数理解析研究所講究録
第1603巻 2008年 120-129

Computation of Conditional Independence Using Cain Polynomials

千葉大学・大学院理学研究科　汪　金芳（Jinfang Wang）
Graduate School of Science, Chiba University

1 Introduction

The concept of probabilistic conditional independence (PCI) is a key concept in many areas such as the study of probabilistic theory of causation. Dawid (1979) is among the first to give axiomatic treatments of PCI. See Dawid (1988) for further developments. Pearl and Paz (1987) proposed a closely related axiomatic system termed the graphoid, which plays an important role in graphical modeling of statistical causation. For instance the contraction property of the graphoid says that $x \perp\!\!\!\perp y | z$ and $x \perp\!\!\!\perp z$ jointly imply $x \perp\!\!\!\perp (y \vee z)$. We shall see later that these PCI relations are isomorphic to the following so called cain polynomial equations, $f_1 = xyz - xz - yz + z = 0$, $f_2 = xz - x - z = 0$ and $g = xyz - yz - x = 0$, respectively. Cain polynomials are defined in §2. Note that $g = f_1 + f_2$, where addition is performed in a usual way. One property of the cain polynomials states that $f_1 = 0, f_2 = 0$ implies $f_1 + f_2 = 0$. Thus we have proved in an 'algebraic fashion' that $x \perp\!\!\!\perp y|z, x \perp\!\!\!\perp z \Rightarrow x \perp\!\!\!\perp (y \vee z)$. In Section 6 we shall see how to derive $x \perp\!\!\!\perp (y \vee z) \Rightarrow x \perp\!\!\!\perp (y \vee z)$.

The purpose of the paper is to give a formal study on deriving PCI relations using the cain polynomials. §2 defines the concept of cain polynomials. The cain algebra is introduced in §3. PCI are studies in §4 based on the cain algebra. §5 connects the two algebraic systems, namely the cain and the cain polynomials. §6 studies methods for deriving PCI relations based on the cain polynomials.

2 Cain Polynomials

Let $\boldsymbol{x} = \{x_1, \ldots, x_p\}$, where $p \geq 1$, and $\mathbb{L}$ be the Boolean lattice of $\boldsymbol{x}$. Let

$$\{0,1\}^p = \{(a_1, \ldots, a_p)\,;\, a_1, \ldots, a_p \in \{0,1\}\}$$

The elements of $\{0,1\}^p$ are denoted by α, β, γ, etc. Let $e_i(\alpha)$ denote the $i$th component of $\alpha \in \{0,1\}^p$. Let $\mathbf{1} = (1, \ldots, 1)$ and $\mathbf{0} = (0, \ldots, 0)$. There is a one-to-one correspondence between $\mathbb{L}$ and $\{0,1\}^p$: for any $A \in \mathbb{L}$, a subset of $\boldsymbol{x}$, there is a unique $\alpha(A) \in \{0,1\}^p$ so that $\alpha(A)$ has its $i$th element $e_i(\alpha(A)) = 1$ if and only if $x_i \in A$. In particular, $\mathbf{1}$ corresponds to $\boldsymbol{x}$ and $\mathbf{0}$ to the empty set $\emptyset$.

Definition 2.1 (cain polynomial). *An expression,*

$$\boldsymbol{x}^\alpha = x_1^{e_1(\alpha)} \ldots x_p^{e_p(\alpha)},$$

where $\alpha \in \{0,1\}^p$, is called a cain monomial. *If $e_i(\alpha) = 0$, we omit $x_i^{e_i(\alpha)}$. As a convention, let*

$$\boldsymbol{x}^{\mathbf{0}} = 0\,.$$

A cain polynomial,

$$f = \sum_{\alpha \in A} c_\alpha \boldsymbol{x}^\alpha ,$$

is a finite linear combination of cain monomials $\boldsymbol{x}^\alpha$ *with integer coefficients* $c_\alpha \in \mathbb{N}$. *The set of all cain polynomials, denoted by* $\mathbb{N}[x_1, \ldots, x_p]$, *is called the* cain polynomial domain.

For instance, if $\boldsymbol{x} = \{x, y, z\}$, then both $f = xy - x - y$ and $g = xyz - xz - yz + z$ are cain polynomials in $\mathbb{N}[x, y, z]$. A cain polynomial f is special in that (i) the coefficients of f are integers; (ii) no 'constant' term appears in f; and (iii) for each monomial of f, each variable x_i appears at most once.

Now we discuss operations of cain polynomials. Scalar products by integers, and additions of cain polynomials are done in a usual way. For example, adding $f_1 = xyz - xz - yz + z$ and $f_2 = xz - x - z$, we get $g = f_1 + f_2 = xyz - yz - x$. Later we shall see that f_1, f_2 and g are isomorphic to the relations

$$x \perp\!\!\!\perp y | z \, , x \perp\!\!\!\perp z$$

and $x \perp\!\!\!\perp (y \vee z)$, respectively.

Now we define cain divisions. It will be convenient to introduce the operators, $\bar{}$, $\vee$ and $\wedge$ in $\{0,1\}^p$: for any $\alpha, \beta \in \{0,1\}^p$, let $\bar{\alpha} = \mathbf{1} - \alpha$, $e_i(\alpha \vee \beta) = \min\{1, e_i(\alpha) + e_i(\beta)\}$ and $e_i(\alpha \wedge \beta) = e_i(\alpha)e_i(\beta)$. So $\boldsymbol{x}^{\bar{\alpha}}$ is the monomial constituting the remaining variables in $\boldsymbol{x}^\alpha$, $\boldsymbol{x}^{\alpha \vee \beta}$ is the monomial by joining all variables in $\boldsymbol{x}^\alpha$ and $\boldsymbol{x}^\beta$, counting each common variable once, and $\boldsymbol{x}^{\alpha \wedge \beta}$ is the monomial consisting of the common variables in both $\boldsymbol{x}^\alpha$ and $\boldsymbol{x}^\beta$. For $f = \sum_{\alpha \in A} c_\alpha \boldsymbol{x}^\alpha$, where $c_\alpha \neq 0$, we call $\mathcal{I}(f) = \vee_{\alpha \in A} \alpha$ the *context* of f. Context is sub-additive: $\mathcal{I}(f + g) \preceq \mathcal{I}(f) \vee \mathcal{I}(g)$.

DEFINITION 2.2 (cain division). *The cain division is specified through three rules. (i)* $\boldsymbol{x}^\alpha \div \boldsymbol{x}^\beta = \boldsymbol{x}^{\alpha \wedge \bar{\beta}}$ *for any* $\alpha, \beta \in \{0,1\}^p$*; (ii) If* $\gamma \wedge \mathcal{I}(f) \wedge \mathcal{I}(g) = \mathbf{0}$, *then division distributes:*

$$(f + g) \div \boldsymbol{x}^\gamma = f \div \boldsymbol{x}^\gamma + g \div \boldsymbol{x}^\gamma ;$$

(iii) If $\beta \wedge \mathcal{I}(f) = \mathbf{0}$, *then* $f \div \boldsymbol{x}^\beta = f$.

$\boldsymbol{x}^\alpha \div \boldsymbol{x}^\beta$ is obtained by doing a 'usual' division first, and then ignoring the denominator. For example,

$$xy \div xy = 0, xy \div x = y, xy \div xz = y, xy \div z = xy .$$

(ii) says that if common variables appearing in both f and g do *not* appear in the denominator $\boldsymbol{x}^\gamma$, then division distributes over sum. If neither of the rules (i) -(iii) applies to $f \div \boldsymbol{x}^\alpha$, we say that f is *not* cain divisible by $\boldsymbol{x}^\alpha$. For instance, $f = xyz - xz - yz + z$ (a polynomial isomorphic to $x \perp\!\!\!\perp y|z$) is not cain divisible by z. Here are some of the properties on cain division.

PROPOSITION 2.1. *For any* $\alpha, \beta, \gamma \in \{0,1\}^p$ *we have (i)* $\boldsymbol{x}^{\alpha \vee \beta} \div \boldsymbol{x}^\beta = \boldsymbol{x}^{\alpha \wedge \bar{\beta}}$*; (ii) If* $\mathcal{I}(g) \wedge \alpha = \mathbf{0}$,

$$(f + g) \div \boldsymbol{x}^\alpha = f \div \boldsymbol{x}^\alpha + g$$

; (iii)

$$\boldsymbol{x}^{\alpha \vee \beta} - \boldsymbol{x}^\beta = \boldsymbol{x}^{\alpha \vee \gamma} - \boldsymbol{x}^\gamma$$

implies

$$\boldsymbol{x}^{\alpha \vee \beta} - \boldsymbol{x}^\beta = \boldsymbol{x}^{\alpha \vee (\beta \wedge \gamma)} - \boldsymbol{x}^{\beta \wedge \gamma}$$

3 The Cain Algebra: the finitery case

3.1 The cainoid

Now we briefly review the theory of cain algebra(Wang, 2007). Although the cain algebra is defined for a possibly infinite lattice, it suffices for the present purpose to consider the special case of a Boolean lattice $\mathbb{L} = 2^{\boldsymbol{x}}$, where $\boldsymbol{x} = \{x_1, \ldots, x_p\}$. We shall use the notation $\boldsymbol{x}^{\alpha}$, where $\alpha \in \{0,1\}^p$, to denote an element of $\mathbb{L}$. The symbol $\mathbb{T}^{\alpha}_{\beta}$ is called an *atom coin*, with the following conventions,

$$\mathbb{T}^{\alpha} = \mathbb{T}^{\alpha}_{\mathbf{0}}, \mathbb{T}_{\beta} = \mathbb{T}^{\mathbf{0}}_{\beta}, \mathbb{T}^{\mathbf{0}}_{\mathbf{0}} = 1 .$$

$\mathbb{T}^{\alpha}$, $\mathbb{T}_{\beta}$ and $\mathbb{T}^{\alpha}_{\beta}$ are called a *raising coin*, a *lowering coin* and a *mixed coin*, respectively. A *coin* is a finite concatenation $\mathbb{T} = \mathbb{T}^{\alpha_1}_{\beta_1} \ldots \mathbb{T}^{\alpha_n}_{\beta_n}$ of n atom coins, where n is any positive integer. Let $\mathfrak{C}$ be the set of all coins.

The atom coins $\mathbb{T}^{\alpha}$, $\mathbb{T}_{\alpha}$ and $\mathbb{T}^{\alpha}_{\beta}$ are cain algebraic counterparts of the classical joint probability density function $f(\boldsymbol{x}^{\alpha})$, the reciprocal $1/f(\boldsymbol{x}^{\alpha})$, and the conditional density function $f(\boldsymbol{x}^{\alpha}|\boldsymbol{x}^{\beta})$, respectively. A coin $\mathbb{T}^{\alpha_1}_{\beta_1} \ldots \mathbb{T}^{\alpha_n}_{\beta_n}$ is a cain algebraic expression of the likelihood function, $f(\boldsymbol{x}^{\alpha_1}|\boldsymbol{x}^{\beta_1}) \times \cdots \times f(\boldsymbol{x}^{\alpha_n}|\boldsymbol{x}^{\beta_n})$.

DEFINITION 3.1 (Cainoid). *Define the dot product of*

$$\mathbb{T} = \mathbb{T}^{\alpha_1}_{\beta_1} \ldots \mathbb{T}^{\alpha_m}_{\beta_m}$$

and

$$\mathbb{T}' = \mathbb{T}^{\alpha'_1}_{\beta'_1} \ldots \mathbb{T}^{\alpha'_n}_{\beta'_n}$$

by

$$\mathbb{T} \cdot \mathbb{T}' = \mathbb{T}^{\alpha_1}_{\beta_1} \ldots \mathbb{T}^{\alpha_m}_{\beta_m} \cdot \mathbb{T}^{\alpha'_1}_{\beta'_1} \ldots \mathbb{T}^{\alpha'_n}_{\beta'_n} .$$

The algebraic structure $(\mathfrak{C}, \cdot)$ *is called a* cainoid *if for any* $\mathbb{T}, \mathbb{T}', \mathbb{T}'' \in \mathfrak{C}$ *and* $\alpha, \beta \in \{0,1\}^p$, *the following hold: C1:* $\mathbb{T} \cdot \mathbb{T}' = \mathbb{T}' \cdot \mathbb{T}$*; C2:* $(\mathbb{T} \cdot \mathbb{T}') \cdot \mathbb{T}'' = \mathbb{T} \cdot (\mathbb{T}' \cdot \mathbb{T}'')$*; C3:* $1 \cdot \mathbb{T} = \mathbb{T}$*; C4:* $\mathbb{T}^{\alpha} \cdot \mathbb{T}_{\alpha} = 1$*; C5:* $\mathbb{T}^{\alpha}_{\beta} = \mathbb{T}^{\alpha\vee\beta} \cdot \mathbb{T}_{\beta}$*, where* $\alpha \succ \mathbf{0}$.

C5 is motivated by the definition of conditional density function, $f(\boldsymbol{x}^{\alpha}|\boldsymbol{x}^{\beta}) = f(\boldsymbol{x}^{\alpha}, \boldsymbol{x}^{\beta})/f(\boldsymbol{x}^{\beta})$. As immediate consequences of C1-C5 we have the *raising-up law*: If $\alpha \succ \mathbf{0}$ then

$$\mathbb{T}^{\alpha\vee\beta}_{\gamma} = \mathbb{T}^{\alpha}_{\beta\vee\gamma}\mathbb{T}^{\beta} \Leftrightarrow \mathbb{T}^{\beta\vee\gamma} = \mathbb{T}^{\beta}\mathbb{T}^{\gamma} ;$$

and the *lowering-down law*: If $\alpha \succ \mathbf{0}$ then

$$\mathbb{T}^{\alpha}_{\beta\vee\gamma} = \mathbb{T}^{\alpha\vee\beta}_{\gamma}\mathbb{T}_{\beta} \Leftrightarrow \mathbb{T}^{\beta\vee\gamma} = \mathbb{T}^{\beta}\mathbb{T}^{\gamma} .$$

The cain. Two raising coins $\mathbb{T}^{\alpha}$ and $\mathbb{T}^{\beta}$ ($\alpha, \beta \succ \mathbf{0}$), are said *mutually prime*, if $\alpha \neq \beta$. Let $\mathbb{T} \neq 1$. If $\mathbb{T} = \mathbb{T}^{\alpha_1}_{\beta_1} \ldots \mathbb{T}^{\alpha_r}_{\beta_r}$, where $\mathbb{T}^{\alpha_j}_{\beta_j} \neq 1$, then we say that $\mathbb{T}$ has an expression with length r. There are infinitely many expressions which are equivalent to one another. It can shown that there exists nonzero integers n_i and mutually prime coins $\mathbb{T}^{\alpha_i}, i = 1, \ldots, r$, so that $\mathbb{T}$ can be expressed as

$$\mathbb{T} = (\mathbb{T}^{\alpha_1})^{n_1} \ldots (\mathbb{T}^{\alpha_r})^{n_r} ,$$

A raising coin $\mathbb{T}^{\alpha}$ is called a *prime coin* if there does not exist an expression

$$\mathbb{T}^{\alpha} = (\mathbb{T}^{\alpha_1})^{n_1} \ldots (\mathbb{T}^{\alpha_r})^{n_r}$$

so that each $\alpha_i \prec \alpha$ for each $i = 1, \ldots, r$.

THEOREM 3.1 (Wang, 2007). *For every coin $\mathbb{T} \in \mathfrak{C}$, there exist nonzero integers $n_1, \ldots, n_r$ so that $\mathbb{T}$ has a unique expression*

$$\mathbb{T} = (\mathbb{T}^{\alpha_1})^{n_1} \ldots (\mathbb{T}^{\alpha_r})^{n_r} \tag{1}$$

where (i) $\mathbb{T}^{\alpha_1}, \ldots, \mathbb{T}^{\alpha_r}$ are prime; and (ii) $\mathbb{T}^{\alpha_1}, \ldots, \mathbb{T}^{\alpha_r}$ are mutually prime.

(1) is called the canonical expression *of $\mathbb{T}$, r the order of $\mathbb{T}$, written as $|\mathbb{T}| = r$, and $\alpha = \vee_{i=1}^{r}\alpha_i$ the* context *of $\mathbb{T}$, written as $\mathfrak{I}(\mathbb{T}) = \alpha$.*

Notations: Let $\mathbb{T}\{\alpha\}$ denote an arbitrary coin with context $\mathfrak{I}(\mathbb{T}\{\alpha\}) = \alpha$, and $\mathbb{T}[\alpha]$ denote an arbitrary coin with $\mathfrak{I}(\mathbb{T}[\alpha]) \preceq \alpha$.

DEFINITION 3.2 (cain). *The* coin integration *is an unary operation on $\mathbb{T} \in \mathfrak{C}$ satisfying C6:*

$$\int \mathbb{T}^{\alpha}\, d\beta = \mathbb{T}^{\alpha\wedge\bar{\beta}}$$

holds for any $\alpha, \beta \in \{0, 1\}^p$; C7:

$$\int (\mathbb{T}\{\alpha_1\}\mathbb{T}\{\alpha_2\})\ d(\beta_1 \vee \beta_2) = \int \mathbb{T}\{\alpha_1\} d\beta_1$$

$\int \mathbb{T}\{\alpha_2\} d\beta_2$ *holds if $\beta_1 \wedge \beta_2 = \mathbf{0}$, and $\beta_1 \wedge \alpha_2 = \beta_2 \wedge \alpha_1 = \mathbf{0}$; C8:*

$$\int \mathbb{T} d\mathbf{0} = \mathbb{T}$$

holds for any $\mathbb{T} \in \mathfrak{C}$.

A cain *is a cainoid $\mathfrak{C}$ furnished with C6-C8.*

C6 is analogous to the definition of marginal probability density functions. C7 is an analogue of a well-known property of the Riemann integral, namely,

$$\int f(x, z)g(y, z)\, dxdy = \int f(x, z)\, dx \int g(y, z)\, dy\,.$$

The following basic properties are special cases of the corresponding properties for a general cain studied in Wang (2007).

THEOREM 3.2. *(i) If $\alpha \wedge \beta = 0$, then*

$$\int \Pi[\beta]\Pi \, d\alpha = \Pi[\beta] \int \Pi \, d\alpha$$

(ii) For any $\alpha, \beta, \gamma \in \{0,1\}^p$,

$$\Pi^\alpha_\beta = \Pi^\alpha_\gamma \;\Rightarrow\; \Pi^\alpha_\beta = \Pi^\alpha_{\beta\wedge\gamma}$$

(iii) If $\beta \wedge \gamma = 0$ then

$$\int \Pi^\alpha_\beta \, d\gamma = \Pi^{\alpha\wedge\bar{\gamma}}_\beta$$

(iv) If $\alpha \succ 0$ then

$$\int \Pi^\alpha_\beta \, d(\alpha \wedge \bar{\beta}) = 1$$

holds for any $\beta$.

(iv) is an analogy of the fact that the conditional probability density functions are normalized functions.

4 Probabilistic Conditional Independence

DEFINITION 4.1 (conditional independence). *If $\Pi^\alpha_{\beta\vee\gamma} = \Pi^\alpha_\gamma$, then $\alpha$ is said independent of $\beta$ conditional on $\gamma$, written $\alpha \perp\!\!\!\perp \beta|\gamma$. When $\alpha \perp\!\!\!\perp \beta|0$ holds, we write $\alpha \perp\!\!\!\perp \beta$ and say that $\alpha$ is independent of $\beta$.*

If $\alpha, \beta, \gamma \succ 0$ then $\Pi^\alpha_{\beta\vee\gamma} = \Pi^\alpha_\gamma$ is equivalent to either

$$\Pi^{\alpha\vee\beta}_\gamma = \Pi^\alpha_\gamma \Pi^\beta_\gamma$$

or

$$\Pi^{\alpha\vee\beta\vee\gamma} = \Pi^{\beta\vee\gamma}\Pi^\alpha_\gamma \,.$$

Note that $\alpha \perp\!\!\!\perp \beta|\gamma$ is symmetric for $\alpha$ and $\beta$, i.e., $\alpha \perp\!\!\!\perp \beta|\gamma \Rightarrow \beta \perp\!\!\!\perp \alpha|\gamma$ in all cases except when $\beta = 0, \alpha \succ 0$ and $\alpha \not\preceq \gamma$. If α, β, γ are nontrivial and mutually exclusive, then $\alpha \perp\!\!\!\perp \beta|\gamma$ holds if and only if $\Pi^{\alpha\vee\beta\vee\gamma} = \Pi[\bar{\alpha}]\Pi[\bar{\beta}]$ holds. This seemingly weaker condition is often convenient. It can be further shown that the ternary relation $\cdot \perp\!\!\!\perp \cdot|\cdot$ on $\{0,1\}^p$ satisfies the axioms of a graphoid of Pearl and Paz (1987). That is, for all nontrivial and mutually exclusive elements $\alpha, \beta, \gamma, \xi \in \{0,1\}^p$ we have

G1 :	$\alpha \perp\!\!\!\perp \beta|\gamma \;\Rightarrow\; \beta \perp\!\!\!\perp \alpha|\gamma$	(symmetry)
G2 :	$\alpha \perp\!\!\!\perp (\beta \vee \xi)|\gamma \;\Rightarrow\; \alpha \,\mathrm{II}\, \beta|\gamma$	(decomposition)
G3 :	$\alpha \perp\!\!\!\perp (\beta \vee \gamma)|\xi \;\Rightarrow\; \alpha \,\mathrm{II}\, \beta|(\gamma \vee \xi)$	(weak union)
G4 :	$\alpha \perp\!\!\!\perp \beta|(\gamma \vee \xi),\ \alpha \perp\!\!\!\perp \gamma|\xi \;\Rightarrow\; \alpha \,\mathrm{II}\, (\beta \vee \gamma) \mid \xi$	(contraction)
G5 :	$\alpha \perp\!\!\!\perp \beta|(\gamma \vee \xi),\ \alpha \perp\!\!\!\perp \gamma|(\beta \vee \xi) \;\Rightarrow\; \alpha \perp\!\!\!\perp (\beta \vee \gamma) \mid \xi$	(intersection)

Properties G1-G5 were also discussed by Dawid (1979) and Spohn (1980). $\cdot \perp\!\!\!\perp \cdot \mid \cdot$ is consistent with a stronger system known as the separoid(Dawid, 2001), which includes several axiomatic systems, such as the orthogonoid and the graphoid, relevant for formal reasoning using the concept of *irrelevance* of information.

5 Isomorphism Between the Polynomial Domain and the Cain Algebra

5.1 Cain polynomials of coins

In this section we make a connection between the two algebraic systems of the cain polynomial domain $\mathbb{N}[x_1, \ldots, x_p]$ and the cain algebra $\mathfrak{C}$. A map $\phi : \mathfrak{C} \to \mathbb{N}[x_1, \ldots, x_p]$ is called a *cain homomorphism*, if for any $\alpha, \beta \in \{0,1\}^p$, $\mathbb{T}, \mathbb{T}' \in \mathfrak{C}$ the following properties hold, (H1): $\phi(1) = 0$; (H2): $\phi(\mathbb{T}\mathbb{T}') = \phi(\mathbb{T}) + \phi(\mathbb{T}')$; and (H3): $\phi\left(\int \mathbb{T}^\alpha \, d\beta\right) = \phi(\mathbb{T}^\alpha) \div \boldsymbol{x}^\beta$.

DEFINITION 5.1 (cain polynomial of a coin). *Let* $\mathrm{Po} : \mathfrak{C} \to \mathbb{N}[x_1, \ldots, x_p]$ *be defined as follows. If* $\mathbb{T}$ *has an expression,* $\mathbb{T} = (\mathbb{T}^{\alpha_1})^{n_1} \ldots (\mathbb{T}^{\alpha_r})^{n_r}$*, then let*

$$\mathrm{Po}(\mathbb{T}) = n_1 \boldsymbol{x}^{\alpha_1} + \cdots + n_r \boldsymbol{x}^{\alpha_r} \tag{2}$$

and $\mathrm{Po}(1) = 0$ *if and only if* $\mathbb{T} = 1$. $\mathrm{Po}(\mathbb{T})$ *is called a* cain polynomial *of* $\mathbb{T}$.

Since a given coin can have infinitely many expressions, the corresponding cain polynomials (2) vary accordingly. The following theorem says that these polynomials are equal to one another.

THEOREM 5.1. *Equivalent coins induce the same cain polynomial, namely,*

$$\mathbb{T} \simeq \mathbb{T}' \Rightarrow \mathrm{Po}(\mathbb{T}) = \mathrm{Po}(\mathbb{T}') \tag{3}$$

where $\mathrm{Po}(\mathbb{T})$ *and* $\mathrm{Po}(\mathbb{T}')$ *are any cain polynomials of* $\mathbb{T}$ *and* $\mathbb{T}'$*, respectively.*

It follows then $\mathrm{Po}(\cdot)$ satisfies (H1)-(H3) and is a cain homomorphism. Let

$$\ker(\phi) = \left\{(\mathbb{T}, \mathbb{T}') \in \mathfrak{C}^2 : \phi(\mathbb{T}) = \phi(\mathbb{T}')\right\}$$

be the *kernel* of a homomorphism $\phi$, then $\ker(\phi)$ is a *congruence* on $\mathfrak{C}$. That is, $\ker(\phi)$ is an equivalence relation on $\mathfrak{C}$, and, in addition, is compatible with the coin product. If $\phi$ is a cain homomorphism, then $\phi(\mathbb{T}^{-1}) = -\phi(\mathbb{T})$. Hence, $(\mathbb{T}, \mathbb{T}') \in \ker(\phi)$ if and only if $\phi(\mathbb{T}(\mathbb{T}')^{-1}) = 0$. Thus, $(\mathbb{T}, \mathbb{T}') \in \ker(\mathrm{Po})$ if and only if $\mathbb{T} \simeq \mathbb{T}'$. These results lead to the following important Theorem.

THEOREM 5.2 (Isomorphism). $\mathrm{Po}(\cdot)$ *is an isomorphism from* $\mathfrak{C}/\simeq$*, the quotient of the cain* $\mathfrak{C}$ *with respect to* $\simeq$*, to the cain polynomial domain* $\mathbb{N}[x_1, \ldots, x_p]$.

5.2 Polynomial representation of conditional independence

We have seen that a cain $\mathfrak{C}$ is equal to the polynomial domain $\mathbb{N}[x_1, \ldots, x_p]$ up to an isomorphism. Addition of two cain polynomials corresponds to the coin product, while cain division corresponds to the coin integration.

Now we turn to consider the conditional independence relations. We have seen that $\alpha \perp\!\!\!\perp \beta|\gamma$ holds if and only if $\mathbb{T}^{\alpha}_{\beta\vee\gamma} = \mathbb{T}^{\alpha}_{\gamma}$, or $\mathbb{T}^{\alpha\vee\beta}_{\gamma} = \mathbb{T}^{\alpha}_{\gamma}\mathbb{T}^{\beta}_{\gamma}$, or $\mathbb{T}^{\alpha\vee\beta\vee\gamma} = \mathbb{T}^{\beta\vee\gamma}\mathbb{T}^{\alpha}_{\gamma}$ holds. In fact there are infinitely many such equivalent forms obtained by multiplying an arbitrary coin to both sides of such an identity. All these identities however may be rewritten in the form $\mathbb{T} = 1$. If $\mathbb{T}' = 1$ is an equivalent identity to $\mathbb{T} = 1$, then $\mathbb{T}' \simeq \mathbb{T}$ necessarily. Recall that the canonical expression of equivalent coins is unique. For example, the unique canonical expression of all equivalent coin identities for the relation $\alpha \perp\!\!\!\perp \beta|\gamma$ is given by

$$\mathbb{T}^{\alpha\vee\beta\vee\gamma}(\mathbb{T}^{\alpha\vee\gamma})^{-1}(\mathbb{T}^{\beta\vee\gamma})^{-1}\mathbb{T}^{\gamma} = 1\,.$$

DEFINITION 5.2 (CI polynomial). *For any* $\alpha, \beta, \gamma \in \{0,1\}^p$, *we call*

$$f(\alpha, \beta|\gamma) = \boldsymbol{x}^{\alpha\vee\beta\vee\gamma} - \boldsymbol{x}^{\alpha\vee\gamma} - \boldsymbol{x}^{\beta\vee\gamma} + \boldsymbol{x}^{\gamma} \tag{4}$$

a conditional independence (CI) polynomial *of* $(\alpha, \beta; \gamma)$*; and*

$$f(\alpha, \beta|\gamma) = 0$$

a CI equation *of* $(\alpha, \beta; \gamma)$.

THEOREM 5.3. *For any* $\alpha, \beta, \gamma \in \{0,1\}^p$, $\alpha \perp\!\!\!\perp \beta|\gamma \Leftrightarrow f(\alpha, \beta|\gamma) = 0$.

Theorem 5.3 allows us to deduce conditional independence relations by deriving CI equations in the cain polynomial domain. The tools for deriving CI equations are supplied by the Theorem 5.2.

6 The Basic Problem and a Partial Solution

The basic problem: Given a set of cain equations $f_1 = \cdots = f_r = 0$, is it true that $g = 0$? Here $g, f_1, \ldots, f_r$ are cain polynomials given a priori. The concept of *marginals* of a cain polynomial will play an essential role for solving the basic problem. To motivate the definition of the marginals, suppose that $x \perp\!\!\!\perp (y, z)$ holds. Then we obtain the CI equation $f = xyz - x - yz = 0$, or $xyz = yz - f_1$, with $f_1 = -x$. Although, $f = 0$ is equivalent to $xyz = yz - f_1$, there is an important difference between the two forms. While the elements x, y, z each appears twice in $f$, each of these elements appers exactly *once* in both sides of $xyz = yz - f_1$. Dividing both sides of $xyz = yz - f_1$ by $y$ gives the equation $xyz \div y = (x + yz) \div y \Leftrightarrow xz - x - z = 0$, the CI equation of $x \perp\!\!\!\perp z$. Note that this equation is not obtainable by directly dividing $f$.

DEFINITION 6.1 (marginal cain polynomial). *Suppose that cain polynomial f can be decomposed as $f = c_\alpha \boldsymbol{x}^\alpha + c_\beta \boldsymbol{x}^\beta + f_1$, with properties (i) $c_\alpha c_\beta = -1$, $|c_\alpha| = |c_\beta| = 1$, and (ii) $\alpha \wedge \beta \succ 0$, and (iii) there exists $0 \prec \gamma \preceq \alpha \wedge \beta$ so that $\gamma \wedge \mathfrak{I}(f_1) = 0$. Then we call*

$$M(f; \boldsymbol{x}^\gamma) = c_\alpha \boldsymbol{x}^{\alpha \wedge \bar{\gamma}} + c_\beta \boldsymbol{x}^{\beta \wedge \bar{\gamma}} + f_1 \tag{5}$$

a marginal (cain polynomial) *of f with respect to $\boldsymbol{x}^\gamma$.*

Since $\gamma \wedge \mathfrak{I}(f_1) = 0$, we have $f_1 \div \boldsymbol{x}^\gamma = f_1$. (5) can then be alternatively written as

$$M(f; \boldsymbol{x}^\gamma) = c_\alpha \boldsymbol{x}^\alpha \div \boldsymbol{x}^\gamma + c_\beta \boldsymbol{x}^\beta \div \boldsymbol{x}^\gamma + f_1$$

The next theorem says that a 'rich' CI polynomial always has many marginals.

THEOREM 6.1. *Suppose that α, β, γ are mutually exclusive, and $\alpha \succ 0$. Then*

$$f(\alpha \wedge \bar{\delta}, \beta | \gamma) = \boldsymbol{x}^{(\alpha \wedge \bar{\delta}) \vee \beta \vee \gamma} - \boldsymbol{x}^{(\alpha \wedge \bar{\delta}) \vee \gamma} - \boldsymbol{x}^{\beta \vee \gamma} + \boldsymbol{x}^\gamma \tag{6}$$

is a marginal of the CI polynomial

$$f(\alpha, \beta | \gamma) = \boldsymbol{x}^{\alpha \vee \beta \vee \gamma} - \boldsymbol{x}^{\alpha \vee \gamma} - \boldsymbol{x}^{\beta \vee \gamma} + \boldsymbol{x}^\gamma$$

for any $\delta \preceq \alpha$. We call $f(\alpha \wedge \bar{\delta}, \beta | \gamma) = M(f(\alpha, \beta | \gamma); \boldsymbol{x}^\delta)$ a marginal CI polynomial *of $f(\alpha, \beta | \gamma)$.*

THEOREM 6.2. *If q is a marginal of f, then $f = 0$ implies $q = 0$.*

Theorem 6.2 gives a partial answer to the basic problem: given $f = 0$, one can show $g = 0$ by showing that g is a marginal of f. The condition of being a marginal is however much stronger than necessary as illustrated in the following example.

EXAMPLE 6.1. *Let $f = xyz - x - yz$ be the CI polynomial of $x \perp\!\!\!\perp (y, z)$ again. Let $g = xyz - yz - xz + z$ be the CI polynomial of $x \perp\!\!\!\perp y | z$. Note that $M(f; y) = xz - x - z$ and $g = f - M(f; y)$. If $f = 0$ then $M(f; y) = 0$ by Theorem 6.2. The equations $f = 0$ and $M(f; y) = 0$ in turn jointly imply $g = f - M(f; y) = 0$, proving that $x \perp\!\!\!\perp (y, z) \Rightarrow x \perp\!\!\!\perp y | z$.*

Generalizing Example 6.1 we have

THEOREM 6.3. *Let f be a cain polynomial and $q_1, \ldots, q_s$ be marginals of f. If $g = c_0 f + \sum_{j=1}^{s} c_j q_j$, where $c_0, c_1, \ldots, c_s \in \mathbb{N}$, then $f = 0$ implies $g = 0$.*

The conditions in Theorem 6.3 to ensure $g = 0$ are again only sufficient.

EXAMPLE 6.2. *Let $f = xyzw - yzw - x$ be the CI polynomial of $x \perp\!\!\!\perp (y, z, w)$. Let $g = xzw - zw - xw + w$ be the CI polynomial of $x \perp\!\!\!\perp z | w$. It can be verified that $g = M(f - M(f; yz); y)$. That is, g is a marginal of the difference between f and a marginal of f. Thus, $f = 0 \Rightarrow g = 0$, or $x \perp\!\!\!\perp (y, z, w) \Rightarrow x \perp\!\!\!\perp z | w$.*

Alternatively, we write g as

$$g = M(f; y) - M(M(f; y); z),$$

the difference between a marginal of f and a marginal of a marginal of f. If $f = 0$, then $M(f; y) = 0$, which in turn implies that $M(M(f; y); z) = 0$, proving $g = 0$.

DEFINITION 6.2 (higher-order marginals). *A marginal $M(f;\boldsymbol{x}^{\alpha})$ of $f$ is called a first-order marginal of $f$, a marginal $M(M(f;\boldsymbol{x}^{\alpha});\boldsymbol{x}^{\beta})$ of $M(f;\boldsymbol{x}^{\alpha})$ is called a second-order marginal of $f$; and so on.*

Now we generalize the ideas expressed in Example 6.2, If $F$ is a finite set of cain polynomials, and $g$ is a $\mathbb{N}$-linear combination of polynomials in $F$, then we say that $g$ is *linearly expressible* by $F$, and write $g = \bar{F}$.

THEOREM 6.4. *Let $f$ be a nonzero cain polynomial. Let F_1 be a finite set containing $f$ and some higher-order marginals of $f$. Let F_2 contain F_1 and some higher-order marginals of some $\mathbb{N}$-linear combinations of polynomials in F_1. Let F_3 contain F_2 and some higher-order marginals of some $\mathbb{N}$-linear combinations of polynomials in F_2, and so on to obtain F_s for some $s \geq 1$.*

If $g = \bar{F}_s$ then $f = 0 \Rightarrow g = 0$.

7 Discussions

The methods discussed in Section 6 only give sufficient conditions for the basic problem when $r = 1$. Both sufficient and necessary conditions for general $r$ need further study.

References

DAWID, A. P. (1979). Conditional independence in statistical theory (with discussion), *J. Roy. Statist. Soc. Ser. B* **41** 1–31.

DAWID, A. P. (1988). Conditional independence. In S. Kotz, C. B. Read and D. L. Banks, (Eds.): *Encyclopedia of Statistical Science (Update Volume 2)*, Wiley Interscience, New York, 146-155.

DAWID, A. P. (2001). Separoids: a mathematical framework for conditional independence and irrelevance, *Ann. Math. Art. Intell.* **32** 335–372.

GARCIA, L. D., STILLMAN, M. AND STURMFELS (2005): Algebraic geometry of Bayesian networks, *Journal of Symbolic Computation* **39**, 331–355.

MILLER, E. AND STURMFELS, B. (2005). *Combinatorial Commutative Algebra*, Springer, New York.

PEARL, J. (2000). *Causality–models, reasoning and inference*, Cambridge University Press, New York.

PEARL, J. AND PAZ, A. (1987). Graphoids: a graph-based logic for reasoning about relevance relations, In D. Hogg and L. Steels (Eds.): *Advances in Artificial Intelligence*, North-Holland, Amsterdam, 357–363.

SPIRTES, P., GLYMOUR, C., AND SCHEINES, R. (2000). *Causation, Prediction, and Search.* 2nd Edition, MIT Press,

STUDENÝ, M (2005). *Probabilistic Conditional Independence Structures*, London, Springer-Verlag.

WANG, J. (2007). Wang (2007). The cain algebra, the semigraphoid and the separoid, *RIMS Kokyuroku*, Vol. 1560, Statistical Decision for Multiple Comparison and Its Related Topics, 145-154.

WANG, J. (2007). The cain algebra: a universal algebraic approach for probabilistic conditional independence, *Submitted.*