_3, \ldots, C_n) \leq \Pr(e \mid c_1, \bar{c}_2, C_3, \ldots, C_n)$$

from which it is impossible to determine the sign σ'. A similar property holds for the product synergy. Hence, additive and product synergies do indeed offer something extra.

There is, however, one exception to this, as is shown in the following proposition.

Proposition 16 *Let $S^0(C_j, E)$ be the qualitative influence that is satisfied for cause variable C_j and effect variable E in the Bayesian network $\mathcal{B} = (G, \Pr)$, then it holds that $Y^0(\{C_i, C_j\}, E)$, $X^0(\{C_i, C_j\}, e)$ and $X^0(\{C_i, C_j\}, \bar{e})$, with $i \neq j$.*

Proof: From $S^0(C_i, E)$ it follows that $\Pr(e \mid C_1, \ldots, c_j, \ldots, C_n) = \Pr(e \mid C_1, \ldots, \bar{c}_j, \ldots, C_n)$. Substituting this in the definitions of the additive and product synergy yields the requested result. □

There are also some other QC patterns which are identical to each other for the 26 Boolean functions considered; in summary the tables 3–6 and 8–17 yield 18 different patterns. These are the patterns that can be used in selecting an appropriate Boolean function in Bayesian-network design.

5 Conclusions and further research

The qualitative characteristics of interactions in Bayesian-network probability tables have been analysed and described in this paper, taking causal independence and QPNs as a foundation. This paper builds upon results regarding causal independence obtained previously by other researchers. Heckerman et al. [13, 14, 15] have previously studied causal independence assuming that the chosen interaction functions are well understood, and that their expected probabilistic behaviour matches the intuition underlying this choice. This may no longer be the case if the interactions to be modelled become more complex. Zhang and Poole have previously proposed to use algebraic methods to formalise causal independence [30]. However, the subject of Zhang and Poole's work is the algorithmic complexity of probabilistic inference – which is why they restrict to commutative and associative operators – not trying to understand the qualitative nature of causal independence. New in the present paper is, therefore, the utilisation of QPNs in a systematic analysis of probabilistic interactions in causal independence models, and this is seen as its main scientific contribution. 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 theory can thus be used in the process of designing a Bayesian network, where, dependent on the problem at hand, a particular QC pattern can be selected, and be used in the design process and in acquiring the necessary probabilistic information.

Not all QC patterns realisable by Boolean functions may have been identified. As different Boolean functions may yield identical QC patterns, 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 find more examples from reality matching the various QC patterns, such that the use of QC patterns can be more easily understood and used by Bayesian-network researchers interested in developing applications.

References

[1] S. Andreassen, M. Woldbye, B. Falck and S.K. Andersen. MUNIN — A causal probabilistic network for interpretation of electromyographic findings. *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, Morgan-Kaufman, 1987, pp. 366–372.

[2] B.G. Buchanan and E.H. Shortliffe (Eds.). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1984.

[3] R.G. Cowell, A.P. Dawid, S.L. Lauritzen and D.J. Spiegelhalter. *Probabilistic Networks and Expert Systems*. Springer, New York, 1999.

[4] F.J. Díez. Parameter adjustment in Bayes networks: the generalized noisy OR-gate. In: *Proc of the 9th Conference on Uncertainty in Artificial Intelligence*, 1993, pp. 99–105.

[5] F.J. Díez, J. Mira, E. Iturralde and S. Zubillage, DIAVAL: a Bayesian expert system for echocardiography, *Artificial Intelligence in Medicine*, 10 (1997) 59–73.

[6] M.J. Druzdzel and L.C. van der Gaag. Elicitation of probabilities for belief networks: combining qualitative and quantitative information. In: *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, 1995, pp. 141 – 148.

[7] M.J. Druzdzel and M. Henrion. Efficient reasoning in qualitative probabilistic networks. In: *Proceedings of the Eleventh National Conference on Artificial Intelligence*, 1993, pp. 548 – 553.

[8] D. Dubois, J. Lang and H. Prade. Possibilistic logic. In: D. Gabay et al. (eds.), *Handbook of Logics in Artificial Intelligence and Logic Programming* 3. Oxford University Press, Oxford, pp. 439–513, 1994.

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

[10] I.J. Good. *Good Thinking: the Foundations of Probability and its Applications*. University of Minnesota Press, Mineapolis, MN, 1983.

[11] D.E. Heckerman, E.J. Horvitz and B.N. Nathwani. Towards normative expert systems: part I – The Pathfinder project. *Methods of Information in Medicine*, 31 (1992) 90–105.

[12] D.E. Heckerman and B.N. Nathwani. Towards normative expert systems: part II – probability-based representations for efficient knowledge acquisition and inference. *Methods of Information in Medicine*, 31 (1992) 106–116.

[13] D. Heckerman. Causal independence for knowledge acquisition and inference. In: *Proc of the 9th Conference on Uncertainty in Artificial Intelligence*, 1993, pp. 122–127.

[14] D. Heckerman and J.S. Breese. A new look at causal independence. In: *Proc of the 10th Conference on Uncertainty in Artificial Intelligence*, 1994, pp. 286–292.

[15] 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) 826–831.

[16] M. Henrion. Some practical issues in constructing belief networks. In: J.F. Lemmer and L.N. Kanal (Eds.), *Uncertainty in Artificial Intelligence*, 3, Elsevier, Amsterdam, 1989, pp. 161–173.

[17] M. Henrion and M.J. Druzdzel. Qualitative propagation and scenario-based approaches to explanation in probabilistic reasoning. In: *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence*, 1991, pp. 17 – 32.

[18] F.V. Jensen. *Bayesian Networks and Decision Graphs*. Springer, New York, 2001

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

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

[21] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.

[22] E.H. Shortliffe and B.G. Buchanan. A model of inexact reasoning in medicine. *Mathematical Biosciences*, 23 (1975) 351–379.

[23] S.L. Lauritzen and D.J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society (Series B)*, 50, 1987, 157–224.

[24] P.J.F. Lucas, H. Boot and B.G. Taal. Computer-based decision-support in the management of primary gastric non-Hodgkin lymphoma. *Methods of Information in Medicine*, 37 (1998) 206–219.

[25] P.J.F. Lucas, N.C. de Bruijn, K. Schurink, I.M. Hoepelman. A Probabilistic and decision-theoretic approach to the management of infectious disease at the ICU. *Artificial Intelligence in Medicine*, 19(3) (2000) 251–279.

[26] S. Srinivas. A generalization of the noisy-OR model. In: *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, 1993, pp. 208–215.

[27] I. Wegener. *The Complexity of Boolean Functions*. John Wiley & Sons, New York, 1987.

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

[29] L.A. Zadeh. Fuzzy logic and approximate reasoning. *Synthese*, 30 (1975) 407–428.

[30] N.L. Zhang and D. Poole. Exploiting causal independence in Bayesian network inference. *Journal of Artificial Intelligence Research*, 5 (1996) 301–328.

[31] N. L. Zhang and Li Yan. Independence of Causal Influence and Clique Tree Propagation. *International Journal of Approximate Reasoning*, 19 (1997) 335–349.