

# A Qualitative Characterisation of Causal Independence Models using Boolean Polynomials

Marcel van Gerven, Peter Lucas, and Theo van der Weide

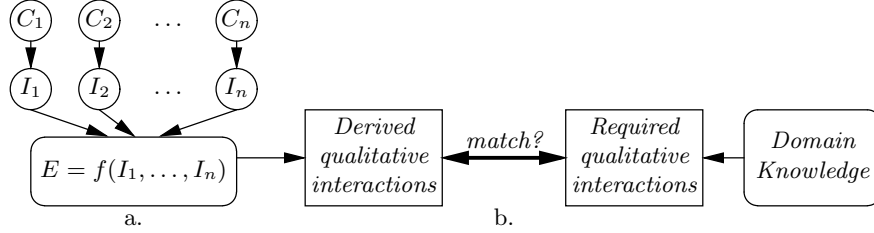
Institute for Computing and Information Sciences, Radboud University Nijmegen,  
Toernooiveld 1, 6525 ED Nijmegen, The Netherlands  
{marcelge,peter1,th.p.vanderweide}@cs.ru.nl

**Abstract.** Causal independence models offer a high level starting point for the design of Bayesian networks but are not maximally exploited as their behaviour is often unclear. One approach is to employ qualitative probabilistic network theory in order to derive a qualitative characterisation of causal independence models. In this paper we exploit polynomial forms of Boolean functions to systematically analyse causal independence models, giving rise to the notion of a polynomial causal independence model. The advantage of the approach is that it allows understanding qualitative probabilistic behaviour in terms of algebraic structure.

## 1 Introduction

Since the end of the 1980s, Bayesian networks have gained a lot of attention as models for reasoning with uncertainty. A Bayesian network is essentially a graphical specification of independence assumptions underlying a joint probability distribution, allowing for the compact representation of probabilistic information in terms of local probability tables [8]. However, in many cases the amount of probabilistic information required is still too large. The theory of *causal independence*, CI for short, offers one way to reduce this amount of probabilistic information [4]. Basically, a probability table is specified in terms of a linear number of parameters  $P(I_k | C_k)$ , as schematically indicated in Fig. 1.a, which are combined by means of a combination function  $f$ . A well-known example of a CI model is the noisy OR model, which is employed to model the disjunctive interaction of multiple independent causes of an effect [1, 5].

In principle, the choice of the combination function is free and can be any of the  $2^{2^n}$  possible Boolean functions. Given the attractive nature of the properties of causal independence models, it is regrettable that only few of the possible CI models are used in practice. This is caused by the fact that it is often unclear with what behaviour a particular CI model is endowed. In [7] *qualitative probabilistic network* (QPN) theory [10] was adopted in order to characterise the behaviour of *decomposable* CI models [4]. Such a qualitative characterisation may then be matched to the behaviour that is dictated by the domain (Fig. 1.b). In this paper, we provide an alternative, systematic characterisation of Boolean combination functions in terms of their polynomial form. The resulting models are called *polynomial CI model*. On the basis of this canonical representation, a number of important qualitative properties of CI models are derived.



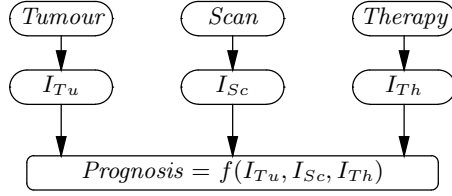
**Fig. 1.** Comparing the observed qualitative behaviour of a CI model with the desired qualitative behaviour as specified by a domain expert.

## 2 Preliminaries

In order to illustrate the theory we introduce a CI model for the domain of medical oncology. Carcinoid tumours synthesise various compounds which leads to a complex symptomatology. Patients may be diagnosed by performing a radioactive scan and can be treated by means of radiotherapy. Patients that are known to have a carcinoid tumour but have a negative radioactive scan (i.e. the tumour does not show up on the scan) will have a decreased probability of survival. This is a counter-intuitive result, which is due to the fact that given a negative radioactive scan, radiotherapy will not be effective. The CI model in Fig. 2 represents this interaction, where *Tumour* ( $Tu$ ) denotes whether or not the tumour has been identified during surgery, *Scan* ( $Sc$ ) denotes whether a radioactive scan is positive or negative and *Therapy* ( $Th$ ) denotes whether radiotherapy was or was not performed. The main task in building a CI model is then to estimate  $P(I_{Tu} | Tu)$ ,  $P(I_{Sc} | Sc)$  and  $P(I_{Th} | Th)$ , and to determine the combination function  $f(I_{Tu}, I_{Sc}, I_{Th})$  that models the interaction between these factors with respect to *Prognosis* ( $Pr$ ), where  $Pr = \top$  refers to a good prognosis and  $Pr = \perp$  refers to a poor prognosis. We will refer to this example as the *carcinoid example*.

Bayesian networks provide for a concise factorisation of a joint probability distribution over random variables. A *Bayesian network*  $B$  is defined as a pair  $B = (G, P)$ , where  $G$  is an acyclic digraph with vertices  $V(G)$  and arcs  $A(G)$  and  $P$  is a joint probability distribution over a set  $\mathbf{X}$  of random variables. It is assumed that there is a one-to-one correspondence between the vertices  $V(G)$  and the random variables  $\mathbf{X}$  such that  $P(\mathbf{X})$  factorises according to the structure of the acyclic digraph  $G$ . To simplify notation, we will use vertices  $V(G)$  and random variables in  $\mathbf{X}$  interchangeably, where the interpretation will be clear from context. In this paper it is assumed that all random variables are binary and we use  $v_i$  to denote  $V_i = \top$  and  $\bar{v}_i$  to denote  $V_i = \perp$ .

CI is the notion that causes  $\mathbf{C}$  are independently contributing to the occurrence of an effect  $E$  through some pattern of interaction. As indicated in Fig. 1.a, intermediate variables  $\mathbf{I}$  are used not only to connect causal variables  $\mathbf{C}$  to the effect variable  $E$ , but also in defining the combination function  $f$ . In this paper it is assumed that the interaction among causes is represented by means



**Fig. 2.** Prognosis of carcinoid cancer using a CI model.

of a *Boolean function*  $f : \mathbb{B}^n \rightarrow \mathbb{B}$  over the domain  $\mathbb{B} = \{\perp, \top\}$  with  $\perp < \top$ . We assign Boolean values to a set  $\mathbf{S}$  of Boolean variables by means of a *valuation*, which is a function  $v : \mathbf{S} \rightarrow \mathbb{B}$  assigning either  $\top$  or  $\perp$  to each variable in  $\mathbf{S}$ . We use  $\sum_{\mathbf{I}} g(\mathbf{I}) = \sum_{(I_1, \dots, I_n) \in \mathbb{B}^n} g(I_1, \dots, I_n)$  to denote a summation over all valuations of  $\mathbf{I}$ . A CI model is then defined as follows.

**Definition 1 (Causal independence model).** Let  $B = (G, P)$  be a Bayesian network with vertices  $V(G) = \mathbf{C} \cup \mathbf{I} \cup \{E\}$  where  $\mathbf{C}$  is a set of cause variables,  $\mathbf{I}$  is a set of intermediate variables with  $\mathbf{C} \cap \mathbf{I} = \emptyset$  and  $E \notin \mathbf{C} \cup \mathbf{I}$  denotes the effect variable. The set of arcs is given by  $A(G) = \{(C, I_C) \mid C \in \mathbf{C}\} \cup \{(I, E) \mid I \in \mathbf{I}\}$ .  $B$  is said to be a causal independence (CI) model, mediated by the combination function  $f : \mathbb{B}^n \rightarrow \mathbb{B}$  if

$$P(e \mid \mathbf{C}) = \sum_{\mathbf{I}} f(\mathbf{I}) \prod_{C \in \mathbf{C}} P(I_C \mid C). \quad (1)$$

We use  $P[f]$  to denote this probability function and assume that  $P(i_C \mid \bar{c}) = 0$  and  $P(i_C \mid c) > 0$ , where an intermediate variable  $I_C$  can be thought to inhibit the occurrence of a cause  $C$  whenever  $P(i_C \mid c) < 1$ .

Qualitative probabilistic networks (QPNs) were introduced by Wellman [10] and are a qualitative abstraction of ordinary Bayesian networks. In the following, let  $(G, P)$  be a Bayesian network, let  $A, B, C \in V(G)$  represent binary random variables and let  $(A, C)$  and  $(B, C)$  be arcs in  $G$ .

A qualitative influence expresses how the value of one vertex influences the probability of observing values for another vertex. Let  $X$  denote  $\pi_G(C) \setminus \{A\}$ . We say that there is a *positive qualitative influence* of  $A$  on  $C$  if

$$P(c \mid a, x) - P(c \mid \bar{a}, x) \geq 0$$

for all valuations  $x \in \mathbb{B}^{|X|}$ . *Negative* and *zero* qualitative influences are defined analogously, replacing  $\geq$  by  $\leq$  and  $=$  respectively. If there are valuations  $x, x' \in \mathbb{B}^{|X|}$  such that  $P(c \mid a, x) - P(c \mid \bar{a}, x) > 0$  and  $P(c \mid a, x') - P(c \mid \bar{a}, x') < 0$  then we say that the qualitative influence is *non-monotonic*. If none of these cases hold (i.e. when there is incomplete information about the probability distribution) then we say that the qualitative influence is *ambiguous*.

An additive synergy expresses how the interaction between two variables influences the probability of observing values for a third vertex. Let  $X$  denote

$\pi_G(C) \setminus \{A, B\}$ . There is a *positive additive synergy* of  $A$  and  $B$  on  $C$  if

$$P(c \mid a, b, x) + P(c \mid \bar{a}, \bar{b}, x) - P(c \mid \bar{a}, b, x) - P(c \mid a, \bar{b}, x) \geq 0$$

for all valuations  $x \in \mathbb{B}^{|X|}$ . Negative, zero, non-monotonic and ambiguous additive synergies are defined analogous to qualitative influences.

A product synergy expresses how upon observation of a common child of two vertices, observing the value of one parent vertex influences the probability of observing a value for the other parent vertex. The original definition of a product synergy is as follows [6]. Let  $X$  denote  $\pi_G(C) \setminus \{A, B\}$ . We say that there is a *positive product synergy* of  $A$  and  $B$  with regard to the value  $c_0$  of variable  $C$  if

$$P(c_0 \mid a, b, x)P(c_0 \mid \bar{a}, \bar{b}, x) - P(c_0 \mid \bar{a}, b, x)P(c_0 \mid a, \bar{b}, x) \geq 0$$

for all valuations  $x \in \mathbb{B}^{|X|}$ . Again, the other types of product synergies are defined analogous to the corresponding types of qualitative influences. Modifications to product synergies have been made after the observation that this definition is incomplete when parent vertices in  $X$  are uninstantiated [2]. However, since we are considering the CI model in isolation; i.e. we assume that a cause  $C$  is independent of  $\mathbf{C} \setminus \{C\}$ , we are entitled to use the original definition of the product synergy in the qualitative analysis of CI models.

In this paper, CI models are analysed by rewriting the combination function in terms of well-formed formulas (wffs) of propositional logic [3]. We will make use of the following concepts. Let  $b$  be a Boolean variable. A *literal*  $l$  refers to  $b$  or its negation  $\neg b$ . In the following we will also write a conjunction of literals as a set of literals  $\bigcup_{l \in m} \{l\}$  where we interpret the empty set as  $\top$ . A *monomial*  $m \equiv \bigwedge_{l \in m} l$  is a conjunction of literals  $l$ . Throughout, we will use a disjunction of monomials as a set of monomials  $\bigcup_{m \in p} \{m\}$  where we interpret the empty set as  $\perp$ . A *Boolean polynomial*  $p \equiv \bigvee_{m \in p} m$  stands for a disjunction of monomials  $m$ . We will use the equivalent notation  $p = \bigvee_{m \in p} \bigwedge_{l \in m} l \equiv \{\{l_{11}, \dots, l_{1n_1}\}, \dots, \{l_{k1}, \dots, l_{kn_k}\}\}$  to denote a Boolean polynomial. We use  $m^+$  to denote the set of positive literals in  $m$ , such that if  $l \in m^+$ , then  $l = b$  and  $m^-$  to denote the set of negative literals in  $m$ , such that if  $l \in m^-$  then  $l = \neg b$ . Since a monomial may consist of positive and negative literals, we may write  $m \equiv \bigwedge_{l \in m^+} l \wedge \bigwedge_{l \in m^-} l$ .

The relation between Boolean functions and well-formed formulas is made explicit by the fact that any Boolean function can be realised by a well-formed formula. This is guaranteed by the fact that any Boolean function can be realised by a Boolean polynomial which is in *disjunctive normal form* (DNF) [3]. A Boolean polynomial  $p$  is in DNF if every monomial in  $p$  contains the same Boolean variables and every two distinct monomials are mutually exclusive. A disadvantage of the disjunctive normal form is that in the worst case, we need to specify  $2^n$  different monomials for an  $n$ -ary Boolean function. Therefore, often the notion of *Boolean function minimisation* is employed, where we find a more compact Boolean polynomial  $p'$  that is logically equivalent to the disjunctive normal form  $p$  of some Boolean function  $f$  [9]. In this paper, we will use Boolean functions  $f$  and wffs  $\phi$  that realise  $f$  interchangeably. Particularly, we will not

distinguish between combination functions of CI models that are specified in terms of either  $f$  or  $\phi$ , where we assume a bijection  $B : \mathbf{C} \rightarrow \mathbf{B}$  between the cause variables  $\mathbf{C}$  and the Boolean variables in  $\mathbf{B}$ , which we abbreviate by  $b_C$ . We will use the notion of substitution to write  $f_\phi(\mathbf{I})$  more compactly as  $\phi(\mathbf{I})$ .

**Definition 2 (Substitution).** Let  $\phi[t_1/x_1, \dots, t_n/x_n]$  denote the simultaneous substitution of each term  $t_i$  in  $\phi$  by  $x_i$ , with  $1 \leq i \leq n$ . We will use  $\phi(\mathbf{I})$  to denote  $\phi[b_{C_1}/I_{C_1}, \dots, b_{C_n}/I_{C_n}]$  for  $\mathbf{C} = \{C_1, \dots, C_n\}$ .

Consider for instance the carcinoid example. At some point it is postulated that the combination function  $f(I_{Tu}, I_{Sc}, I_{Th})$  might be realised by the DNF:  $(\neg b_{Tu} \wedge \neg b_{Sc} \wedge \neg b_{Th}) \vee (\neg b_{Tu} \wedge \neg b_{Sc} \wedge b_{Th}) \vee (\neg b_{Tu} \wedge b_{Sc} \wedge b_{Th}) \vee (b_{Tu} \wedge b_{Sc} \wedge b_{Th})$ , expressing the background knowledge about the causal mechanism underlying the model. This DNF  $p$  is equivalent to the minimal polynomial  $p' = (\neg b_{Tu} \wedge \neg b_{Sc}) \vee (b_{Sc} \wedge b_{Th})$ . We may then write  $p'(i_{Tu}, \bar{i}_{Sc}, i_{Th})$  to denote the substitution of  $b_{Tu}$  by  $\top$ ,  $b_{Sc}$  by  $\perp$  and  $b_{Th}$  by  $\top$  in  $p'$ , which evaluates to  $(\perp \wedge \top) \vee (\perp \wedge \top) = \perp$ .

### 3 Polynomial CI Models

In this section, we introduce *polynomial CI models*. These models enable us to zoom in on the characteristics of Boolean functions mediating a CI model. In the next section, we will derive the qualitative properties of these polynomial CI models. We will first prove a number of general properties of CI models. For the sake of readability we will often write  $P[\phi]$  instead of  $P[\phi](e \mid \mathbf{C})$ , and if we state a property of  $P[\phi]$  then the property holds for all valuations of  $\mathbf{C}$ . We list most properties without proof due to space considerations.

**Lemma 1.**  $P[\neg\phi] = 1 - P[\phi]$ .

**Lemma 2.**  $P[\phi \vee \psi] = 1 - P[\neg\phi \wedge \neg\psi] = P[\phi] + P[\psi] - P[\phi \wedge \psi]$ .

**Lemma 3.** If  $\phi \wedge \psi = \perp$  then  $P[\phi \vee \psi] = P[\phi + \psi] = P[\phi] + P[\psi]$ .

**Lemma 4.**  $P[\phi - \psi] = P[\phi] - P[\psi]$ .

**Lemma 5.**  $P[\phi \wedge \psi] \leq P[\phi]$ .

In general, we can model the behaviour of an combination function in terms of any equivalent wff using the basis functions  $\vee, \wedge$  and  $\neg$ , but in this paper, we will resort to the use of Boolean polynomials. We will use  $l_m(C)$  to refer to a literal in a monomial  $m$  that is associated with a cause variable  $C$ , where  $l_m(C) = b_C$  if  $b_C \in m$ ,  $l_m(C) = \neg b_C$  if  $\neg b_C \in m$  and  $l_m(C) = \top$  otherwise.

We refer to a CI model that employs a Boolean polynomial  $p$  as its combination function as a *polynomial CI model*. The probability of observing an effect  $E$  given causes  $\mathbf{C}$  for such a model is determined by the following proposition.

**Proposition 1.** For a polynomial CI model mediated by  $p$  it holds that

$$P[p](e \mid \mathbf{C}) = 1 - \sum_{\mathbf{I}} \prod_{m \in p} (1 - \prod_{l \in m^+} l(\mathbf{I}) \prod_{l \in m^-} l(\mathbf{I})) P(\mathbf{I} \mid \mathbf{C}). \quad (2)$$

*Proof.* By DeMorgan's law,  $p$  is equivalent to  $\neg \bigwedge_{m \in p} \neg m$ . From lemma 1 it then follows that  $P[p](e \mid \mathbf{C}) = P[\neg \bigwedge_{m \in p} \neg m](e \mid \mathbf{C}) = 1 - P[\bigwedge_{m \in p} \neg m](e \mid \mathbf{C})$ . Due to the analogy between Boolean algebra and ordinary logic we may write  $\bigwedge_{m \in p} \neg m$  as  $\prod_{m \in p} (1 - m(\mathbf{I}))$ . Likewise, and using the equivalence of  $m$  and  $\bigwedge_{l \in m^+} l \wedge \bigwedge_{l \in m^-} \neg l$  we may write  $m(\mathbf{I})$  as  $\prod_{l \in m^+} l(\mathbf{I}) \prod_{l \in m^-} \neg l(\mathbf{I})$ . By plugging this in into the previous equation we obtain the required result.  $\square$

The use of Boolean polynomials instead of Boolean functions is valid since any Boolean function can be realised by a Boolean polynomial in DNF. The properties of the DNF lead to a different form of Equation (2).

**Proposition 2.** *If for a polynomial CI model mediated by  $p$  it holds that  $m \wedge m' \equiv \perp$  for all  $m, m' \in p$  with  $m \neq m'$  then  $P[p] = \sum_{m \in p} P[m]$ .*

*Proof.* Let  $p$  be such that in  $\forall m, m' \in p: m \neq m' \Rightarrow m \wedge m' \equiv \perp$ . Then, according to lemma 3,  $P[m_1 \vee \dots \vee m_k](e \mid \mathbf{C})$  equals  $\sum_{m \in p} P[m](e \mid \mathbf{C})$ .  $\square$

We may compute the probability that a monomial yields  $\top$  given a valuation of the causes  $\mathbf{C}$  by

$$P[m](e \mid \mathbf{C}) = \prod_{l_m(C) \in m^+} P(i_C \mid C) \prod_{l_m(C) \in m^-} P(\bar{i}_C \mid C). \quad (3)$$

We list the following two properties of polynomial CI models, as they are used in the proof of qualitative properties in the next section.

**Proposition 3.** *Let  $B$  be a polynomial CI model mediated by  $p$ . If  $\forall m \in p: m^+ \neq \emptyset$  then we can choose a valuation  $\mathbf{c}$  of  $\mathbf{C}$  such that  $P[p](e \mid \mathbf{c}) = 0$ .*

**Proposition 4.** *Let  $B$  be a polynomial CI model mediated by a polynomial  $p \neq \perp$ . Then, there is some valuation  $\mathbf{c}$  of  $\mathbf{C}$  such that  $P[p](e \mid \mathbf{c}) > 0$ .*

## 4 Qualitative Behaviour of Polynomial CI Models

CI models will now be described qualitatively in terms of concepts taken from QPN theory. Note that we can assume that the causes are direct parents of  $E$  as the intermediate variables are marginalised out in the final computation of  $P[f](e \mid \mathbf{C})$  (cf. Equation (1)). For our analysis, we assume some fixed CI model over a set  $\mathbf{C}$  of  $n$  cause variables, in which we focus on the interaction between different cause variables  $C$  and  $C'$  and the effect variable  $E$ , where we abbreviate  $I_C$  by  $I$  and  $I_{C'}$  by  $I'$ . Throughout this paper we will use  $\mathbf{C}^1$  to denote  $\mathbf{C} \setminus \{C\}$  and  $\mathbf{C}^2$  to denote  $\mathbf{C} \setminus \{C, C'\}$ . Likewise, we will use  $\mathbf{I}^1$  to denote  $\mathbf{I} \setminus \{I\}$  and  $\mathbf{I}^2$  to denote  $\mathbf{I} \setminus \{I, I'\}$ . We use  $\mathbf{c}$  to denote a valuation of  $\mathbf{C}^1$  or  $\mathbf{C}^2$ , where the interpretation will be clear from context. We will also use the notion of a *curry*  $f_{x_1=v_1, \dots, x_k=v_k}(\mathbf{x})$  with  $x_1, \dots, x_k \in \mathbf{x}$  to denote the function  $f(\mathbf{x})$  where  $x_i$  is set to  $v_i$  for  $1 \leq i \leq k$ . For example, let  $I$  and  $I'$  be the intermediate variables as defined above and let  $f(I, I')$  be a Boolean function. Then, the curry  $f_i(I')$  is the function  $f(\perp, I')$ . In the following sections we will analyse the different types of qualitative interactions in CI models. We remark that the listed conditions are sufficient but may not be necessary. We will therefore use the ambiguous category to collect those interactions for which the qualitative behaviour is uncertain.

## 4.1 Qualitative Influences

A qualitative influence  $\sigma_C$  between a cause  $C$  and effect  $E$  denotes how the observation of  $C$  influences the observation of the effect  $e$ . The sign of a qualitative influence for a CI model mediated by  $f$  is then determined by the sign of

$$\delta_C(\mathbf{C}^1) = P[f](e | c, \mathbf{C}^1) - P[f](e | \bar{c}, \mathbf{C}^1) \quad (4)$$

such that there is a *positive* qualitative influence ( $\sigma_C = +$ ) if the sign of  $\delta_C(\mathbf{C}^1)$  is zero or positive for every valuation of  $\mathbf{C}^1$ . *Negative* ( $\sigma_C = -$ ), *zero* ( $\sigma_C = 0$ ), *ambiguous* ( $\sigma_C = ?$ ) and *non-monotonic influences* ( $\sigma_C = \sim$ ) are defined analogously. The analysis requires that we isolate the contribution of a cause variable  $C$  with respect to the effect  $E$ . By writing

$$P[f](e | C, \mathbf{C}^1) = P[f_{\bar{i}}](e | \mathbf{C}^1) + P(i | C)P[\Delta_C(f)](e | \mathbf{C}^1) \quad (5)$$

where  $\Delta_C(f)$  denotes the *difference function*  $f_i - f_{\bar{i}}$ , we obtain this isolation. Additionally, we isolate the contribution of a variable  $I$  to the results of a Boolean function  $f$ . To this end, we use the following notation regarding the isolation of one Boolean variable associated with a cause variable  $C$  and a polynomial  $p$ .  $q_C \equiv \{m \setminus \{l_m(C)\} \mid m \in p, l_m(C) \in m^+\}$  represents those monomials where  $l_m(C)$  is positive,  $q_{\bar{C}} \equiv \{m \setminus \{l_m(C)\} \mid m \in p, l_m(C) \in m^-\}$  represents those monomials where  $l_m(C)$  is negative and  $q_{\dot{C}} \equiv \{m \mid m \in p, l_m(C) \notin m\}$  represents those monomials where  $l_m(C)$  is absent. Let  $X \in \{C, \bar{C}, \dot{C}\}$ . We use  $p_X \equiv \{m \setminus \{l_m(C)\} \mid m \in q_X\}$  to denote  $q_X$  from which  $l_m(C)$  is removed and  $\bar{p}_X \equiv \{m \setminus \{l_m(C)\} \mid m \in p, m \notin q_X\}$  to denote those monomials that do not occur in  $q_X$ , where again  $l_m(C)$  is removed from the monomials. For instance, in the minimal polynomial  $p = (-b_{Tu} \wedge \neg b_{Sc}) \vee (b_{Sc} \wedge b_{Th})$  of the carcinoid example we have  $p_{\bar{T}u} = \{\{-b_{Sc}\}\}$ ,  $p_{Sc} = \{\{b_{Th}\}\}$  and  $p_{Th} = \{\{-b_{Tu}, \neg b_{Sc}\}\}$ . Using this notation, we can decompose a Boolean polynomial  $p$  as follows:

$$p(I, \mathbf{I}^1) = ((I \wedge p_C) \vee (\neg I \wedge p_{\bar{C}}) \vee p_{\dot{C}})(\mathbf{I}^1). \quad (6)$$

If we substitute (5) into (4) and under the assumption that  $P(i | c) > P(i | \bar{c})$  we obtain  $P[\Delta_C(f)](e | \mathbf{C}^1)$  as the specialisation of (4) to qualitative influences in CI models. We may further specialise this to polynomial CI models. The difference  $\Delta_C(f)$  is non-zero if either  $f_i(\mathbf{I}^1) = \top$  and  $f_{\bar{i}}(\mathbf{I}^1) = \perp$  or  $f_i(\mathbf{I}^1) = \top$  and  $f_{\bar{i}}(\mathbf{I}^1) = \perp$ . With the use of (6), this leads to  $\Delta_C(f) = (p_C \wedge \neg \bar{p}_C) - (p_{\bar{C}} \wedge \neg \bar{p}_{\bar{C}})$ . Then, using lemma 4, the sign of the qualitative influence for polynomial CI models, is determined by the sign of

$$d_C(\mathbf{C}^1) = P[p_C \wedge \neg \bar{p}_C](e | \mathbf{C}^1) - P[p_{\bar{C}} \wedge \neg \bar{p}_{\bar{C}}](e | \mathbf{C}^1). \quad (7)$$

Lemma 6 then lists a sufficient condition for observing a positive value of  $d_C(\mathbf{C}^1)$ .

**Lemma 6.** *If  $\exists m \in p_C \forall m' \in \bar{p}_C : m^+ \wedge \neg m'^+$  then  $\exists_{\mathbf{c} \in \mathbb{B}^{n-1}} : d_C(\mathbf{c}) > 0$ .*

This follows from the observation that according to lemmas 3 and 5, we can find a valuation of causes such that  $P[p_{\bar{C}} \wedge \neg \bar{p}_{\bar{C}}](e | \mathbf{c}) = 0$ , reducing (7) to  $P[p_C \wedge \neg \bar{p}_C](e | C)$ , which is larger than zero for some valuation of causes and intermediate variables. The same reasoning holds for negative values of  $d_C(\mathbf{C}^1)$ .

**Table 1.** Determining the qualitative influences for the carcinoid example.

Condition	<i>Tumour</i>	<i>Scan</i>	<i>Therapy</i>
1	$b_{Sc}$	$b_{Tu} \vee b_{Th}$	$\top$
2	$\top$	$\neg b_{Th} \vee \neg b_{Tu}$	$\neg b_{Sc}$
$\sigma_C$	$-$	$?$	$+$

**Lemma 7.** *If  $\exists m \in p_C \forall m' \in \bar{p}_C : m^+ \wedge \neg m'^+$  then  $\exists c \in \mathbb{B}^{n-1} : d_C(c) < 0$ .*

We may use Equation (7) to derive the following proposition, characterising the qualitative influences for polynomial CI models.

**Proposition 5.** *Qualitative influences are characterised as follows:*

1. *If  $p_{\bar{C}} \Rightarrow \bar{p}_{\bar{C}}$  then  $\sigma_C = +$ .*
2. *If  $p_C \Rightarrow \bar{p}_C$  then  $\sigma_C = -$ .*
3. *If (1) and (2) hold, then  $\sigma_C = 0$ .*
4. *If lemmas 6 and 7 hold then  $\sigma_C = \sim$ .*
5.  *$\sigma_C = ?$ , otherwise.*

We prove just case (1), since case (2) proceeds analogously and the rest follows directly from the definitions of the different types of qualitative influences. Case (1) states that  $p_{\bar{C}} \Rightarrow \bar{p}_{\bar{C}}$ , which is equal to  $\neg p_{\bar{C}} \vee \bar{p}_{\bar{C}}$  or  $\neg(p_{\bar{C}} \wedge \neg \bar{p}_{\bar{C}})$ . But then (7) reduces to  $P[p_C \wedge \neg \bar{p}_C](e \mid \mathbf{C}^1) - P[\perp](e \mid \mathbf{C}^1) \geq 0$ , since  $P[\perp](e \mid \mathbf{C}^1) = 0$ . Therefore, the sign of the qualitative influence is positive.

We illustrate these results with the carcinoid example. Using proposition 5 we can easily determine the signs of the qualitative influences. The conditions of proposition 5 and the outcomes for the clinical variables are listed in Table 1. Recall the conventions that the empty monomial  $\emptyset$  is equal to  $\top$ , whereas the empty polynomial  $\emptyset$  is equal to  $\perp$ . For instance, we determine condition 2 for the clinical variable *Tumour* by  $p_{Tu} \Rightarrow \bar{p}_{Tu} \vee p_{Tu}$ , which is equal to  $\perp \Rightarrow \neg b_{Sc} \vee (b_{Sc} \wedge b_{Th})$ , or  $\top$ . Table 1 represents the situations in which a qualitative influence is positive, negative or ambiguous. The results show that observing a tumour has a negative effect on patient prognosis. The qualitative influence of a scan on prognosis cannot be determined by proposition 5 alone. We may then use lemmas 6 and 7 to determine whether there is a non-monotonicity present. However, the condition  $\exists m \in p_{Sc} \forall m' \in \bar{p}_{Sc} : m^+ \wedge \neg m'^+$  does not hold since  $b_{Th} \wedge \neg \top = \perp$ . This implies that the qualitative influence of a scan on patient prognosis is of the ambiguous type. *Therapy* has a positive qualitative influence on patient prognosis. Note that if the scan is negative then the influence of therapy on prognosis is zero, since a therapy is only fruitful when the scan is positive.



## 4.2 Additive Synergies

Additive synergies express how two cause variables jointly influence the probability of observing the effect. The additive synergy  $\sigma_{C,C'}$  between two causes  $C$  and  $C'$  is determined by

$$\begin{aligned} \delta_{C,C'}(\mathbf{C}^2) &= P[f](e | c, c', \mathbf{C}^2) + P[f](e | \bar{c}, \bar{c}', \mathbf{C}^2) - \\ &P[f](e | \bar{c}, c', \mathbf{C}^2) - P[f](e | c, \bar{c}', \mathbf{C}^2) \end{aligned} \quad (8)$$

where the different types of additive synergies are defined similarly to the different types of qualitative influences. The analysis requires an isolation of  $C$  and  $C'$ . We apply the decomposition (5) twice and obtain by straight computation:

$$\begin{aligned} P[f] &= P(i | C)P(i' | C')P[\Delta_{C,C'}(f)] + P[f_{i,\bar{i}'}] + \\ &P(i | C)P[\Delta_C(f_{i'})] + P(i' | C')P[\Delta_{C'}(f_{\bar{i}})], \end{aligned} \quad (9)$$

where the *difference function*  $\Delta_{C,C'}(f) = f_{i,i'} + f_{\bar{i},\bar{i}'} - f_{\bar{i},i'} - f_{i,\bar{i}'}$ , can also be expressed as  $\Delta_{C'}(f_i) - \Delta_{C'}(f_{\bar{i}})$  or  $\Delta_C(f_{i'}) - \Delta_C(f_{\bar{i}'})$ . With regard to the analysis of Boolean variables associated with  $C$  and  $C'$  we introduce the following notation. Let  $X \in \{C, \bar{C}, \dot{C}\}$  and  $Y \in \{C', \bar{C}', \dot{C}'\}$ . Then  $p_{X,Y} \equiv (p_X)_Y$  refers to polynomials in which both  $X$  and  $Y$  are present,  $p_{X|Y} \equiv p_{X,Y} \cup p_{\dot{C},Y} \cup p_{X,\dot{C}'}$  refers to polynomials in which both or either of  $X$  and  $Y$  are present and  $p_{X;Y} \equiv p_{X|Y} \cup p_{\dot{C},\dot{C}'}$  refers to polynomials in which both, either or none of  $X$  and  $Y$  are present. We use  $\bar{p}_{X,Y} \equiv \{m \setminus \{l_m(C), l_m(C')\} \mid m \in p, m \notin q_X \cap q_Y\}$  to refer to the complement  $q_{X,Y}$  from which literals  $l_m(C)$  and  $l_m(C')$  are removed. For instance, for the minimal polynomial associated with the running example we have  $p_{Tu,Sc} = \{\emptyset\}$ ,  $p_{\bar{T}u|Sc} = \{\{b_{Th}\}\}$ ,  $p_{\bar{S}c;Th} = \{\{-b_{Tu}\}\}$  and  $\bar{p}_{Tu,Th} = \{\{-b_{Sc}\}, \{b_{Sc}\}\}$ . Now we can decompose a Boolean polynomial  $p$  as follows:

$$\begin{aligned} p(I, I', \mathbf{I}^2) &= (I \wedge I' \wedge p_{C;C'}) \vee (\neg I \wedge I' \wedge p_{\bar{C};C'}) \vee \\ &(I \wedge \neg I' \wedge p_{C;\bar{C}'}) \vee (\neg I \wedge \neg I' \wedge p_{\bar{C};\bar{C}'})(\mathbf{I}^2). \end{aligned} \quad (10)$$

By inserting (9) into (8), and under the assumptions that  $P(i | c) > P(i | \bar{c})$  and  $P(i' | c') > P(i' | \bar{c}')$  we obtain  $P[\Delta_{C,C'}(f)](e | \mathbf{C}^2)$  for computing the sign of the additive synergy in CI models. In terms of polynomials, we can write  $\Delta_{C,C'}(f)$  using (10) as:  $p_{C;C'} + p_{\bar{C};\bar{C}'} - p_{\bar{C};C'} - p_{C;\bar{C}'}$ . This difference is positive if either  $p_1 = p_{C|C'} \wedge p_{\bar{C}|\bar{C}'} \wedge \neg(p_{\bar{C};C'} \wedge p_{C;\bar{C}'})$  or  $p_2 = p_{C;C'} \wedge \neg p_{\bar{C};C'}$  or  $p_3 = p_{\bar{C};\bar{C}'} \wedge \neg p_{\bar{C};C'}$  hold. The difference is negative if either  $p_4 = p_{\bar{C}|\bar{C}'} \wedge p_{C|\bar{C}'} \wedge \neg(p_{C;C'} \wedge p_{\bar{C};\bar{C}'})$  or  $p_5 = p_{\bar{C};C'} \wedge \neg p_{\bar{C};\bar{C}'}$  or  $p_6 = p_{C;\bar{C}'} \wedge \neg p_{\bar{C};\bar{C}'}$  holds. As these cases are mutually exclusive, this results in the following equation:

$$\begin{aligned} d_{C,C'}(\mathbf{C}^2) &= P[p_1](e | \mathbf{C}^2) + P[p_2](e | \mathbf{C}^2) + P[p_3](e | \mathbf{C}^2) - \\ &P[p_4](e | \mathbf{C}^2) - P[p_5](e | \mathbf{C}^2) - P[p_6](e | \mathbf{C}^2). \end{aligned} \quad (11)$$

We proceed by examining the positive and negative contributions to (11). We use  $(U, V) \in \{(C, C'), (\bar{C}, \bar{C}')\}$  and  $(X, Y) \in \{(C, C'), (C', C)\}$  in the following.

**Lemma 8.**  $\exists_{\mathbf{c} \in \mathbb{B}^{n-2}} : d_{C,C'}(\mathbf{c}) > 0$  if any of the following cases hold:

1.  $\exists_{m \in p_{U,V}} \forall_{m' \in \bar{p}_{U,V}} : m^+ \wedge \neg m'^+$ .
2.  $\exists_{m_u \in p_{C|C'}, m_v \in p_{C|\bar{C}'}} \forall_{m \in p_{X,\bar{Y}}} : m_u^+ \wedge m_v^+ \wedge \neg m^+$ .

This lemma can be proved using the same line of thought as the proof of lemma 6. The second case is just the decomposition of  $p_1$ .

**Lemma 9.**  $\exists_{\mathbf{c} \in \mathbb{B}^{n-2}} : d_{C,C'}(\mathbf{c}) < 0$  if any of the following cases hold:

1.  $\exists_{m \in p_{X,\bar{Y}}} \forall_{m' \in \bar{p}_{X,\bar{Y}}} : m^+ \wedge \neg m'^+$ .
2.  $\exists_{m_u \in p_{C|C'}, m_v \in p_{C|\bar{C}'}} \forall_{m \in p_{U,V}} : m_u^+ \wedge m_v^+ \wedge \neg m^+$ .

The characterisation of additive synergies is analogous to that of qualitative influences and follows from Equation (11).

**Proposition 6.** *Additive synergies are characterised as follows:*

1. If  $p_{\bar{C},C'} \Rightarrow \bar{p}_{\bar{C},C'}$  and  $p_{C,\bar{C}'} \Rightarrow \bar{p}_{C,\bar{C}'}$  and  $p_{\bar{C}|C'} \wedge p_{C|\bar{C}'} \Rightarrow p_{C,C'} \wedge p_{\bar{C};\bar{C}'}$  hold then  $\sigma_{C,C'} = +$ .
2. If  $p_{C,C'} \Rightarrow \bar{p}_{C,C'}$  and  $p_{\bar{C},\bar{C}'} \Rightarrow \bar{p}_{\bar{C},\bar{C}'}$  and  $p_{C|C'} \wedge p_{\bar{C}|\bar{C}'} \Rightarrow p_{\bar{C};C'} \wedge p_{C;\bar{C}'}$  hold then  $\sigma_{C,C'} = -$ .
3. If (1) and (2) hold, then  $\sigma_{C,C'} = 0$ .
4. If lemmas 8 and 9 hold then  $\sigma_{C,C'} = \sim$ .
5.  $\sigma_{C,C'} = ?$ , otherwise.

We determine the signs of the additive synergies for the carcinoid example using this proposition. *Tumour* and *Scan* are then found to exhibit a positive additive synergy. This is because observing a tumour and a positive scan or not observing a tumour and having a negative scan is in general better for prognosis than observing one of both. A positive additive synergy between *Scan* and *Therapy* is caused by the fact that they also amplify each other; i.e. a positive scan and the administration of therapy will yield a better prognosis than when either one of both is present. A zero additive synergy between *Tumour* and *Therapy* is caused by the fact that  $b_{Sc}$  renders both independent; i.e. if a scan is negative, then the prognosis is dependent on *Tumour* only, whereas if a scan is positive, then the prognosis is dependent on *Therapy* only.

### 4.3 Product Synergies

Product synergies describe the dependence between two causes when the value of the effect variable is observed. The sign  $\sigma_{C,C'}^e$  of a product synergy between  $C$  and  $C'$  is determined by

$$\delta_{C,C'}^E(\mathbf{C}^2) = P[f](E | c, c', \mathbf{C}^2)P[f](E | \bar{c}, \bar{c}', \mathbf{C}^2) - P[f](E | \bar{c}, c', \mathbf{C}^2)P[f](E | c, \bar{c}', \mathbf{C}^2) \quad (12)$$

where the different types of product synergies are defined similarly to the different types of qualitative influences. For binary variables,  $\sigma_{C,C'}^e$  is fully determined

by  $\sigma_{C,C'}^e$  and  $\sigma_{C,C'}$  through the equation  $\delta_{C,C'}^{\bar{e}}(\mathbf{C}^2) = \delta_{C,C'}^e(\mathbf{C}^2) - \delta_{C,C'}(\mathbf{C}^2)$  and we will therefore restrict ourselves to the case where  $E = \top$ . According to (9) and under the standard assumptions, we can compute the product synergy by:

$$P[\Delta_{C,C'}(f)](e \mid \mathbf{C}^2)P[f_{\bar{i},\bar{i}'}](e \mid \mathbf{C}^2) - P[\Delta_C(f_{\bar{i}'})](e \mid \mathbf{C}^2)P[\Delta_{C'}(f_{\bar{i}})](e \mid \mathbf{C}^2).$$

As  $\Delta_C(f_{\bar{i}'}) = f_{i,\bar{i}'} - f_{\bar{i},\bar{i}'}$ ,  $\Delta_{C'}(f_{\bar{i}}) = f_{\bar{i},i'} - f_{\bar{i},\bar{i}'}$ , and  $\Delta_{C,C'}(f) = f_{i,i'} + f_{\bar{i},\bar{i}'} - f_{\bar{i},i'} - f_{i,\bar{i}'}$  we can alternatively write this as

$$P[f_{i,i'}](e \mid \mathbf{C}^2)P[f_{\bar{i},\bar{i}'}](e \mid \mathbf{C}^2) - P[f_{\bar{i},i'}](e \mid \mathbf{C}^2)P[f_{i,\bar{i}'}](e \mid \mathbf{C}^2),$$

which, with the use (10), reduces for polynomial CI models to

$$d_{C,C'}^e(\mathbf{C}^2) = P[p_{C,C'}](e \mid \mathbf{C}^2)P[p_{\bar{C},\bar{C}'}](e \mid \mathbf{C}^2) - P[p_{\bar{C},C'}](e \mid \mathbf{C}^2)P[p_{C,\bar{C}'}](e \mid \mathbf{C}^2). \quad (13)$$

Again, we determine conditions for which  $d_{C,C'}^e(\mathbf{C}^2)$  is positive or negative. The lemmas follow from (13) and their proof is analogous to that of lemma 6. We use  $(X, Y) \in \{(C, C'), (C', C)\}$  in the following.

**Lemma 10.**  $\exists_{\mathbf{c} \in \mathbb{B}^{n-2}}: d_{C,C'}^e(\mathbf{c}) > 0$  if any of the following cases hold:

1.  $\exists m_u \in p_{X,Y}, m_v \in p_{\bar{X},\bar{Y}} \forall m \in p_{\bar{X},Y}: m_u^+ \wedge m_v^+ \wedge \neg m^+$ .
2.  $\exists m_u \in p_{\bar{X},Y}, m_v \in p_{\bar{X},\bar{Y}} \forall m \in p_{X,\bar{Y}}: m_u^+ \wedge m_v^+ \wedge \neg m^+$ .

**Lemma 11.**  $\exists_{\mathbf{c} \in \mathbb{B}^{n-2}}: d_{C,C'}^e(\mathbf{c}) < 0$  if any of the following cases hold:

1.  $\exists m_u \in p_{\bar{X},Y}, m_v \in p_{X,\bar{Y}} \forall m \in p_{\bar{X},Y}: m_u^+ \wedge m_v^+ \wedge \neg m^+$ .
2.  $\exists m_u \in p_{\bar{X},Y}, m_v \in p_{X,\bar{Y}} \forall m \in p_{\bar{X},\bar{Y}}: m_u^+ \wedge m_v^+ \wedge \neg m^+$ .

The characterisation of product synergies is analogous to that of qualitative influences and additive synergies and follows from Equation (13).

**Proposition 7.** *Product synergies are characterised as follows:*

1. If either  $p_{\bar{X},\bar{Y}} \vee p_{\bar{X},Y} \Rightarrow p_{X,Y}$  and  $p_{X,\bar{Y}} \vee p_{X,Y} \Rightarrow p_{\bar{X},\bar{Y}}$  or  $\neg p_{U,V}$  with  $(U, V) \in \{(\bar{C}, C'), (C, \bar{C}')\}$  holds then  $\sigma_{C,C'}^e = +$ .
2. If either  $p_{\bar{X},Y} \vee p_{X,Y} \Rightarrow p_{X,\bar{Y}}$  and  $p_{\bar{X},\bar{Y}} \vee p_{\bar{X},Y} \Rightarrow p_{\bar{X},Y}$  or  $\neg p_{U,V}$  with  $(U, V) \in \{(C, C'), (\bar{C}, \bar{C}')\}$  holds then  $\sigma_{C,C'}^e = -$ .
3. If both (1) and (2) hold then  $\sigma_{C,C'}^e = 0$ .
4. If lemmas 10 and 11 hold then  $\sigma_{C,C'}^e = \sim$ .
5.  $\sigma_{C,C'}^e = ?$ , otherwise.

For product synergies we use proposition 7 to determine the signs of the product synergies for the carcinoid example. We find a positive product synergy between *Tumour* and *Scan*, which is caused by the fact that given a good prognosis, it is more likely that a tumour is accompanied by a positive scan rather than that a tumour is accompanied by a negative scan. The positive product synergy between *Scan* and *Therapy* is caused by the fact that given a good prognosis, it is more likely that a positive scan is accompanied by therapy rather than that a positive scan is not accompanied by therapy. The positive product synergy between *Tumour* and *Therapy* is caused by the fact that given a good prognosis, it is more likely that the tumour is present and therapy is given rather than that the tumour is present and no therapy is given.

## 5 Conclusions

In this paper we analysed the qualitative properties of Boolean CI models. Polynomial CI models, where the combination function is rewritten in terms of a Boolean polynomial, were introduced. They enable the analysis of a CI model's qualitative characteristics by examining the structure of the Boolean polynomial. Qualitative influences, additive synergies and product synergies were examined and conditions under which positive, negative, zero, non-monotonic and ambiguous signs are observed were determined. This facilitates the use of CI models in the construction of Bayesian networks since one can determine whether a particular model fulfils a qualitative specification of cause-effect interactions. The carcinoid example illustrated the usefulness of the theory in practice.

## References

1. F. J. Díez. Parameter adjustment in Bayes networks. the generalized noisy or-gate. In *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence*, San Francisco, CA, 1993. Morgan Kaufmann Publishers.
2. M. J. Druzdzel and M. Henrion. Intercausal reasoning with uninstantiated ancestor nodes. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 317–325. Morgan Kaufmann Publishers, San Mateo, California, 1993.
3. H. B. Enderton. *A Mathematical Introduction to Logic*. Academic Press, Inc., 1972.
4. D. Heckerman and J. Breese. Causal independence for probability assessment and inference using Bayesian networks. *IEEE, Systems, Man, and Cybernetics*, 26:826–831, 1996.
5. M. Henrion. Some practical issues in constructing belief networks. In *Proceedings of the Third Conference on Uncertainty in Artificial Intelligence*, pages 161–173. Elsevier, Amsterdam, 1989.
6. 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*, pages 17–32, 1991.
7. P.J.F. Lucas. Bayesian network modelling by qualitative patterns. *Artificial Intelligence*, 163:233–263, 2005.
8. J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, 1988.
9. I. Wegener. *The Complexity of Boolean Functions*. John Wiley & Sons, New York, 1987.
10. M.P. Wellman. Fundamental concepts of qualitative probabilistic networks. *Artificial Intelligence*, 44:257–303, 1990.