数理解析研究所講究録
第1560巻 2007年 145-154

The cain algebra, the semigraphoid and the separoid

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 cain algebra was proposed recently by Wang (2006) and Wang (2007) to enable study of properties of probability density functions, especially those relevant for probabilistic conditional independence, in a purely universal algebraic fashion. In the cain algebra the familiar likelihood functions are replaced by the corresponding algebraic objects called the coins. A cain is an Abelian group with the coins as its elements equipped with a dot product. In the cain algebra, properties such as conditional independence relations are represented by equations using appropriate coins.

One major advantage of the universal algebraic approach is that properties concerning probability density functions in general and conditional independence in particular can be derived automatically by transforming one coin equation to another coin equation by appealing to the cain axioms, which themselves are in equational from. Thus the cain algebra is in contrast with other well-known axiomatic systems for conditional independence, such as the graphoid of Pearl and Paz (1987) and the separoid of Dawid (2001). The latter systems are built on some principal properties of conditional independence regarded useful for general probabilistic reasoning, none of the axioms in these systems being expressed in equations. A possible limitation of these latter approaches is that not all properties on probabilistic conditional independence can be derived from these finite set of axioms (Studený, 1992).

The purpose of this paper is to briefly introduce the cain algebra (Wang 2006, 2007) and show that this algebraic system is consistent with the graphoid and the separoid.

2 The Cain Algebra

2.1 The Cainoid

Let $(\mathbb{L}, \leq)$ be a lattice, where $\mathbb{L}$ is a nonempty set and $\leq$ is a partial order in $\mathbb{L}$. It is further supposed that $(\mathbb{L}, \leq)$ is bounded below, that is, there exists a bottom $\emptyset$ with $\emptyset \leq x$ for any $x \in \mathbb{L}$. An element $x \in \mathbb{L}$ is said nontrivial if $x > \emptyset$, and a set of elements are said mutually exclusive if their pairwise meets are trivial. The direct product $\mathbb{L} \otimes \mathbb{L} = \{(x, y) \mid x, y \in \mathbb{L}\}$ plays an important role in the theory. Note that $(x, y) \neq (y, x)$ if $x \neq y$. To emphasize this asymmetry, we replace (x, y) by using a new symbol π_y^x (reads as *coin-x-over-y*) and making the following conventions $\pi^x = \pi_\emptyset^x$, $\pi_y = \pi_y^\emptyset$, $\pi_\emptyset^\emptyset = 1$.

We call π^x the *raising coin* with *context* x, π_y the *lowering coin* with context y, and π_y^x the *mixed coin* with raising context x and lowering context y. All these coins are called *atom coins*.

A *coin* is a concatenation $\mathbb{T} = \mathbb{T}_{y_1}^{x_1} \cdots \mathbb{T}_{y_n}^{x_n}$ of n atom coins, with some atom coins (adjacent or not) being possibly identical. Now denote by

$$\mathfrak{C} = \left\{ \mathbb{T}_{y_1}^{x_1} \cdots \mathbb{T}_{y_n}^{x_n} \;\middle|\; \mathbb{T}_{y_1}^{x_1}, \cdots, \mathbb{T}_{y_n}^{x_n} \in \mathbb{L} \otimes \mathbb{L}, n \in \mathbb{N} \right\} \tag{1}$$

the set of all coins. Since n can be any natural number, $\mathfrak{C}$ is an infinite set even if $\mathbb{L}$ is finite. Note that $\mathbb{L} \otimes \mathbb{L} = \{\mathbb{T}_y^x \mid x, y \in \mathbb{L}\}$ is a proper subset of $\mathfrak{C}$. In the set of coins $\mathfrak{C}$, we introduce a binary dot operator, $\cdot : \mathfrak{C} \times \mathfrak{C} \to \mathfrak{C}$, so that for any $\mathbb{T} = \mathbb{T}_{y_1}^{x_1} \cdots \mathbb{T}_{y_m}^{x_m}$ and $\mathbb{T}' = \mathbb{T}_{v_1}^{u_1} \cdots \mathbb{T}_{v_n}^{u_n}$ we have

$$\mathbb{T} \cdot \mathbb{T}' = \mathbb{T}_{y_1}^{x_1} \cdots \mathbb{T}_{y_m}^{x_m} \cdot \mathbb{T}_{v_1}^{u_1} \cdots \mathbb{T}_{v_n}^{u_n} \tag{2}$$

DEFINITION 2.1 (CAINOID). *Let $(\mathbb{L}, \leq)$ be a lattice with bottom $\emptyset$. Let $\mathfrak{C}$ of (1) be the set of coins, and $\cdot$ be defined by (2). An algebraic structure $(\mathfrak{C}, \cdot)$ is called a* cainoid *if for any $\mathbb{T}, \mathbb{T}', \mathbb{T}'' \in \mathfrak{C}$ and $x, y \in \mathbb{L}$, 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: $\mathbb{T}^x \cdot \mathbb{T}_x = 1$

C5: $\mathbb{T}_y^x = \mathbb{T}^{x \vee y} \cdot \mathbb{T}_y \qquad (\textit{where } x > \emptyset)$

From the definition it is not hard to see that a cainoid $(\mathfrak{C}, \cdot)$ forms an Abelian group. The coins $\mathbb{T}^x$, $\mathbb{T}_x$ and $\mathbb{T}_y^x$ may be regarded as algebraic abstractions of the joint probability density function $f(x)$, the reciprocal $1/f(x)$, and the conditional density function $f(x|y)$, respectively. A coin $\mathbb{T}_{y_1}^{x_1} \cdots \mathbb{T}_{y_n}^{x_n}$ may then be regarded as the abstraction of the product $f(x_1|y_1) \cdots f(x_n|y_n)$ of n conditional probability density functions. Axiom C5 may be regarded as the abstraction of the definition of conditional density function, $f(x|y) = f(x,y)/f(y)$. With these interpretations, the identity $\mathbb{T}_y^x = \mathbb{T}_x^y \mathbb{T}^x \mathbb{T}_y$ $(x > \emptyset, y > \emptyset)$ may be regarded as an anologue of the familiar Bayes' theorem.

As a direct consequence of the axioms, we have the following important result. In a cainoid the embedding partial order $\leq$ can be represented by the identity: $x \leq y \Leftrightarrow \mathbb{T}_y^x = 1$ $(x > \emptyset)$. In the cain algebra most of the problems will be reduced to the problem of solving or transforming some coin equations to other coin equations. For this purpose two following properties are of basic importance.

PROPOSITION 2.1 (RAISING-UP LAW). *For any $x > \emptyset, y, z$ of $\mathbb{L}$ we have*

$$\mathbb{T}_z^{x \vee y} = \mathbb{T}_{y \vee z}^x \mathbb{T}^y \iff \mathbb{T}^{y \vee z} = \mathbb{T}^y \mathbb{T}^z \qquad (x > \emptyset) \tag{3}$$

The coin $\mathbb{T}_z^{x \vee y}$ on the l.h.s. of (3) is obtained by *detaching* the context y from $y \vee z$ of $\mathbb{T}_{y \vee z}^x$ and *raising* y using $\mathbb{T}^y$. To ensure the validity of the detachment, we need the 'independence condition', $\mathbb{T}^{y \vee z} = \mathbb{T}^y \mathbb{T}^z$.

PROPOSITION 2.2 (LOWERING-DOWN LAW). *For any $x > \emptyset, y, z$ of $\mathbb{L}$ we have*

$$\mathbb{T}_{y \vee z}^x = \mathbb{T}_z^{x \vee y} \mathbb{T}_y \iff \mathbb{T}^{y \vee z} = \mathbb{T}^y \mathbb{T}^z \qquad (x > \emptyset) \tag{4}$$

The coin $\mathbb{T}^x_{y\vee z}$ on the l.h.s. of (4) is obtained by *lowering* y in $\mathbb{T}^{x\vee y}_z$ using $\mathbb{T}_y$, and *joining* y with z. To ensure the validity of joining y with z, again we need the 'independence condition', $\mathbb{T}^{y\vee z} = \mathbb{T}^y\mathbb{T}^z$. The following fact is also helpful for transforming coin equations.

PROPOSITION 2.3. *Let $x\vee z, x\vee w, y\vee z, y\vee w$ be nontrivial. Then*

$$\mathbb{T}^{y\vee z}_x = \mathbb{T}^{y\vee w}_x\mathbb{T} \quad\cong\quad \mathbb{T}^{x\vee z}_y = \mathbb{T}^{x\vee w}_y\mathbb{T} \tag{5}$$

where $\mathbb{T}\in\mathfrak{C}$ is an arbitrary coin.

2.2 Canonical expressions

To introduce further structures in a cainoid we now study canonical expressions of coins. Let $\mathbb{T}\neq 1$ be an arbitrary coin. A coin identity of the form $\mathbb{T} = \mathbb{T}^{x_1}_{y_1}\cdots\mathbb{T}^{x_r}_{y_r}$ is called an expression of $\mathbb{T}$ with length r, where $\mathbb{T}^{x_1}_{y_1},\cdots,\mathbb{T}^{x_r}_{y_r}$ are atom coins other than unity. There are infinitely many expressions which are equivalent to one another. Two raising coins $\mathbb{T}^x$ and $\mathbb{T}^y$, with $x>\emptyset$ and $y>\emptyset$, are said *mutually prime*, if $x\neq y$.

THEOREM 2.1. *For any arbitrary coin $\mathbb{T}\neq 1$, there exists nonzero integers n_i and mutually prime coins $\mathbb{T}^{x_i}, i=1,\cdots,r$ so that $\mathbb{T}$ can be expressed as*

$$\mathbb{T} = (\mathbb{T}^{x_1})^{n_1}\cdots(\mathbb{T}^{x_r})^{n_r} \tag{6}$$

DEFINITION 2.2 (PRIME COIN). *A raising coin $\mathbb{T}^x$ is called a* prime coin *if there does not exist an expression*

$$\mathbb{T}^x = (\mathbb{T}^{x_1})^{n_1}\cdots(\mathbb{T}^{x_r})^{n_r}$$

so that each $x_i<x$ for each $i=1,\cdots,r$.

If $\mathbb{L}$ satisfies the descending chain condition (DCC) then we have a stronger result than theorem 2.1.

THEOREM 2.2. *Suppose that $\mathbb{L}$ satisfies the DCC. Then there exist nonzero integers $n_1,\cdots,n_r$ so that every coin $\mathbb{T}\in\mathfrak{C}$ has a unique expression*

$$\mathbb{T} = (\mathbb{T}^{x_1})^{n_1}\cdots(\mathbb{T}^{x_r})^{n_r} \tag{7}$$

where (i) $\mathbb{T}^{x_1},\cdots,\mathbb{T}^{x_r}$ are prime; and (ii) $\mathbb{T}^{x_1},\cdots,\mathbb{T}^{x_r}$ are mutually prime.

The unique expression given by (7) is called the *canonical expression* of $\mathbb{T}$. r is called the order of $\mathbb{T}$, written as $|\mathbb{T}| = r$; and $x = \vee_{i=1}^r x_i$ is called the *context* of $\mathbb{T}$, written as $\mathfrak{I}(\mathbb{T}) = x$. The context $\mathfrak{I}$ defines a function from the set of coins $\mathbb{I}$ to the lattice $\mathbb{L}$ with the properties: (i) $\mathfrak{I}(\mathbb{T}) = \mathfrak{I}(\mathbb{T}^{-1})$, and (ii) sub-additivity: $\mathfrak{I}(\mathbb{T}\mathbb{T}')\leq\mathfrak{I}(\mathbb{T})\vee\mathfrak{I}(\mathbb{T}')$.

Now we make a further assumption that each nontrivial element x has a *canonical join representation*

$$x = x_1\vee x_2\vee\cdots\vee x_n \tag{8}$$

That is, (i) (8) is *irredundant* so that x cannot be the join of a proper subset of $\{x_1, \cdots, x_n\}$, and (ii) if $A \subset \mathbb{L}$ then $x = \vee A$ implies $\{x_1, \cdots, x_n\} \ll A$, that is, for all x_i we have $x_i \ll a$ for some $a \in A$. This is possible for any finite join-semidistributive lattice (Jónsson and Kiefer, 1962).

For (8) we shall write $\mathfrak{A}(x) = \{x_1, \cdots, x_n\}$. Expanding $\mathfrak{A}(x)$ by including all possible joins and meets of subsets of $\mathfrak{A}(x)$ we get a sublattice $\mathbb{L}_x$ of $\mathbb{L}$. We call $\mathbb{L}_x$ the sublattice of $\mathbb{L}$ generated by x. The sub-lattice $\mathbb{L}_x$ also induces a cainoid, which will be denoted by $\mathfrak{C}_x$. Note that an element of $\mathfrak{C}_x$ is necessarily contained in $\mathfrak{C}$.

DEFINITION 2.3 (MARGINAL CAINOID). *The cainoid* $(\mathfrak{C}_x, \cdot)$ *induced by the sub-lattice* $\mathbb{L}_x$ *is called the* marginal cainoid *of* x. *An element of* $\mathfrak{C}_x$ *is called a* marginal coin *of* x.

We use the notation $\mathbb{T}[x]$ to denote an arbitrary marginal coin of x. The set of all coins equivalent to $\mathbb{T}$ with respect to x is called, using standard group terminology, the *coset* of $\mathbb{T}$ with respect to x. This coset is given by $\mathbb{T}\mathfrak{C}_x \equiv \{\mathbb{T}[x]\mathbb{T} \mid \mathbb{T}[x] \in \mathfrak{C}_x\}$. The coset $\mathbb{T}\mathfrak{C}_x$ is also referred to as the *orbit* of $\mathbb{T}$ *caused* by the subgroup $\mathfrak{C}_x$.

2.3 The Cain

Now we shall introduce further structures into a cainoid. This is done by considering an operation analogous to the integration of ordinary functions.

We have already introduced the symbols $\mathbb{T}^x$ and $\mathbb{T}[x] \in \mathfrak{C}_x$. Now denote by $\mathbb{T}\{x\}$ an arbitrary coin in $\mathfrak{C}$ with context $\mathfrak{I}(\mathbb{T}\{x\}) = x$. Note that while $\mathfrak{I}(\mathbb{T}^x) = \mathfrak{I}(\mathbb{T}\{x\}) = x$, we have $\mathfrak{I}(\mathbb{T}[x]) \leq x$.

DEFINITION 2.4 (x-INTEGRABILITY). *Let* $x \in \mathbb{L}$. *A coin* $\mathbb{T}$ *is said* x-integrable *if* $\mathbb{T}$ *has context no less than* x, *that is,* $\mathbb{T} = \mathbb{T}\{y\}$ *with* $y \geq x$.

Note that when both $\mathbb{T}$ and $\mathbb{T}'$ are x-integrable, it does not follow that $\mathbb{T}\mathbb{T}'$ is x-integrable since some terms can cancell out. For instance, let $x = x_1 \vee x_2$ irredundantly, where x_1, x_2 are join irreducible. Let $\mathbb{T} = \mathbb{T}^{x_1}\mathbb{T}^{x_2}, \mathbb{T}' = \mathbb{T}_{x_1}\mathbb{T}^{x_2}$ then $\mathbb{T}\mathbb{T}' = (\mathbb{T}^{x_2})^2$ is not x-integrable because $\mathfrak{I}(\mathbb{T}\mathbb{T}') = x_2 < x$. It follows then the set of x-integrable coins is not closed under dot product.To define the coin integration we assume that $\mathbb{L}$ is a complemented distributive lattice.

DEFINITION 2.5 (COIN INTEGRATION). *For an arbitrary* $x \in \mathbb{L}$, *let* $\mathfrak{D}(x)$ *be the set of all* x-*integrable coins. The* x-integration *is a function, denoted by* $\int_x$, *from* $\mathfrak{D}(x)$ *into* $\mathfrak{C}$, $\int_x : \mathfrak{D}(x) \longrightarrow \mathfrak{C}$, *so that for any* $\mathbb{T}\{y\} \in \mathfrak{D}(x)$, *there is a unique coin* $\mathbb{T}\{y \wedge \bar{x}\} \in \mathfrak{C}$ *such that*

$$\int_x (\mathbb{T}\{y\}) = \mathbb{T}\{y \wedge \bar{x}\} \tag{9}$$

Further, the function $\int_x$ *satisfies the following properties (writing* $\int \mathbb{T}\, dx$ *instead of* $\int_x(\mathbb{T})$*):*

(i) If $\mathbb{T}^y$ *is* x-*integrable then*

$$\int \mathbb{T}^y\, dx = \mathbb{T}^{y \wedge \bar{x}} \tag{10}$$

(ii) Let $x = x_1 \vee x_2$ with $x_1 \wedge x_2 = \emptyset$ (that is, x_1 and x_2 are relative complements w.r.t. x). Let $\mathbb{T} = \mathbb{T}\{y_1\}\mathbb{T}\{y_2\}$ be x-integrable, where $\mathbb{T}\{y_1\}$ is x_1-integrable and $\mathbb{T}\{y_2\}$ is x_2-integrable. Further assume that $x_1 \wedge y_2 = x_2 \wedge y_1 = \emptyset$. Then it holds

$$\int (\mathbb{T}\{y_1\}\mathbb{T}\{y_2\})\, d(x_1 \vee x_2) = \int \mathbb{T}\{y_1\}dx_1 \int \mathbb{T}\{y_2\}dx_2 \tag{11}$$

(iii) For any $\mathbb{T} \in \mathfrak{C}$, it holds

$$\int \mathbb{T} d\emptyset = \mathbb{T} \tag{12}$$

Note that $\int_x(\mathbb{T})$ is defined for any $x \leq \mathfrak{I}(\mathbb{T})$. Axiom (10) is analogous to the definition of marginal probability density functions. Note that if $x = \emptyset$ then $\bar{x} = \top$, so $y \wedge \bar{x} = y$. Thus (10) implies $\int \mathbb{T}^y\, d\emptyset = \mathbb{T}^y$, a special case of (12). Axiom (11) is an analogue of the following property of the conventional integration $\int f(x,z)g(y,z)\,dxdy = \int f(x,z)\,dx \int g(y,z)\,dy$.

DEFINITION 2.6 (CAIN). *A cainoid $\mathfrak{C}$ is called a* cain *if $\mathfrak{C}$ further satisfies the axioms (10)-(12).*

Now we list some of the properties of the cain.

THEOREM 2.3. *If $\mathbb{T}$ is x-integrable and $x \wedge y = \emptyset$, then $\int \mathbb{T}[y]\mathbb{T}\,dx = \mathbb{T}[y] \int \mathbb{T}\,dx$*

THEOREM 2.4. *For any $x, y, z \in \mathbb{L}$, $\mathbb{T}_y^x = \mathbb{T}_z^x \implies \mathbb{T}_y^x = \mathbb{T}_{y\wedge z}^x$.*

THEOREM 2.5. *If $x \geq z, y \wedge z = \emptyset$ then $\int \mathbb{T}_y^x\, dz = \mathbb{T}_y^{x\wedge\bar{z}}$*

Theorem 2.5 tells us how to compute a 'conditionally marginal' coin from a 'conditional joint' coin. The next theorem is an analogy of the fact that (conditional) probability density functions integrate to unity.

THEOREM 2.6. *For any $x, y \in \mathbb{L}$ with $x > \emptyset$ we have $\int \mathbb{T}_y^x\, d(x \wedge \bar{y}) = 1 \qquad (x > \emptyset)$*

See Wang (2007) for more properties useful for transforming coin equations.

The following two theorems give two important rules concerning certian types of coin equations. These rules will play important roles in transforming coin identities.

THEOREM 2.7 (LAW OF NORMALIZATION). *Let $\bar{x}, \bar{y}$ be the complements of x and y respectively. Then*

$$\mathbb{T}_{y\vee z}^x = \mathbb{T}[\bar{z}] \;\Rightarrow\; \mathbb{T}_{y\vee z}^x = \mathbb{T}_{y\wedge\bar{z}}^{x\wedge\bar{z}} \qquad (x > \emptyset) \tag{13}$$

$$\mathbb{T}^{x\vee y\vee z} = \mathbb{T}[\bar{y}]\mathbb{T}[\bar{x}] \;\Rightarrow\; \mathbb{T}^{x\vee y\vee z} = \mathbb{T}^{(x\vee z)\wedge\bar{y}}\mathbb{T}^{(y\vee z)\wedge\bar{x}}\mathbb{T}_{z\wedge\bar{x}\wedge\bar{y}} \tag{14}$$

The following results are immediate consequences of Theorem 2.7. These special cases will be more frequently met in applications than general forms (13)-(14).

COROLLARY 2.1. *If x, y, z are nontrivial and mutually exclusive then*

$$\pi^x_{y\vee z} = \pi[\bar{z}] \;\Rightarrow\; \pi^x_{y\vee z} = \pi^x_y \qquad (x > \emptyset) \tag{15}$$

$$\pi^{x\vee y\vee z} = \pi[\bar{y}]\pi[\bar{x}] \;\Rightarrow\; \pi^{x\vee y\vee z} = \pi^{x\vee z}\pi^{y\vee z}\pi_z \tag{16}$$

COROLLARY 2.2. *For any x, y we have*

$$\pi^x_y = \pi[\bar{y}] \;\Rightarrow\; \pi^x_y = \pi^{x\wedge\bar{y}} \qquad (x > \emptyset) \tag{17}$$

$$\pi^{x\vee y} = \pi[\bar{y}]\pi[\bar{x}] \;\Rightarrow\; \pi^{x\vee y} = \pi^{x\wedge\bar{y}}\pi^{y\wedge\bar{x}} \tag{18}$$

When x, y are nontrivial and mutually exclusive then (17) and (18) reduce to, respectively

$$\pi^x_y = \pi[\bar{y}] \;\Rightarrow\; \pi^x_y = \pi^x \tag{19}$$

$$\pi^{x\vee y} = \pi[\bar{y}]\pi[\bar{x}] \;\Rightarrow\; \pi^{x\vee y} = \pi^x\pi^y \tag{20}$$

We shall refer to the *Law of Normalization* as the *N-Law* for short. The N-Law is a powerful principle that enables one to write an 'ambiguous' coin equation in an 'exact' form. This is useful, for instance, in situations when many atom coins enter into a coin equation but we are only interested in relations concerning a small portion of them. Those 'nuisance' coins can be treated as 'proportionality' constant.

A 'large' coin identity can give rise to many 'small' identities using the following *Law of Marginalization*, or the *M-Law*.

THEOREM 2.8 (LAW OF MARGINALIZATION). *If $x \wedge y = \emptyset$, then for any $a, b \in \mathbb{L}$,*

$$\pi^{x\vee y}_z = \pi^x_z\pi^y_z \;\Rightarrow\; \pi^{(x\wedge a)\vee(y\wedge b)}_z = \pi^{x\wedge a}_z\pi^{y\wedge b}_z \tag{21}$$

In particular, if $z = \emptyset$ then $\pi^{x\vee y} = \pi^x\pi^y \;\Rightarrow\; \pi^{(x\wedge a)\vee(y\wedge b)} = \pi^{x\wedge a}\pi^{y\wedge b}$.

The following rule is also useful in marginalizing a 'fat' coin equation.

THEOREM 2.9. *If x, y, z are mutually exclusive then*

$$\pi^{x\vee y\vee z} = \pi[\bar{z}]\pi^{y\vee z} \Rightarrow \pi^{x\vee y} = \pi[\bar{z}]\pi^y \tag{22}$$

3 The Graphoid and the Separoid

3.1 Conditional Independence

Now we shall assume that $\mathbb{L}$ is a Boolean algebra. That is, $\mathbb{L}$ is a complemented distributive lattice with bottom $\emptyset$ and top $\top > \emptyset$. Let $\mathfrak{C}$ be the cain defined on $\mathbb{L}$.

DEFINITION 3.1. *x is said independent of y conditional on z, written $x \perp\!\!\!\perp y|z$, iff $\pi^x_{y\vee z} = \pi^x_z$.*

We shall write $x \perp\!\!\!\perp y$ if $x \perp\!\!\!\perp y|\emptyset$, and say in this case that x is independent of y. So independence is a special case of conditional independence. Applying the R- and L-Laws, we immediately see that, if $x, y, z > \emptyset$ then $\pi^x_{y \vee z} = \pi^x_z$ is equivalent to either $\pi^{x \vee y}_z = \pi^x_z \pi^y_z$ or $\pi^{x \vee y \vee z} = \pi^{y \vee z} \pi^x_z$. If $x, y > \emptyset$, then $\pi^x_y = \pi^x$ and $\pi^{x \vee y} = \pi^x \pi^y$ give equivalent conditions for $x \perp\!\!\!\perp y$. Note that $x \perp\!\!\!\perp y|z$ is symmetric for x and y, i.e., $x \perp\!\!\!\perp y|z \Rightarrow y \perp\!\!\!\perp x|z$, unless $\emptyset = y < x \not\leq z$.

It can be shown that the embedding lattice order $x \leq y$ is equivalent to $x \perp\!\!\!\perp x|y$. It can also be shown that $y \leq z \Rightarrow x \perp\!\!\!\perp y|z$ for any x. For instance, $x \perp\!\!\!\perp \emptyset|y$, $x \perp\!\!\!\perp y|\top$, $x \perp\!\!\!\perp y|y$ for any x, y. If $\emptyset < x$ then for any y we have $x \leq z \Rightarrow x \perp\!\!\!\perp y|z$.

As an immediate consequence of the N-Law (16), we have the following seemingly weaker but equivalent condition for conditional independence. For the usual probabilistic conditional independence, this condition is commonly known as the Factorization Theorem.

THEOREM 3.1 (FACTORIZATION). *If nontrivial elements x, y, z are also mutually exclusive, then*

$$x \perp\!\!\!\perp y|z \Leftrightarrow \pi^{x \vee y \vee z} = \pi[\bar{x}]\pi[\bar{y}] \tag{23}$$

where $\pi[\bar{x}]$ and $\pi[\bar{y}]$ are appropriate coins of the marginal cains $\mathfrak{C}_{\bar{x}}$ and $\mathfrak{C}_{\bar{y}}$ respectively.

As an illustration, if $D = \{1, 2, 3\}$ then any one of the following equations, $\pi^{12}_3 = \pi^1_3 \pi^2_3$, $\pi^{123} = \pi^{13}\pi^2_3$, $\pi^{123} = \pi^{23}\pi^1_3$, $\pi^1_{23} = \pi^1_3$, $\pi^2_{13} = \pi^2_3$, gives a necessary and sufficient condition for $1 \perp\!\!\!\perp 2|3$.

3.2 The Graphoid

The following properties of conditional independence can be derived from the axioms of the cain. These properties play an important role in statistical graphical models (Lauritzen, 1996).

THEOREM 3.2 (DECOMPOSITION). *If $x \wedge y = \emptyset$, then for any $a \leq x, b \leq y$ it holds that $x \perp\!\!\!\perp y|z \Rightarrow a \perp\!\!\!\perp b|z$.*

THEOREM 3.3 (WEAK UNION). *If x, y, z, w are mutually exclusive, then $x \perp\!\!\!\perp (y \vee z)|w \Rightarrow x \perp\!\!\!\perp y|(z \vee w)$.*

The following theorem gives a seemingly weaker sufficient and necessary condition for the joint conditional independence $x \perp\!\!\!\perp (y \vee z)|w$. This condition is sometimes referred to as the *contraction* property (Pear, 2000, p.11).

THEOREM 3.4 (CONTRACTION). *If x, y, z, w are nontrivial and mutually exclusive, then we have*

$$\left.\begin{array}{c} x \perp\!\!\!\perp y \mid (z \vee w) \\ x \perp\!\!\!\perp z \mid w \end{array}\right\} \Leftrightarrow x \perp\!\!\!\perp (y \vee z) \mid w$$

THEOREM 3.5 (INTERSECTION). *If x, y, z, w are nontrivial and mutually exclusive, then*

$$\left.\begin{array}{c} x \perp\!\!\!\perp y \mid (z \vee w) \\ x \perp\!\!\!\perp z \mid (y \vee w) \end{array}\right\} \Leftrightarrow x \perp\!\!\!\perp (y \vee z) \mid w \tag{24}$$

Recall that a ternary relation $\cdot \amalg \cdot | \cdot$ defined on a Boolean algebra $\mathbb{L}$ is called a graphoid, if for all nontrivial and mutually exclusive elements we have

$$
\begin{array}{lll}
\text{G1}: & x \amalg y | z \Rightarrow y \amalg x | z & \text{(Symmetry)} \\
\text{G2}: & x \amalg (y \vee w) | z \Rightarrow x \amalg y | z & \text{(Decomposition)} \\
\text{G3}: & x \amalg (y \vee z) | w \Rightarrow x \amalg y | (z \vee w) & \text{(Weak union)} \\
\text{G4}: & x \amalg y | (z \vee w),\ x \amalg z | w \Rightarrow x \amalg (y \vee z) \mid w & \text{(Contraction)} \\
\text{G5}: & x \amalg y | (z \vee w),\ x \amalg z | (y \vee w) \Rightarrow x \amalg (y \vee z) \mid w & \text{(Intersection)}
\end{array}
$$

In the finitary case, properties G1-G5 were discussed by Dawid (1979) and Spohn (1980). The name graphoid was due to Pearl and Paz (1987), who used G1-G5 as axioms to characterize the relation between graphs and informational relevance; see also Pearl(200). The results of this section show that if $\mathbb{L}$ is a Boolean algebra, then $\cdot \perp\!\!\!\perp \cdot | \cdot$ is a graphoid. Properties of the classical probabilistic conditional independence are not limited to those discussed in the previous subsection. In fact Studený (1992) showed that it is not possible to give a finite characterization of conditional independence relations. See Wang (2007) for discussions on more useful properties concerning conditional independence.

3.3 The Separoid

The cain algebra is consistent with a stronger system than the graphoid. In this section we show that the cain algebra satisfies the defining axioms of a strong separoid of Dawid (2001). The separoid includes several axiomatic systems, such as the orthogonoid and the graphoid, relevant for formal reasoning using the concept of *irrelevance* of information.

The concept of a separoid was invented by Dawid (2001). In the very last paragraph of his paper, Professor Dawid concluded that "New concepts of 'irrelevance' in a wide variety of settings, have been proliferating rapidly, and this can be expected to continue. It should be regarded as a matter of course, when any such new definition is proposed, that a check be made to see whether it satisfies the separoid axioms P1-P5. If not, departures from the axioms should be carefully identified and justified". Dawid showed that both the orthgoid and the graphoid can be regarded as a special case of the separoid. In this section we show that the universal algebraic definition of the cain conditional independence does satisfy the defining axioms of the strong separiod.

DEFINITION 3.2 (SEPAROID). *Let $(S, \leq)$ be a join-semilattice. Let $\cdot \perp\!\!\!\perp \cdot | \cdot$ be a ternary relation on S. Then $(S, \leq, \perp)$ is a* separoid *if the following hold*

$$
\begin{array}{lll}
\text{P1}: & x \perp y \mid x & \\
\text{P2}: & x \perp y \mid z & \Longrightarrow y \perp x \mid z \\
\text{P3}: & x \perp y \mid z \,\&\, w \leq y & \Longrightarrow x \perp w \mid z \\
\text{P4}: & x \perp y \mid z \,\&\, w \leq y & \Longrightarrow x \perp y \mid (z \vee w) \\
\text{P5}: & x \perp y \mid z \,\&\, x \perp w \mid (y \vee z) & \Longrightarrow x \perp (y \vee w) \mid z
\end{array}
$$

REMARK 3.1. *None of the axioms* P1-P5 *is* equational *in form, thus the description of a separoid $(S, \leq, \perp\!\!\!\perp)$ as given in Definition 3.2 does not constitute a universal algebra.*

DEFINITION 3.3 (STRONG SEPAROID). *A separoid* $(S, \leq, \perp)$ *is said* strong *if* $(S, \leq)$ *is a lattice and the following additional property holds*

$$\text{P6: If } z \leq y \;\&\; w \leq y \text{ then}$$
$$x \perp y \mid z \;\&\; x \perp y \mid w \implies x \perp y \mid (z \wedge w)$$

Now we show that P1-P6 are satisfied by a cain conditional independence, $\cdot \perp\!\!\!\perp \cdot \mid \cdot$.

THEOREM 3.6. *Let* $\mathfrak{C}$ *be a cain generated by a complemented distributive lattice* $(\mathbb{L}, \leq)$. *We say that a ternary relation* $x \perp\!\!\!\perp y | z$ *holds for nontrivial elements* x, y, z *in* $\mathbb{L}$ *if the coin idenitity* $\pi_z^{x \vee y} = \pi_z^x \pi_z^y$ *holds in* $\mathfrak{C}$. *Then* $(\mathbb{L}, \leq, \perp\!\!\!\perp)$ *is a strong separoid.*

See Wang (2007) for details of the proof.

4 Discussions

In this paper we introduced a new universal algebraic system called the cain. This framework is built upon some essential algebraic properties of the probability density functions. In the cain algebra conditional independence is defined through a coin equation, which can be easily transformed into other equivalent equations using the rules of the algebra. In particular we have showed the consistency with the well-known axiomatic systems of graphoid and the separoid. This universal algebraic approach provides possibilities for making statistical causal inference using the new algebraic tools combing with modern techniques such as the Gröbner basis theories.

References

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

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

[3] Jónsson, B. and Kiefer, J. (1962). Finite sublattices of a free lattice. *Canad. J. Math.* **14** 487–497.

[4] Lauritzen, S. L. (1996). *Graphical Models*. Clarendon, Oxford.

[5] Pearl, J. (2000). *Causality–models, reasoning and inference*. Cambridge University Press, New York.

[6] Pearl, J. and Paz, A. (1987). Graphoids: a graph-based logic for reasoning about relevance relations. In *Advances in Artificial Intelligence* (D. Hogg and L. Steels, eds.) 357–363, North-Holland, Amsterdam.

[7] Spohn, W. (1980). Stochastic independence, causal independence, and shieldability. *J. Phil. Logic* **9** 73–99.

[8] Studený, M. (1992). Conditional independence relations have no finite complete characterization. In *Transactions of the 11th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes* 377–396. Academia, Prague.

[9] Wang, J. (2006). The coin algebra for conditional independence. 数理解析研究所講究録 **1506**, Statistical Conditional Inference and Its Related Topics, 2006 年 7 月, 177-196.

[10] Wang, J. (2007). The cain algebra: a universal algebraic approach for probabilistic conditional independence. Submitted to *The Annals of Statistics*.