Ann Inst Stat Math (2010) 62:747–773
DOI 10.1007/s10463-010-0278-5

# A universal algebraic approach for conditional independence

**Jinfang Wang**

Received: 27 April 2009 / Revised: 4 December 2009 / Published online: 31 March 2010


**Abstract** In this paper we show that elementary properties of joint probability density functions naturally induce a universal algebraic structure suitable for studying probabilistic conditional independence (PCI) relations. We call this algebraic system the cain. In the cain algebra, PCI relations are represented in equational forms. In particular, we show that the cain satisfies the axioms of the graphoid of Pearl and Paz (Advances in artificial intelligence. North-Holland, Amsterdam, 1987) and the separoid of Dawid (Ann. Math. Artif. Intell. 32:335–372, 2001), these axiomatic systems being useful for general probabilistic reasoning.

**Keywords** Cain · Conditional independence · Graphical model · Graphoid · Probability density function · Separoid

## 1 Introduction

The concept of conditional independence (CI), both probabilistic and non-probabilistic, has found its use in many fields including statistics and artificial intelligence. Various formal theories exist for exploring the nature of CI. These include, but are not limited to, the probabilistic approach of Dawid (1979a) and Spohn (1980), Dempster–Shafer's theory of evidence [see Shafer (1976) and Shenoy (1994). See also Schoken and Hummel (1993) and Teixeira de Silva and Milidiu (1993) for applications], theories in relational databases (Sagiv and Walecka 1982), Spohn's kappa

This article was submitted to "Special Section on Algebraic Methods in Computational Statistics".

J. Wang (✉)
Department of Mathematics and Informatics, Graduate School of Science,
Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan
e-mail: wang@math.s.chiba-u.ac.jp

calculus (Spohn 1988), Shenoy's theory of valuation-based system (Shenoy 1994), which unifies Dempster–Shafer's theory of evidence and Zadeh's possibility theory approach (Zadeh 1978). Studený (1993, 2005) gave a comparison for some of these theories. Studený (2005) also proposes the use of structural imsets for describing CI structures for discrete random variables.

The importance of the concept of probabilistic conditional independence (PCI) in areas such as probabilistic theory of causation was recognized as early as Reichenbach (1956, Sect. 23), while Dawid (1979a) is generally regarded as the first attempt of an axiomatic treatment of PCI. Dawid's (1979a) work has been further developed in statistics by Dawid (1979b, 1980a,b, 1985); Lauritzen et al. (1990); Dawid and Lauritzen (1993); Dawid and Mortera (1996); Dawid and Evett (1997); Dawid and Studený (1999), among others. For an overview, see Dawid (1988). In epistemology, Spohn (1980) gave an essentially the same but more explicit axiomatic formation of PCI; see also Spohn (1994). In artificial intelligence, Pearl and Paz (1987) proposed a closely related axiomatic system termed the graphoid with a strong emphasis on graphical applications. For further developments of their work, see Paz and Pearl (1994), Paz et al. (1996), Galles and Pearl (1997), and Pearl (1988, 2000). Recently, Dawid (2001) introduced an axiomatic system termed the *separoid*, which unifies a couple of different notions of 'irrelevance' including the orthogonoid and the graphoid.

The motivation of this paper is to develop a theory of PCI in a purely universal algebraic fashion so that PCI relations are all represented in equational forms. One potential advantage of this approach is that one may derive PCI relations from other PCI relations automatically using computer algorithms. Thus, our theory is different from other existing axiomatic systems for PCI, such as the graphoid of Pearl and Paz (1987) and the separoid of Dawid (2001), in that the latter systems are built on some principal properties of PCI useful for general probabilistic reasoning.

Our idea is to (a) formalize a set of key properties to probability density functions and (b) to use these properties in a purely axiomatic fashion to derive PCI relations. In particular, we show that the cain satisfies the axioms of both the graphoid and the separoid. We will begin our theory by introducing a mathematical object termed the *coin* (deriving from *co*nditional *in*dependence). A coin is an algebraic abstraction of a joint probability density function. Coins can be multiplied to produce another coin, just as density functions are multiplied for independent random variables to produce joint density functions. The coin product obeys the rules of an Abelian group, with a further requirement parallel to the definition of a conditional density function. The set of all coins equipped with this product defines a universal algebra called a *cainoid*. A cainoid is not enough for studying PCI relations. We further allow a coin to be *marginalized* to give rise to another coin, just as a joint probability density function can be integrated to produce a marginal density function. Again, all rules concerning the marginalization will be stated in equational forms. A cainoid with this marginalization operator is called a *cain* (deriving from *ca*usal *in*ference). In a cain, PCI relations are defined through *coin equations*.

The rest of the paper is organized as follows. In Sect. 2 we introduce the concept of a coin, which is an algebraic abstraction of a joint density function. Elementary properties intrinsic to density functions then naturally lead to an algebraic system called the *cainoid*. In order to study PCI relations, it is necessary to give a careful treatment of the relations between marginal and joint probability density functions, and this is the topic of Sect.3. The *cain* is defined in Sect. 3 as a cainoid equipped with a further operator called *marginalization*. Section 4 studies CI relations using the axioms of a cain. In particular, we show that the cain-algebraic PCI relations satisfy the requirements of a (strong) graphoid of Dawid (2001), which includes the graphoid of Pearl and Paz (1987) as a special case. Finally, Sect. 5 gives a short discussion and the Appendix contains proofs of some major results.

## 2 The Cainoid

### 2.1 Definition and basic properties

Let $(\mathbb{L}, \leq)$ be a *lattice*, where $\mathbb{L}$ is a nonempty (possibly countably infinite) set, and $\leq$ is a partial order in $\mathbb{L}$. We denote by $x \vee y$ the *join* (least upper bound) and $x \wedge y$ the *meet* (greatest lower bound) for any $x, y \in \mathbb{L}$. Throughout the paper we shall suppose that $(\mathbb{L}, \leq)$ is *bounded below*, i.e., there exists a *bottom* $\emptyset$ with $\emptyset \leq x$ for any $x \in \mathbb{L}$. By definition, $\emptyset$ is unique. An element $x \in \mathbb{L}$ is said to be *nontrivial* if $x > \emptyset$. A set of elements are said to be *mutually exclusive* if their pairwise meets are trivial.

*Example 1 (The finitary case)* When investigating the cause and effect relations among a set of finite number of random variables, we, for convenience, can represent the set of variables by $D = \{1, 2, \dots, d\}$. Each element of $D$ corresponds to a *random variable*. In this case, the underlying lattice $(\mathbb{L}, \leq) = (2^D, \subseteq)$ is the Boolean algebra, where $2^D$ is the power set of $D$ and $\subseteq$ is the usual set inclusion. Note that if $\emptyset$ denotes the empty set, then $x \geq \emptyset$ holds for any $x \subset D$. The bottom $\emptyset$ plays the role of representing the state of 'ignorance'. Note also that all nonempty subsets (random vectors) are nontrivial.

The above example represents any finite Boolean lattice $\mathbb{L}$ by the Stone's representation theorem (Halmos 1974), which says that any Boolean lattice is isomorphic to a power set.

To introduce more algebraic structures in $\mathbb{L}$ for studying PCI relations among the elements in $\mathbb{L}$, consider the direct product $\mathbb{L} \otimes \mathbb{L} = \{(x, y) \,|\, x, y \in \mathbb{L}\}$. Note that $(x, y) \neq (y, x)$ if $x \neq y$. To emphasize this asymmetry, we replace $(x, y)$ by using a new symbol $\mathbb{T}_y^x$ (reads as *coin-x-over-y*). The following conventions

$$\mathbb{T}^x = \mathbb{T}_\emptyset^x, \; \mathbb{T}_y = \mathbb{T}_y^\emptyset, \; \mathbb{T}_\emptyset^\emptyset = 1$$

will be adopted throughout the paper. We call $\mathbb{T}^x$ the *raising coin* with *context* $x$, $\mathbb{T}_y$ the *lowering coin* with context $y$, and $\mathbb{T}_y^x$ the *mixed coin* with raising context $x$ and lowering context $y$. All these coins are called *atom coins*. Justification for these terminologies will be made later.

We define a *coin* to be a string or 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. Denote by

$$\mathfrak{C} = \left\{ \mathbb{T}_{y_1}^{x_1} \cdots \mathbb{T}_{y_n}^{x_n} | x_i, y_i \in \mathbb{L}, i = 1, \dots, n, n \in \mathbb{N} \right\} \tag{1}$$

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

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

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

$$\begin{aligned}
&\text{C1:} \;\; \mathbb{T} \cdot \mathbb{T}' = \mathbb{T}' \cdot \mathbb{T} \\
&\text{C2:} \;\; \left(\mathbb{T} \cdot \mathbb{T}'\right) \cdot \mathbb{T}'' = \mathbb{T} \cdot \left(\mathbb{T}' \cdot \mathbb{T}''\right) \\
&\text{C3:} \;\; 1 \cdot \mathbb{T} = \mathbb{T} \\
&\text{C4:} \;\; \mathbb{T}^x \cdot \mathbb{T}_x = 1 \\
&\text{C5:} \;\; \mathbb{T}^x_y = \mathbb{T}^{x \vee y} \cdot \mathbb{T}_y \quad (\text{where } x > \emptyset).
\end{aligned}$$

C1 and C2 state that the dot product is *commutative* and *associative*. It follows that a coin $\mathbb{T}^{x_1}_{y_1} \cdots \mathbb{T}^{x_n}_{y_n}$ is the unique product of $n$ atom coins $\mathbb{T}^{x_1}_{y_1}, \ldots, \mathbb{T}^{x_n}_{y_n}$. C3 says that $\mathbb{T}^{\emptyset}_{\emptyset} = 1$ is the *unit element*. C4 says that a lowering coin is an *inverse* of the corresponding raising coin with the same context and vice versa. C5 says that, a mixed coin $\mathbb{T}^x_y$, with nontrivial raising context $x$, is the product of the raising coin $\mathbb{T}^{x \vee y}$ and the lowering coin $\mathbb{T}_y$. Note that by C3, C5 also holds for $x = y = \emptyset$. So C5 holds for all $x$ and $y$ except when $x = \emptyset$, $y > \emptyset$. Among these axioms, C5 is the most important one, which defines a mixed coin in terms of a raising coin and a lowering coin. In the rest of the paper we sometimes omit the dot symbol in coin products for easy readability.

*Remark 1* In Definition 1, $\mathbb{T}^x$, $\mathbb{T}_x$ and $\mathbb{T}^x_y$ may be regarded as algebraic abstractions of the joint probability density function (p.d.f.) $f(x)$, the reciprocal $1/f(x)$, and the conditional p.d.f. $f(x|y)$, respectively. So a coin $\mathbb{T}^{x_1}_{y_1} \cdots \mathbb{T}^{x_n}_{y_n}$ may then be regarded as an abstraction of the product $f(x_1|y_1) \cdots f(x_n|y_n)$ of $n$ conditional p.d.f.'s. Axiom C5 corresponds to the definition of conditional density functions.

**Lemma 1** *(i) For any $x$, $\mathbb{T}^x_x = 1$; (ii) If $x \geq y$, then $\mathbb{T}^x_y = \mathbb{T}^x \mathbb{T}_y$; and (iii) If $x > \emptyset$ and $w \leq x \vee y$, then $\mathbb{T}^x_y = \mathbb{T}^{x \vee w}_y$.*

*Proof* For (i), $\mathbb{T}^x_x = \mathbb{T}^{x \vee x} \mathbb{T}_x = \mathbb{T}^x \mathbb{T}_x = 1$, the first equality being due to C5, the second one due to the absorption law of $\vee$, and the third one due to C4.

For (ii), $\mathbb{T}^x_y = \mathbb{T}^{x \vee y} \mathbb{T}_y = \mathbb{T}^x \mathbb{T}_y$, the first equality being due to C5, the second one due to $x = x \vee y \Leftrightarrow x \geq y$.

For (iii), $\mathbb{T}^x_y = \mathbb{T}^{x \vee y} \mathbb{T}_y = \mathbb{T}^{(x \vee y) \vee w} \mathbb{T}_y = \mathbb{T}^{(x \vee w) \vee y} \mathbb{T}_y = \mathbb{T}^{x \vee w}_y$, the first equality being due to C5, the second one due to $x \vee y = (x \vee y) \vee w \Leftrightarrow x \vee y \geq w$, the third one due to associative law of $\vee$, the fourth one due to C5 again. □

The following result represents the lattice order using a coin identity:

**Proposition 1** *If $x > \emptyset$, then $x \leq y \Leftrightarrow \mathbb{T}^x_y = 1$.*

*Proof* By C5, we have $\mathbb{T}^x_y = \mathbb{T}^{x\vee y}\mathbb{T}_y = 1 \Leftrightarrow \mathbb{T}^{x\vee y} = \mathbb{T}^y \Leftrightarrow x \vee y = y \Leftrightarrow x \leq y$. □

**Proposition 2** *(Bayes' theorem) If $x > \emptyset$, $y > \emptyset$, then $\mathbb{T}^x_y = \mathbb{T}^y_x \mathbb{T}^x \mathbb{T}_y$.*

*Proof* By C1–C5, the r.h.s. of the equation can be computed as follows:

$$\begin{aligned} \mathbb{T}^y_x \mathbb{T}^x \mathbb{T}_y &= \left(\mathbb{T}^{x\vee y}\mathbb{T}_x\right)\left(\mathbb{T}^x \mathbb{T}_y\right) \\ &= \mathbb{T}^{x\vee y}\left(\mathbb{T}_x \mathbb{T}^x\right)\mathbb{T}_y \\ &= \mathbb{T}^{x\vee y}\mathbb{T}_y \\ &= \mathbb{T}^x_y \end{aligned}$$

the first equality being due to C2 and C5, the second one to C2, the third one to C3–C4, and the last one to C5 again. □

**Proposition 3** *A cainoid $(\mathfrak{C}, \cdot)$ is an Abelian group.*

*Proof* Since the other properties of a group are obviously satisfied by $(\mathfrak{C}, \cdot)$, we only need to check that for every coin $\mathbb{T} \in \mathfrak{C}$ there exists $\mathbb{T}' \in \mathfrak{C}$ so that $\mathbb{T}\,\mathbb{T}' = 1$. First consider atom coins. If $\mathbb{T} = \mathbb{T}_y$, then $\mathbb{T}_y \mathbb{T}^y = 1$. If $\mathbb{T} = \mathbb{T}^x_y$ and $x > \emptyset$, then

$$\begin{aligned} \mathbb{T}^x_y\left(\mathbb{T}_{x\vee y}\mathbb{T}^y\right) &= \left(\mathbb{T}^{x\vee y}\mathbb{T}_y\right)\left(\mathbb{T}_{x\vee y}\mathbb{T}^y\right) && \text{(by C5)} \\ &= \left(\mathbb{T}^{x\vee y}\mathbb{T}_{x\vee y}\right)\left(\mathbb{T}_y \mathbb{T}^y\right) && \text{(by C1 \& C2)} \\ &= 1 && \text{(by C3 \& C4).} \end{aligned}$$

Now let $\mathbb{T} = \prod_{i=1}^s \mathbb{T}^{x_i}_{y_i}$ be an arbitrary coin, and $(\mathbb{T}^{x_i}_{y_i})'$ be an inverse of $\mathbb{T}^{x_i}_{y_i}$. Then, by C1–C3, $\mathbb{T}\prod_{i=1}^s(\mathbb{T}^{x_i}_{y_i})' = \prod_{i=1}^s(\mathbb{T}^{x_i}_{y_i}(\mathbb{T}^{x_i}_{y_i})') = 1$, showing that $\prod_{i=1}^s(\mathbb{T}^{x_i}_{y_i})'$ is an inverse of $\mathbb{T}$. □

Note that since a cainoid $(\mathfrak{C}, \cdot)$ is a group, both the unit element and the inverse element are unique. By the uniqueness of an inverse element, we shall use the symbol $(\mathbb{T})^{-1}$ or simply $\mathbb{T}^{-1}$ to denote the inverse of $\mathbb{T}$. Due to the fact that $\mathbb{T}^x_y = \mathbb{T}^{x\vee y}\mathbb{T}_y = \mathbb{T}^{x\vee y}(\mathbb{T}^y)^{-1}$, we have

**Proposition 4** *Any cainoid $\mathfrak{C}$ is generated by raising coins.*

For instance, consider the finitary case with $D = \{1, 2\}$. Let $(\mathbb{T})^0 = 1$; we can then write $\mathfrak{C} = \{(\mathbb{T}^1)^{n_1}(\mathbb{T}^2)^{n_2}(\mathbb{T}^{12})^{n_3} | n_1, n_2, n_3 = 0, \pm 1, \pm 2, \ldots\}$.

## 2.2 Coin identities

We have defined the cainoid using only equational axioms C1–C5. In so doing we are able to present relations among coins in terms of coin identities. The Bayes' Theorem, Proposition 2, for instance, is expressed in a form of coin identity. As a second example, suppose that $x_1, \ldots, x_n$ are nontrivial elements in $\mathbb{L}$ (a condition slightly stronger than necessary); then using C3–C5 we obtain the identity, $\mathbb{T}^{x_1 \vee \cdots \vee x_n} = \mathbb{T}^{x_1} \mathbb{T}^{x_2}_{x_1} \mathbb{T}^{x_3}_{x_1 \vee x_2} \cdots \mathbb{T}^{x_n}_{x_1 \vee \cdots \vee x_{n-1}}$, which corresponds to a well-known formula for density decomposition.

Given any identity, there are infinitely many identities which are *equivalent* to it. For instance, suppose that $x, y, z$ are all nontrivial and $\mathbb{T}^{x \vee y}_z = \mathbb{T}^x_z \mathbb{T}^y_z$ holds. Multiplying both sides by $\mathbb{T}^z$ and using C5, we get $\mathbb{T}^{x \vee y \vee z} = \mathbb{T}^x_z \mathbb{T}^{y \vee z}$, which when multiplied on both sides by $\mathbb{T}_{y \vee z}$ and using C3–C5, then leads to $\mathbb{T}^x_{y \vee z} = \mathbb{T}^x_z$. Thus, if one of these identities is interpreted as representing the situation that '*x is irrelevant to y given z*', then all the other equivalent identities give sufficient and necessary conditions for this 'conditional independence'. Using C1–C5, one can easily derive other equivalent forms.

More formally, we call an equation $\mathbb{T} = \mathbb{T}'$ a *coin identity*, where $\mathbb{T}$ and $\mathbb{T}'$ are two coins in $\mathfrak{C}$, if and only if $\mathbb{T}(\mathbb{T}')^{-1} = 1$. We shall denote the set of all coin identities by $\mathbb{I}$, and the elements of $\mathbb{I}$ by $\mathfrak{Q}_1, \mathfrak{Q}_2$, etc. Now consider the group action of $\mathfrak{C}$ on $\mathbb{I}$. For each $\mathbb{T} \in \mathfrak{C}$ and $\mathfrak{Q} = \{\mathbb{T}' = \mathbb{T}''\} \in \mathbb{I}$, we define

$$\mathbb{T}(\mathfrak{Q}) = \{\mathbb{T}\mathbb{T}' = \mathbb{T}\mathbb{T}''\}. \tag{3}$$

**Definition 2** For each coin identity $\mathfrak{Q} \in \mathbb{I}$, the set

$$\mathfrak{C}\mathfrak{Q} = \left\{\mathbb{T}(\mathfrak{Q}) | \mathbb{T} \in \mathfrak{C}\right\}$$

is called the *orbit* of $\mathfrak{Q}$, where $\mathbb{T}(\mathfrak{Q})$ is defined by (3). Two coin identities $\mathfrak{Q}_1$ and $\mathfrak{Q}_2$ are said *conjugate* or *equivalent* to one another, written $\mathfrak{Q}_1 \cong \mathfrak{Q}_2$, if and only if they belong to the same orbit $\mathfrak{C}\mathfrak{Q}$ for some $\mathfrak{Q} \in \mathbb{I}$.

### 2.3 The R-law and the L-law

The following two rules are useful for transforming coin identities:

**Lemma 2** *(raising-up law) For any $x > \emptyset$, $y, z \in \mathbb{L}$, we have*

$$\mathbb{T}^{x \vee y} = \mathbb{T}^x_y \mathbb{T}^y \quad (x > \emptyset) \tag{4}$$

$$\mathbb{T}^{x \vee y}_z = \mathbb{T}^x_{y \vee z} \mathbb{T}^y \Leftrightarrow \mathbb{T}^{y \vee z} = \mathbb{T}^y \mathbb{T}^z \quad (x > \emptyset). \tag{5}$$

Multiplying $\mathbb{T}^x_y$ by $\mathbb{T}^y$ *raises* the context $y$, giving $\mathbb{T}^{x \vee y}$ on the l.h.s. of (4). Similarly, $\mathbb{T}^{x \vee y}_z$ on the l.h.s. of (5) is obtained by *detaching* the context $y$ from $y \vee z$ of $\mathbb{T}^x_{y \vee z}$ and *raising* $y$ using $\mathbb{T}^y$. To ensure the validity of the detachment, we need the 'independence condition', $\mathbb{T}^{y \vee z} = \mathbb{T}^y \mathbb{T}^z$. Note that (5) reduces to (4) when $z = \emptyset$.

*Proof* We only need to prove (5). First note that $\mathbb{T}^{x\vee y\vee z} = \mathbb{T}_z^{x\vee y}\mathbb{T}^z = \mathbb{T}_{y\vee z}^x\mathbb{T}^{y\vee z}$ holds by C2, C4, and C5. So we have

$$\begin{aligned}\mathbb{T}_z^{x\vee y} &= \mathbb{T}_z^{x\vee y}\left(\mathbb{T}^z\mathbb{T}_z\right)\\ &= \left(\mathbb{T}_z^{x\vee y}\mathbb{T}^z\right)\mathbb{T}_z\\ &= \left(\mathbb{T}_{y\vee z}^x\mathbb{T}^{y\vee z}\right)\mathbb{T}_z\\ &= \mathbb{T}_{y\vee z}^x\left(\mathbb{T}^y\mathbb{T}_y\right)\left(\mathbb{T}^{y\vee z}\mathbb{T}_z\right)\\ &= \left(\mathbb{T}_{y\vee z}^x\mathbb{T}^y\right)\left(\mathbb{T}_y\mathbb{T}^{y\vee z}\mathbb{T}_z\right)\end{aligned}$$

implying that

$$\begin{aligned}\mathbb{T}_z^{x\vee y} = \mathbb{T}_{y\vee z}^x\mathbb{T}^y &\Longleftrightarrow \mathbb{T}^{y\vee z}\mathbb{T}_y\mathbb{T}_z = 1\\ &\Longleftrightarrow \mathbb{T}^{y\vee z} = \mathbb{T}^y\mathbb{T}^z.\end{aligned}$$

□

In the sequel, we shall refer to (4) and (5) as the *raising-up law*, or the *R-law* for short.

**Lemma 3** *(lowering-down law) For any $x > \emptyset$, $y, z \in \mathbb{L}$, we have*

$$\mathbb{T}_y^x = \mathbb{T}^{x\vee y}\mathbb{T}_y \quad (x > \emptyset) \tag{6}$$

$$\mathbb{T}_{y\vee z}^x = \mathbb{T}_z^{x\vee y}\mathbb{T}_y \Leftrightarrow \mathbb{T}^{y\vee z} = \mathbb{T}^y\mathbb{T}^z \quad (x > \emptyset). \tag{7}$$

(6) is nothing but C5. The coin $\mathbb{T}_y^x$ on the l.h.s. of (6) is obtained by *lowering* the context $y$ of $\mathbb{T}^{x\vee y}$ using $\mathbb{T}_y$. Similarly, the coin $\mathbb{T}_{y\vee z}^x$ on the l.h.s. of (7) is obtained by *lowering* $y$ in $\mathbb{T}_z^{x\vee y}$ using $\mathbb{T}_y$, and *joining* $y$ with $z$ when $\mathbb{T}^{y\vee z} = \mathbb{T}^y\mathbb{T}^z$ holds. If $z = \emptyset$ then (7) reduces to (6). We omit the proof, which is similar to that of Lemma 2. We shall refer to (6) and (7) as the *lowering-down law*, or the *L-law* for short. Note that the L-law is a direct consequence of the R-Law and C4. The following result is also sometimes useful:

**Lemma 4** *If $x \vee z$, $x \vee w$, $y \vee z$, $y \vee w$ are nontrivial, then $\mathbb{T}_x^{y\vee z} = \mathbb{T}_x^{y\vee w}\mathbb{T}$ holds for some $\mathbb{T} \in \mathfrak{C}$ if and only if $\mathbb{T}_y^{x\vee z} = \mathbb{T}_y^{x\vee w}\mathbb{T}$ holds true.*

The second identity in Lemma 4 is obtained from the first one by interchanging the roles of $x$ and $y$.

*Proof* Acting $\mathbb{T}^x$ on $\mathbb{T}_x^{y\vee z} = \mathbb{T}_x^{y\vee w}\mathbb{T}$, and using the R-Law, we get $\mathbb{T}^{x\vee y\vee z} = \mathbb{T}^{x\vee y\vee w}\mathbb{T}$, which, when acted upon by $\mathbb{T}_y$ and by the L-Law, transforms to $\mathbb{T}_y^{x\vee z} = \mathbb{T}_y^{x\vee w}\mathbb{T}$. □

## 2.4 Canonical expressions

Let $\mathbb{T} \neq 1$ be an arbitrary coin. A coin identity of the form $\mathbb{T} = \mathbb{T}_{y_1}^{x_1} \cdots \mathbb{T}_{y_r}^{x_r}$ is called an expression of $\mathbb{T}$ with length $r$, where $\mathbb{T}_{y_1}^{x_1}, \ldots, \mathbb{T}_{y_r}^{x_r}$ are atom coins other than the unity. There are, however, infinitely many expressions which are equivalent to one another. For instance, $\mathbb{T} = \mathbb{T}\left\{(\mathbb{T}^x)^n (\mathbb{T}_x)^n\right\}$ holds for any $x \in \mathbb{L}$ and any $n \in \mathbb{N}$. Less trivially, suppose $z = x_1 \vee x_2 = y_1 \vee y_2 \vee y_3$. Then we can write $\mathbb{T}^z$ in many equivalent ways such as $\mathbb{T}^z = \mathbb{T}^{x_1}\mathbb{T}_{x_1}^{x_2} = \mathbb{T}^{x_2}\mathbb{T}_{x_2}^{x_1} = \mathbb{T}^{y_1}\mathbb{T}_{y_1}^{y_2}\mathbb{T}_{y_1 \vee y_2}^{y_3} = \mathbb{T}^{y_2}\mathbb{T}_{y_2}^{y_1}\mathbb{T}_{y_1 \vee y_2}^{y_3}$. The first two expressions have length 2, and the last two expressions have length 3. It will be useful to decompose a coin in a way similar to the prime decomposition of a natural number.

We say that two raising coins $\mathbb{T}^x$ and $\mathbb{T}^y$, with $x > \emptyset$ and $y > \emptyset$, are *mutually distinct*, if $x \neq y$. The following theorem says that a coin can be expressed using mutually distinct coins (see the Appendix for a proof).

**Theorem 1** *For an arbitrary coin $\mathbb{T} \neq 1$, there exist nonzero integers $n_i$ and mutually distinct coins $\mathbb{T}^{x_i}$, $i = 1, \ldots, r$ so that $\mathbb{T}$ can be expressed as $\mathbb{T} = (\mathbb{T}^{x_1})^{n_1} \cdots (\mathbb{T}^{x_r})^{n_r}$.*

**Definition 3** *(prime coin)* A raising coin $\mathbb{T}^x$ is said to be *prime* 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 $x_i < x$ for each $i = 1, \ldots, r$.

Theorem 1 implies that for any $x$, there exist nonzero integers $n_i$ and mutually distinct coins $\mathbb{T}^{x_i}$, $i = 1, \ldots, r$ so that $\mathbb{T}^x = (\mathbb{T}^{x_1})^{n_1} \cdots (\mathbb{T}^{x_r})^{n_r}$. If $\mathbb{T}^x$ is prime, then we must have $r = 1$, $n_1 = 1$ and $x_1 = x$. A coin $\mathbb{T}^x$ is not prime, for instance, if $\mathbb{T}^x = \mathbb{T}^y\mathbb{T}^z$, where $x = y \vee z$ and $y, z > \emptyset$.

A cainoid $\mathfrak{C}$ is said to be *free* if $\mathbb{T}^x$ is prime for any $x \in \mathbb{L}$. In a free cainoid $\mathfrak{C}$, there are no extra relations among coins of $\mathfrak{C}$ other than those derived from the axioms C1–C5.

From now on we shall impose the following *descending chain condition* (DCC) on the underlying lattice $\mathbb{L}$. A lattice $\mathbb{L}$, as an ordered set, is said to satisfy DCC if $\mathbb{L}$ contains no infinite descending chain $x_1 > x_2 > x_3 > \cdots$. DCC is equivalent to the condition that every nonempty subset $S \subset \mathbb{L}$ contains an element minimal in $S$. DCC is an important condition because ordered sets satisfying DCC are those for which the principle of induction holds (Nation, p.4). This condition also allows the join decomposition of an arbitrary element of a lattice. To state this latter fact more accurately, recall that an element $x \in \mathbb{L}$ is said to be *join irreducible* if $F$ is a finite set and $x = \vee F$ then we must have $x \in F$. It can be shown that if $\mathbb{L}$ satisfies DCC then each element of $\mathbb{L}$ is a join of finitely many join irreducible elements (Blyth 2005, p. 59). For a more systematic introduction to the theory of lattice see Blyth (2005) and Davey and Priestley (2002). Birkhoff (1967) gives more advanced treatment. See also Stanley and Sankappanavar (1981) and Jipsen and Rose (1992).

**Theorem 2** *Suppose that $\mathbb{L}$ satisfies DCC. There exist nonzero integers $n_1, \ldots, n_r$ so that every coin $\mathbb{T}$ has the following expression, unique up to reordering of terms:*

$$\mathbb{T} = (\mathbb{T}^{x_1})^{n_1} \cdots (\mathbb{T}^{x_r})^{n_r} \tag{8}$$

*where (i) $\mathbb{T}^{x_1}, \ldots, \mathbb{T}^{x_r}$ is prime; and (ii) $\mathbb{T}^{x_1}, \ldots, \mathbb{T}^{x_r}$ is mutually distinct.*

See the appendix for a proof.

**Definition 4** *(canonical expression)* The unique expression given by (8) is called the *canonical expression* of $\mathbb{T}$. And we call (i) $r$ the order of $\mathbb{T}$, written as $|\mathbb{T}| = r$; and (ii) $x = \vee_{i=1}^{r} x_i$ the *context* of $\mathbb{T}$, written as $\mathfrak{I}(\mathbb{T}) = x$.

The context $\mathfrak{I}$ thus defines a function from the set of coins $\mathfrak{C}$ to the lattice $\mathbb{L}$ with the following properties.

**Theorem 3** *Let* $\mathbb{T}, \mathbb{T}'$ *be any coins. Then the following hold: (i)* $\mathfrak{I}(\mathbb{T}) = \mathfrak{I}(\mathbb{T}^{-1})$*; (ii)* (sub-additivity) $\mathfrak{I}(\mathbb{T}\mathbb{T}') \leq \mathfrak{I}(\mathbb{T}) \vee \mathfrak{I}(\mathbb{T}')$.

*Proof* (i) Suppose that $\mathbb{T}$ has canonical expression $\mathbb{T} = (\mathbb{T}^{x_1})^{n_1} \cdots (\mathbb{T}^{x_r})^{n_r}$; then we have $\mathbb{T}^{-1} = (\mathbb{T}^{x_1})^{m_1} \cdots (\mathbb{T}^{x_r})^{m_r}$, where $m_i = -n_i \neq 0$. This expression for $\mathbb{T}^{-1}$ is canonical because $\mathbb{T}^{x_i}$ are prime and $\mathbb{T}^{x_i}$ and $\mathbb{T}^{x_j}$ are mutually prime for $i \neq j = 1, \ldots, r$. Hence $\mathfrak{I}(\mathbb{T}^{-1}) = \vee_{i=1}^{r} x_i = \mathfrak{I}(\mathbb{T})$.

(ii) Let $\mathbb{T} = (\mathbb{T}^{x_1})^{n_1} \cdots (\mathbb{T}^{x_r})^{n_r}$ and $\mathbb{T}' = (\mathbb{T}^{y_1})^{m_1} \cdots (\mathbb{T}^{y_s})^{m_s}$ be the canonical expressions of $\mathbb{T}$ and $\mathbb{T}'$, respectively. Let $z_1, \ldots, z_t$ be defined by

$$\{z_1, \ldots, z_t\} = \{x_1, \ldots, x_r, y_1, \ldots, y_s\}. \tag{9}$$

Note that $z_i$ equals either $x_j$ or $y_k$. Let $k_i = n_j$ if $z_i = x_j$; $k_i = m_k$ if $z_i = y_k$; and $k_i = n_j + m_k$ if $z_i = x_j = y_k$. It follows then $\mathbb{T}\mathbb{T}'$ can be expressed as $\mathbb{T}\mathbb{T}' = (\mathbb{T}^{z_1})^{k_1} \cdots (\mathbb{T}^{z_t})^{k_t}$ , where $z_i \neq z_j$ for $i \neq j = 1, \ldots, t$. Suppose that $k_1, \ldots, k_w (w \leq t)$ are none-zero and the remaining numbers are zero. Since $\mathbb{T}^{z_i}$ equals either $\mathbb{T}^{x_j}$ or $\mathbb{T}^{y_k}$, so $\mathbb{T}^{z_i}$ is prime. Thus $\mathbb{T}\mathbb{T}'$ has the following canonical expression $\mathbb{T}' = (\mathbb{T}^{z_1})^{k_1} \cdots (\mathbb{T}^{z_w})^{k_w}$. From (9) we then conclude that

$$\begin{aligned}
\mathfrak{I}(\mathbb{T}\hat{\mathbb{T}}) &= z_1 \vee \cdots \vee z_w \\
&\leq (z_1 \vee \cdots \vee z_w) \vee (z_{w+1} \vee \cdots \vee z_t) \\
&= (x_1 \vee \cdots \vee x_r) \vee (y_1 \vee \cdots \vee y_s) \\
&= \mathfrak{I}(\mathbb{T}) \vee \mathfrak{I}(\mathbb{T}')
\end{aligned}$$

which was what required to prove. This completes the proof. □

*Example 2* Suppose that $\mathfrak{C}$ is free. Suppose all the elements in the following are nontrivial. (i) Let $x = x_1 \vee x_2 \vee x_3$ and consider the following expression $\mathbb{T} = \mathbb{T}^{x_1}\mathbb{T}^{x_2}_{x_1}\mathbb{T}^{x_3}_{x_1 \vee x_2}$. Applying the R-law sequentially, we get an equivalent expression $\mathbb{T} = \mathbb{T}^{x}$. Since $\mathfrak{C}$ is free, $\mathbb{T} = \mathbb{T}^{x}$ is canonical, implying that $|\mathbb{T}| = 1$ and $\mathfrak{I}(\mathbb{T}) = x$.

(ii) Similarly, if $\mathbb{T} = \mathbb{T}^{x_1}\mathbb{T}^{x_2}_{x_1}$, then $\mathbb{T} = \mathbb{T}^{x_1 \vee x_2}$ is canonical. So $|\mathbb{T}| = 1$ and $\mathfrak{I}(\mathbb{T}) = x_1 \vee x_2$.

(iii) The mixed coin $\mathbb{T} = \mathbb{T}^{x_1}_{x_2}$, by C5, can be expressed as $\mathbb{T} = \mathbb{T}^{x_1 \vee x_2}\mathbb{T}_{x_2} = \mathbb{T}^{x_1 \vee x_2}(\mathbb{T}^{x_2})^{-1}$ where $\mathbb{T}^{x_1 \vee x_2}$ and $\mathbb{T}^{x_2}$ are prime and mutually distinct. So $|\mathbb{T}| = 2$ and $\mathfrak{I}(\mathbb{T}) = x_1 \vee x_2$. In general, raising and lowering coins have order 1, while mixed coins have order 2.

(iv) Let $\mathbb{T} = \mathbb{T}^{x_2}_{x_1}\mathbb{T}^{x_1}\mathbb{T}_{x_2}$. That $\mathbb{T} = \mathbb{T}^{x_1}_{x_2}$, by the Bayes' theorem, implying $|\mathbb{T}| = 2$.

### 2.5 Marginal cainoids

Let $x \in \mathbb{L}$ be nontrivial. Recall that if $\mathbb{L}$ satisfies DCC then there exist join irreducible $x_1, \ldots, x_n \in \mathbb{L}$ so that $x$ can be written as

$$x = x_1 \vee x_2 \vee \cdots \vee x_n. \tag{10}$$

However, (10) is by no means unique even if $\mathbb{L}$ is finite. (10) is called a *canonical join representation* of $x$ if (i) (10) is *irredundant*, that is, $x$ cannot be the join of a proper subset of $\{x_1, \ldots, x_n\}$, and (ii) if $A \subset \mathbb{L}$, then $x = \vee A$ implies $\{x_1, \ldots, x_n\} \ll A$, that is, for all $x_i$ we have $x_i \leq a$ for some $a \in A$. In this sense, a canonical join representation is the finest join decomposition. We shall write $\mathfrak{A}(x) = \{x_1, \ldots, x_n\}$ if (10) is canonical, which is necessarily unique (Freese et al. (1995, p. 36)). Jónsson and Kiefer (1962) show that if $\mathbb{L}$ is a finite lattice, then every element $x \in \mathbb{L}$ has a canonical join representation if and only if $\mathbb{L}$ is *join-semidistributive*: that is, $x \vee y = x \vee z$ implies $x \vee y = x \vee (y \wedge z)$ for any $x, y, z \in \mathbb{L}$. Note that finite Boolean lattices are join-semidistributive, where each element $x$ is canonically written as $x = \{x_1, \ldots, x_n\} = \{x_1\} \vee \cdots \vee \{x_n\}$.

Back to the general case, suppose that $x$ has a canonical join representation so that $\mathfrak{A}(x) = \{x_1, \ldots, x_n\}$. Recall that a sublattice $\mathbb{S}$ is a nonempty set of a lattice $\mathbb{L}$ so that for every pair of elements $a, b \in \mathbb{S}$, both $a \wedge b$ and $a \vee b$ are in $\mathbb{S}$. Expanding $\mathfrak{A}(x)$ by including all joins and meets of subsets of $\mathfrak{A}(x)$ we get a subset $\mathbb{L}_x$, which is a sublattice of $\mathbb{L}$ with respect to the partial order 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 we denote by $\mathfrak{C}_x$.

**Definition 5** *(marginal cainoid)* Let $x$ be an element of lattice $\mathbb{L}$ so that it has a (unique) canonical join representation. The cainoid $(\mathfrak{C}_x, \cdot)$ induced by the sublattice $\mathbb{L}_x$ is called the *marginal cainoid* (of $x$). An element of $\mathfrak{C}_x$ is called a *marginal coin* (of $x$).

**Notation 1** *We use $\mathbb{T}[x]$ to denote an arbitrary marginal coin of $x$.*

**Proposition 5** *$\mathfrak{C}_x$ is a subgroup of $\mathfrak{C}$.*

*Proof* We only need to show, for any $\mathbb{T}, \mathbb{T}' \in \mathfrak{C}_x$, that (i) $\mathbb{T}\mathbb{T}' \in \mathfrak{C}_x$, and (ii) $\mathbb{T}^{-1} \in \mathfrak{C}_x$. For (i), if $\mathbb{T}, \mathbb{T}' \in \mathfrak{C}_x$, then we have marginal expressions, $\mathbb{T} = \prod_{i=1}^{s} \mathbb{T}_{y_i}^{x_i}$ and $\mathbb{T}' = \prod_{j=1}^{t} \mathbb{T}_{v_j}^{u_j}$ where $\mathbb{T}_{y_i}^{x_i}$ and $\mathbb{T}_{v_j}^{u_j}$ are atom coins in $\mathfrak{C}_x$. It follows that $\mathbb{T}\mathbb{T}' = \prod_{i=1}^{s} \mathbb{T}_{y_i}^{x_i} \prod_{j=1}^{t} \mathbb{T}_{v_j}^{u_j}$ is also a product of a finite sequence of atom coins in $\mathfrak{C}_x$. So $\mathbb{T}\mathbb{T}' \in \mathfrak{C}_x$. For (ii), let $\mathbb{T}$ be expressed marginally as in (i). Since $(\mathbb{T}_{y_i}^{x_i})^{-1} = \mathbb{T}^{y_i}\mathbb{T}_{x_i \vee y_i}$ is also a marginal coin of $x$, it follows that $\mathbb{T}^{-1} = \prod_{i=1}^{s} (\mathbb{T}_{y_i}^{x_i})^{-1} = \prod_{i=1}^{s} \mathbb{T}^{y_i}\mathbb{T}_{x_i \vee y_i} \in \mathfrak{C}_x$. □

The marginal cainoid $\mathfrak{C}_x$, as a subgroup of $\mathfrak{C}$, introduces a natural *equivalence relation* among coins in $\mathfrak{C}$.

**Definition 6** *(equivalent coins)* Two coins $\mathbb{T}, \hat{\mathbb{T}} \in \mathfrak{C}$ are said equivalent with respect to $x$, written $\mathbb{T} \overset{x}{\sim} \hat{\mathbb{T}}$, if and only if $\mathbb{T}\hat{\mathbb{T}}^{-1} \in \mathfrak{C}_x$. That is, $\mathbb{T} \overset{x}{\sim} \hat{\mathbb{T}} \Leftrightarrow \mathbb{T} = \mathbb{T}[x]\hat{\mathbb{T}}$ holds for some marginal coin $\mathbb{T}[x]$ of $x$.

The set of all coins equivalent to $\mathbb{T}$ with respect to $x$ forms a *coset* of $\mathbb{T}$. 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 $\mathfrak{C}_x$.

## 3 The Cain

Now we are in a position to introduce further structure into a cainoid. We achieve this goal by considering an operation analogous to the integration of ordinary functions.

### 3.1 $x$-Integrability

**Notation 2** *We have already introduced the symbols* $\mathbb{T}^x$ *and* $\mathbb{T}[x] \in \mathfrak{C}_x$. *Now we denote by* $\mathbb{T}\{x\}$ *an arbitrary coin in* $\mathfrak{C}$ *with context* $\mathfrak{I}(\mathbb{T}\{x\}) = x$. *Note that* $\mathfrak{I}(\mathbb{T}^x) = \mathfrak{I}(\mathbb{T}\{x\}) = x$, *while* $\mathfrak{I}(\mathbb{T}[x]) \leq x$.

*Example 3* Consider a Boolean lattice generated by $D = \{1, 2, 3\}$. Let $x = \{1, 2\}$. Then we have (i) $\mathbb{T}^x = \mathbb{T}^{12}$; (ii) $\mathfrak{C}_x$ is an infinite set consisting of all marginal coins $\mathbb{T}[x]$ such as 1, $\mathbb{T}^1$, $(\mathbb{T}^1)^2$, $\mathbb{T}_2$, $\mathbb{T}_2^1$, $\mathbb{T}^{12}$, etc.; (iii) Coins $\mathbb{T}^1\mathbb{T}^2$, $\mathbb{T}_2^1\mathbb{T}_1^2$, $\mathbb{T}^{12}$, $\mathbb{T}^{12}\mathbb{T}_2$, $\mathbb{T}^1\mathbb{T}_1^{23}\mathbb{T}_{123}\mathbb{T}^{12}$ are examples of $\mathbb{T}\{x\}$; they all have context $x$.

**Definition 7** *($x$-integrability)* Let $x \in \mathbb{L}$. A coin $\mathbb{T}$ is said to be*$x$-integrable* if $\mathbb{T}$ has context no less than $x$, that is, $\mathbb{T} = \mathbb{T}\{y\}$ with $y \geq x$.

$\mathbb{T}^x$ is $x$-integrable for any $x$. If the top $\top$ exists, then $\top$-integrable coins are those with the largest context $\top$. On the other hand, all coins are $\emptyset$-integrable, where $\emptyset$ is the bottom. Since $\mathfrak{I}(\mathbb{T}) = \mathfrak{I}(\mathbb{T}^{-1})$ by Theorem 3, $\mathbb{T}$ is $x$-integrable if and only if $\mathbb{T}^{-1}$ is $x$-integrable.

Note that, when both $\mathbb{T}$ and $\mathbb{T}'$ are $x$-integrable, it does not follow that $\mathbb{T}\mathbb{T}'$ is $x$-integrable. For instance, let $x = x_1 \vee x_2$ irredundantly. 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 $\mathfrak{I}(\mathbb{T}\mathbb{T}') = x_2 < x$. So the set of $x$-integrable coins is not closed under dot product.

**Lemma 5** *If a lattice* $\mathbb{L}$ *has bottom* $\emptyset$, *then for any* $a, b, x \in \mathbb{L}$, *conditions* $a \vee b \geq x, b < x$ *imply* $a \wedge x > \emptyset$.

*Proof* If $a \wedge x = \emptyset$ holds, then $(a \vee b) \wedge x = (a \wedge x) \vee (b \wedge x) = \emptyset \vee b = b$. However, $a \vee b \geq x$ implies $(a \vee b) \wedge x = x > b$. The contradiction shows $a \wedge x \neq \emptyset$. Hence $a \wedge x > \emptyset$. □

**Lemma 6** *If* $\mathbb{T}$ *is $x$-integrable and* $y \wedge x = \emptyset$, *then* $\mathbb{T}[y]\mathbb{T}$ *is also $x$-integrable.*

*Proof* Let $\mathbb{T} = \mathbb{T}\{z\}$. Since $\mathbb{T}$ is $x$-integrable, $z \geq x$. Let $\mathbb{T}[y] = \mathbb{T}\{w\}$, where $w \leq y$. Note that $w \leq y$ implies $(w \wedge x) \wedge (y \wedge x) = w \wedge x$, or equivalently, $w \wedge x \leq y \wedge x$. Since $y \wedge x = \emptyset$ so $w \wedge x = \emptyset$. Now write $\mathbb{T}\{z\}$ and $\mathbb{T}\{w\}$ canonically as

$$\mathbb{T}\{z\} = (\mathbb{T}^{z_1})^{m_1} \cdots (\mathbb{T}^{z_r})^{m_r}, \quad \mathbb{T}\{w\} = (\mathbb{T}^{w_1})^{n_1} \cdots (\mathbb{T}^{w_s})^{n_s}$$

where $w = \vee_{i=1}^{s} w_i$ and $z = \vee_{i=1}^{r} z_i$. If no $w_i$ is equal to any $z_j$ with $m_i = -n_j$, then $\mathfrak{I}(\mathbb{T}\{z\}\mathbb{T}\{w\}) = z \vee w \geq x$, implying that $\mathbb{T}[y]\mathbb{T}$ is $x$-integrable.

Now suppose, without loss of generality, $w_i = z_i$, $m_i = -n_i$, $i = 1, \ldots, t$. Then the context (say $u$) of $\mathbb{T}\{z\}\mathbb{T}\{w\}$ equals $u = (\vee_{i=t+1}^{r} z_i) \vee (\vee_{i=t+1}^{s} w_i)$. We show $\vee_{i=t+1}^{r} z_i \geq x$, which implies $u \geq x$. If $\vee_{i=t+1}^{r} z_i < x$ holds, then by Lemma 5, $(\vee_{i=1}^{t} z_i) \wedge x > \emptyset$. But $w_i = z_i$ $(i = 1, \ldots, t)$, so $(\vee_{i=1}^{t} w_i) \wedge x > \emptyset$. It follows then $w \wedge x = (\tilde{w} \wedge x) \vee (\hat{w} \wedge x) > \emptyset$, where $\tilde{w} = \vee_{i=1}^{t} w_i$, $\hat{w} = \vee_{i=t+1}^{s} w_i$, contradicting the fact $w \wedge x = \emptyset$. This completes the proof. □

### 3.2 The cain

A bounded lattice $\mathbb{L}$, with a top $\top$ and a bottom $\emptyset$, is called a *complemented lattice* or *ortholattice*, if for each $x \in \mathbb{L}$ there exists a *complement*, denoted by $\bar{x}$, so that $x \vee \bar{x} = \top$, $x \wedge \bar{x} = \emptyset$. If $\mathbb{L}$ is also *distributive*, then $\bar{x}$ is also unique. From now on we shall assume that $\mathbb{L}$ is a complemented distributive lattice, i.e., a Boolean algebra. For details on Boolean algebra, see for instance, Salii (1988, Sect. 3).

**Definition 8** *(coin marginalization)* For an arbitrary $x \in \mathbb{L}$, let $\mathfrak{D}(x)$ be the set of all $x$-integrable coins. The *$x$-marginalization* 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

$$\int_x \mathbb{T}\{y\} = \mathbb{T}\{y \wedge \bar{x}\}. \tag{11}$$

Further, the function $\int_x$ satisfies the following three axiomatic properties:

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

$$\int_x \mathbb{T}^y = \mathbb{T}^{y \wedge \bar{x}}. \tag{12}$$

(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

$$\int_{x_1 \vee x_2} \mathbb{T}\{y_1\}\mathbb{T}\{y_2\} = \int_{x_1} \mathbb{T}\{y_1\} \int_{x_2} \mathbb{T}\{y_2\}. \tag{13}$$

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

$$\int_{\emptyset} \mathbb{T} = \mathbb{T}. \tag{14}$$

*Remark 2* $\int_x \mathbb{T}$ is defined for any $x \leq \mathfrak{I}(\mathbb{T})$.

**Notation 3** *To mimic the conventional notation of integration, we shall write* $\int_x \mathbb{T}$ *by* $\int \mathbb{T}\, dx$. *With this notation, axioms* (12)–(14) *can be rewritten as*

$$\int \mathbb{T}^y \, \mathrm{d}x = \mathbb{T}^{y \wedge \bar{x}} \tag{15}$$

$$\int \left(\mathbb{T}\{y_1\}\mathbb{T}\{y_2\}\right) \, \mathrm{d}(x_1 \vee x_2) = \int \mathbb{T}\{y_1\}\, \mathrm{d}x_1 \int \mathbb{T}\{y_2\}\, \mathrm{d}x_2 \tag{16}$$

$$\int \mathbb{T} \, \mathrm{d}\emptyset = \mathbb{T}. \tag{17}$$

(15) 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 (15) implies $\int \mathbb{T}^y \, \mathrm{d}\emptyset = \mathbb{T}^y$, a special case of (17). (16) is an analog of the following property for conventional integration $\int f(x, z)g(y, z)\, \mathrm{d}x\, \mathrm{d}y = \int f(x, z)\, \mathrm{d}x \int g(y, z)\, \mathrm{d}y$. The ordinary integration also satisfies $\int cf(x, y)\, \mathrm{d}x = c \int f(x, y)\, \mathrm{d}x$, where $c$ is a constant independent of both $x$ and $y$. (17) is an abstraction of this property.

**Definition 9** *(cain)* A cain is a cainoid further satisfying (15)–(17).

### 3.3 Properties of the cain

Now we give some basic properties on the coin marginalization. Using (16)–(17), we have

**Theorem 4** *If* $\mathbb{T}$ *is* $x$*-integrable and* $x \wedge y = \emptyset$, *then* $\int \mathbb{T}[y]\mathbb{T}\, \mathrm{d}x = \mathbb{T}[y] \int \mathbb{T}\, \mathrm{d}x$.

*Proof* By Lemma 6, $\mathbb{T}[y]\mathbb{T}$ is $x$-integrable. So

$$\begin{aligned}
\int \mathbb{T}[y]\mathbb{T}\, \mathrm{d}x &= \int \mathbb{T}[y]\mathbb{T}\, \mathrm{d}(x \vee \emptyset) \\
&= \int \mathbb{T}[y]\, \mathrm{d}\emptyset \int \mathbb{T}\, \mathrm{d}x \quad (\text{by } (16)) \\
&= \mathbb{T}[y] \int \mathbb{T}\, \mathrm{d}x \quad (\text{by } (17)).
\end{aligned}$$

□

**Theorem 5** *For any* $x, y, z \in \mathbb{L}$, $\mathbb{T}^x_y = \mathbb{T}^x_z \Rightarrow \mathbb{T}^x_y = \mathbb{T}^x_{y \wedge z}$.

*Proof* If $x > \emptyset$, by the R-Law, then $\mathbb{T}^x_y = \mathbb{T}^x_z \Rightarrow \mathbb{T}^{x \vee y} = \mathbb{T}^x_z \mathbb{T}^y$. Let $y_1 = y \wedge \bar{x} \wedge \bar{z}$; then

$$\mathbb{T}^{(x \vee y) \wedge \bar{y}_1} = \int \mathbb{T}^{x \vee y}\, \mathrm{d}y_1 = \int \mathbb{T}^x_z \mathbb{T}^y\, \mathrm{d}y_1 = \mathbb{T}^x_z \mathbb{T}^{y \wedge \bar{y}_1}$$

the first and third equality being due to (15) and Theorem 4, respectively. Since $(x \vee y) \wedge \bar{y}_1 = x \vee (y \wedge z)$, so $\mathbb{T}^{x \vee (y \wedge z)} = \mathbb{T}^x_z \mathbb{T}^{y \wedge \bar{y}_1}$. Symmetrical arguments for $z$ will lead to $\mathbb{T}^{x\vee(y\wedge z)} = \mathbb{T}_y^x\mathbb{T}^{z\wedge\bar{z}_1}$, where $z_1 = z \wedge \bar{x} \wedge \bar{y}$. These two equations jointly imply $\mathbb{T}^{y\wedge\bar{y}_1} = \mathbb{T}^{z\wedge\bar{z}_1}$, which in turn implies $y \wedge \bar{y}_1 = z \wedge \bar{z}_1 = (y \wedge \bar{y}_1) \wedge (z \wedge \bar{z}_1) = y \wedge z$. So we conclude that $\mathbb{T}^{x\vee(y\wedge z)} = \mathbb{T}_y^x\mathbb{T}^{y\wedge z}$, which by the L-Law, gives the required result.

The case for $x = \emptyset$ is obvious because $\mathbb{T}_y^x = \mathbb{T}_z^x$ reduced to $\mathbb{T}_y = \mathbb{T}_z$, which implies $y = z$. □

**Lemma 7** *If $\mathbb{L}$ is complemented and distributive, then $x \wedge y = \emptyset \Leftrightarrow y \leq \bar{x}$.*

*Proof* Sufficiency. If $y \leq \bar{x}$, then $y = \bar{x} \wedge y$, so $y \wedge x = (\bar{x} \wedge y) \wedge x = \emptyset$.

Necessity: If $x \wedge y = \emptyset$, then $\overline{x \wedge y} = \bar{x} \vee \bar{y} = \top$, where $\top$ is the top of $\mathbb{L}$. It follows then $y = \top \wedge y = (\bar{x} \vee \bar{y}) \wedge y = \bar{x} \wedge y$, implying $y \leq \bar{x}$. □

**Theorem 6** *If $x \geq z$, $y \wedge z = \emptyset$ then $\int \mathbb{T}_y^x\, dz = \mathbb{T}_y^{x\wedge\bar{z}}$.*

*Proof* If $x = \emptyset$, then $z = \emptyset$ and $x \wedge \bar{z} = x$, so the equation holds by (17).

If $x > \emptyset$, then $\mathbb{T}_y^x = \mathbb{T}^{x\vee y}\mathbb{T}_y$, which is $z$-integrable. By Theorem 4 and (15), $\int \mathbb{T}_y^x\, \mathrm{d}z = \int \mathbb{T}^{x\vee y}\mathbb{T}_y\, \mathrm{d}z = \mathbb{T}_y\mathbb{T}^{(x\vee y)\wedge\bar{z}}$. By Lemma 7, $y \wedge z = \emptyset \Rightarrow y \leq \bar{z}$. It follows then $(x \wedge \bar{z}) \vee y = (x \vee y) \wedge \bar{z}$, which implies $\mathbb{T}_y^{x\wedge\bar{z}} = \mathbb{T}_y\mathbb{T}^{(x\vee y)\wedge\bar{z}}$. □

The following theorem is an analog of the fact that (conditional) probability density functions are normalized functions.

**Theorem 7** *(i) If $x \in \mathbb{L}$, then $\int \mathbb{T}^x\, \mathrm{d}x = 1$; (ii) If $x, y \in \mathbb{L}$ and $x > \emptyset$, then $\int \mathbb{T}_y^x\, \mathrm{d}(x \wedge \bar{y}) = 1$.*

*Proof* For (i), using (15), we have $\int \mathbb{T}^x\, \mathrm{d}x = \mathbb{T}^{x\wedge\bar{x}} = \mathbb{T}^{\emptyset} = 1$.

For (ii), let $z = x \wedge \bar{y}$, then $x \vee y = z \vee y$ and $z \wedge y = \emptyset$. Since $x > \emptyset$, by C5, $\mathbb{T}_y^x = \mathbb{T}^{x\vee y}\mathbb{T}_y = \mathbb{T}^{z\vee y}\mathbb{T}_y$, which is $z$-integrable. So

$$
\begin{aligned}
\int \mathbb{T}_y^x\, \mathrm{d}z &= \int \mathbb{T}^{z\vee y}\mathbb{T}_y\, \mathrm{d}z \\
&= \mathbb{T}_y \int \mathbb{T}^{z\vee y}\, \mathrm{d}z \quad \text{(by Theorem 4)} \\
&= \mathbb{T}_y\mathbb{T}^{(z\vee y)\wedge\bar{z}} \quad \text{(by (15))} \\
&= \mathbb{T}_y\mathbb{T}^{y} \quad \text{(by Lemma 7)} \\
&= 1.
\end{aligned}
$$

□

The following formulae are useful for transforming coin equations:

**Corollary 1** *If $x, y, z, w$ are mutually exclusive, then (i) $\int \mathbb{T}[x]\mathbb{T}^y\, \mathrm{d}y = \mathbb{T}[x]$; (ii) $\int \mathbb{T}[x]\mathbb{T}^{y\vee z}\, \mathrm{d}z = \mathbb{T}[x]\mathbb{T}^y$; (iii) $\int \mathbb{T}[x \vee w]\, \mathbb{T}^{x\vee y\vee z}\, \mathrm{d}z = \mathbb{T}[x \vee w]\mathbb{T}^{x\vee y}$; and (iv) $\int \mathbb{T}_z^{x\vee y}\, \mathrm{d}y = \mathbb{T}_z^x$, where $x > \emptyset$.*

*Proof* For (i), since $x \wedge y = \emptyset$ so $\mathbb{T}[x]\mathbb{T}^y$ is $y$-integrable. By Theorems 4 and 7 we then have

$$\int \mathbb{T}[x]\mathbb{T}^y \,\mathrm{d}y = \mathbb{T}[x] \int \mathbb{T}^y \,\mathrm{d}y = \mathbb{T}[x].$$

For (ii), using Theorem 4 and Lemma 7, we have

$$\int \mathbb{T}[x]\mathbb{T}^{y\vee z} \,\mathrm{d}z = \mathbb{T}[x] \int \mathbb{T}^{y\vee z} \,\mathrm{d}z = \mathbb{T}[x]\mathbb{T}^{(x\vee z)\wedge \bar{z}} = \mathbb{T}[x]\mathbb{T}^{x}.$$

Similarly, (iii) follows from

$$\int \mathbb{T}[x\vee w]\,\mathbb{T}^{x\vee y\vee z} \,\mathrm{d}z = \mathbb{T}[x\vee w] \int \mathbb{T}^{x\vee y\vee z} \,\mathrm{d}z = \mathbb{T}[x\vee w]\mathbb{T}^{x\vee y}.$$

Finally, (iv) follows from

$$\int \mathbb{T}_z^{x\vee y} \,\mathrm{d}y = \int \mathbb{T}_z\mathbb{T}^{x\vee y\vee z} \,\mathrm{d}y = \mathbb{T}_z \int \mathbb{T}^{x\vee y\vee z} \,\mathrm{d}y = \mathbb{T}_z\mathbb{T}^{x\vee z} = \mathbb{T}_z^{x}.$$

□

### 3.4 The N-law and the M-law

Now we discuss two important rules concerning certain types of coin equations. These rules will play important roles in transforming coin equations. Since conditional independence will be defined in terms of coin identities, these rules are thus of primary importance for manipulating conditional independence relations.

**Theorem 8** *(law of normalization) Let $\bar{x}$, $\bar{y}$ be the complements of $x$ and $y$, respectively. Then*

$$\mathbb{T}^x_{y\vee z} = \mathbb{T}[\bar{z}] \Rightarrow \mathbb{T}^x_{y\vee z} = \mathbb{T}^{x\wedge\bar{z}}_{y\wedge\bar{z}} \quad (x > \emptyset) \tag{18}$$

$$\mathbb{T}^{x\vee y\vee z} = \mathbb{T}[\bar{y}]\mathbb{T}[\bar{x}] \Rightarrow \mathbb{T}^{x\vee y\vee z} = \mathbb{T}^{(x\vee z)\wedge\bar{y}}\mathbb{T}^{(y\vee z)\wedge\bar{x}}\mathbb{T}_{z\wedge\bar{x}\wedge\bar{y}}. \tag{19}$$

We use the name *normalization* because (18)–(19) depend essentially on (17). See the Appendix for a proof. In particular, when $x, y, z$ are nontrivial and mutually exclusive we have

$$\mathbb{T}^x_{y\vee z} = \mathbb{T}[\bar{z}] \Rightarrow \mathbb{T}^x_{y\vee z} = \mathbb{T}^x_y \quad (x > \emptyset) \tag{20}$$

$$\mathbb{T}^{x\vee y\vee z} = \mathbb{T}[\bar{y}]\mathbb{T}[\bar{x}] \Rightarrow \mathbb{T}^{x\vee y\vee z} = \mathbb{T}^{x\vee z}\mathbb{T}^{y\vee z}\mathbb{T}_z. \tag{21}$$

As special cases of (18)–(19), we also have, for any $x$, $y$, that

$$\mathbb{T}^x_y = \mathbb{T}[\bar{y}] \Rightarrow \mathbb{T}^x_y = \mathbb{T}^{x\wedge\bar{y}} \quad (x > \emptyset) \tag{22}$$

$$\mathbb{T}^{x\vee y} = \mathbb{T}[\bar{y}]\mathbb{T}[\bar{x}] \Rightarrow \mathbb{T}^{x\vee y} = \mathbb{T}^{x\wedge\bar{y}}\mathbb{T}^{y\wedge\bar{x}}. \tag{23}$$

In particular, when $x$, $y$ are nontrivial and mutually exclusive, we have

$$\mathbb{T}^x_y = \mathbb{T}[\bar{y}] \Rightarrow \mathbb{T}^x_y = \mathbb{T}^x \quad \text{and} \quad \mathbb{T}^{x\vee y} = \mathbb{T}[\bar{y}]\mathbb{T}[\bar{x}] \Rightarrow \mathbb{T}^{x\vee y} = \mathbb{T}^x\mathbb{T}^y. \tag{24}$$

The *law of normalization*, or the *N-law* for short, is a powerful rule that enables one to coerce an 'ambiguous' coin equation into 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 on a small portion of them. Those 'nuisance' coins can be treated as 'proportionality' constant.

On the other hand, a 'large' coin identity can give rise to many 'small' identities using the following *law of marginalization*, or the *M-law*; see the Appendix for a proof.

**Theorem 9** *(law of marginalization) If* $x \wedge y = \emptyset$*, then for any* $a, b \in \mathbb{L}$ *with* $x \wedge a > \emptyset$, $y \wedge b > \emptyset$*, the following holds true:*

$$\mathbb{T}^{x\vee y}_z = \mathbb{T}^x_z\mathbb{T}^y_z \Rightarrow \mathbb{T}^{(x\wedge a)\vee(y\wedge b)}_z = \mathbb{T}^{x\wedge a}_z\mathbb{T}^{y\wedge b}_z. \tag{25}$$

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

The following rule is also useful in marginalizing a coin equation:

**Theorem 10** *If* $x$, $y$, $z$ *are mutually exclusive then* $\mathbb{T}^{x\vee y\vee z} = \mathbb{T}[\bar{z}]\mathbb{T}^{y\vee z} \Rightarrow \mathbb{T}^{x\vee y} = \mathbb{T}[\bar{z}]\mathbb{T}^y$.

*Proof* Integrating $\mathbb{T}^{x\vee y\vee z}$ with respect to $z$, we have

$$\begin{aligned}\mathbb{T}^{x\vee y} &= \int \mathbb{T}^{x\vee y\vee z}\,\mathrm{d}z \\ &= \int \mathbb{T}[\bar{z}]\mathbb{T}^{y\vee z}\,\mathrm{d}z \\ &= \mathbb{T}[\bar{z}]\int \mathbb{T}^{y\vee z}\,\mathrm{d}z \\ &= \mathbb{T}[\bar{z}]\mathbb{T}^y.\end{aligned}$$

We leave to the reader to justify each step in the above. □

## 4 Conditional independence

### 4.1 Definition and basic properties

Let $\mathbb{L}$ be a Boolean algebra. That is, $\mathbb{L}$ is a complemented distributive lattice with bottom $\emptyset$ and top $\top > \emptyset$. Let $\mathfrak{C}$ be a cain defined on $\mathbb{L}$. Recall that for random variables, $x$ is said to be independent of $y$ given $z$ if $f(x|y,z) = f(x|z)$, where $f(\cdot|\cdot)$ denotes a conditional density function (Dawid 1979a). Similarly, we give the following definition:

**Definition 10** *(conditional independence)* $x$ is said to be independent of $y$ conditional on $z$, written $x \perp\!\!\!\perp y|z$, if and only if $\mathbb{T}^x_{y\vee z} = \mathbb{T}^x_z$.

If $x \perp\!\!\!\perp y|\emptyset$ we say that $x$ is independent of $y$, written $x \perp\!\!\!\perp y$.

The R-, and L-laws immediately lead to the following sufficient and necessary conditions:

**Theorem 11** *If $x, y, z > \emptyset$, then $x \perp\!\!\!\perp y|z$ holds if and only if $\mathbb{T}^{x\vee y}_z = \mathbb{T}^x_z \mathbb{T}^y_z$ or $\mathbb{T}^{x\vee y\vee z} = \mathbb{T}^{y\vee z}\mathbb{T}^x_z$ holds true.*

Note that the conditions (A) $\mathbb{T}^x_{y\vee z} = \mathbb{T}^x_z$, (B) $\mathbb{T}^{x\vee y}_z = \mathbb{T}^x_z \mathbb{T}^y_z$ and (C) $\mathbb{T}^{x\vee y\vee z} = \mathbb{T}^{y\vee z}\mathbb{T}^x_z$ are not equivalent to each other if the assumption $x, y, z > \emptyset$ is dropped out. (i) First, (A) implies (B) in all cases except when $y = \emptyset, z > \emptyset$; and (B) implies (A) unless $y > x = \emptyset$. (ii) Next, while (B) implies (C) at all cases, (C) implies (B) unless $z > y = \emptyset$. (iii) Finally, (C) implies (A) except when $y > z = x = \emptyset$; and (A) implies (C) unless $z > x = \emptyset$.

*Remark 3* $x \perp\!\!\!\perp y|z$ is symmetric for $x$ and $y$, unless $\emptyset = y < x \not\leq z$.

**Theorem 12** *If $x, y > \emptyset$, then $x \perp\!\!\!\perp y$ if and only if $\mathbb{T}^{x\vee y} = \mathbb{T}^x\mathbb{T}^y$.*

Notice that (D) $\mathbb{T}^x_y = \mathbb{T}^x$ implies (E) $\mathbb{T}^{x\vee y} = \mathbb{T}^x\mathbb{T}^y$ at all cases, (E) does not imply (D) when $y > x = \emptyset$. The following properties follow immediately from the definition of conditional independence:

**Proposition 6** *For any $x, y$, it holds that $x \perp\!\!\!\perp x|y \Leftrightarrow x \leq y$.*

When $x = \emptyset$, we have $y \geq x$, For $x > \emptyset$, this is true because $x \perp\!\!\!\perp x|y \Leftrightarrow \mathbb{T}^x_y = \mathbb{T}^{x\vee x}_y = \mathbb{T}^x_y\mathbb{T}^x_y \Leftrightarrow \mathbb{T}^x_y = 1 \Leftrightarrow x \leq y$ by Proposition 1. This property says that an inequality in $\mathbb{L}$ can be represented by a CI relation in $\mathfrak{C}$.

**Proposition 7** *(i) For any $x, y, z$, if $y \leq z$, then $x \perp\!\!\!\perp y|z$; (ii) If $x > \emptyset$, then $x \leq z \Rightarrow x \perp\!\!\!\perp y|z$.*

For instance, $x \perp\!\!\!\perp \emptyset|z$, $x \perp\!\!\!\perp y|\top$ and $x \perp\!\!\!\perp y|y$ hold for any $x, y$. If $x > \emptyset$, then $x \perp\!\!\!\perp y|x$.

**Proposition 8** *When $x, y, z$ are nontrivial and mutually exclusive, $x \perp\!\!\!\perp y|z$ holds if and only if $\mathbb{T}_z^{x\vee y} \overset{y\vee z}{\sim} \mathbb{T}_z^x$, or $\mathbb{T}_z^{x\vee y} \overset{x\vee z}{\sim} \mathbb{T}_z^y$.*

*Proof* "⇒". $x \perp\!\!\!\perp y|z \Rightarrow \mathbb{T}_z^{x\vee y} = \mathbb{T}_z^x \mathbb{T}_z^y \in \mathbb{T}_z^x \mathfrak{C}_{y\vee z} \Rightarrow \mathbb{T}_z^{x\vee y} \overset{y\vee z}{\sim} \mathbb{T}_z^x$.

"⇐". $\mathbb{T}_z^{x\vee y} \overset{y\vee z}{\sim} \mathbb{T}_z^x \Rightarrow \mathbb{T}_z^{x\vee y} = \mathbb{T}_z^x \mathbb{T}[y \vee z]$ for some $\mathbb{T}[y \vee z] \in \mathfrak{C}_{y\vee z}$, which implies $\mathbb{T}_z^{x\vee y} = \mathbb{T}_z^x \mathbb{T}_z^y$ by the N-Law (21). So $x \perp\!\!\!\perp y|z$. □

By the above proposition the mixed coin $\mathbb{T}_z^{x\vee y}$ is a representative element belonging to the quotient groups $\mathfrak{C}/\mathfrak{C}_{y\vee z}$ and $\mathfrak{C}/\mathfrak{C}_{x\vee z}$.

The N-law leads to the following seemingly weaker but equivalent condition for conditional independence:

**Theorem 13** *(factorization) If $x, y, z$ are nontrivial and mutually exclusive, then*

$$x \perp\!\!\!\perp y|z \Leftrightarrow \mathbb{T}^{x\vee y\vee z} = \mathbb{T}[\bar{x}]\mathbb{T}[\bar{y}]$$

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

*Proof* If $x \perp\!\!\!\perp y|z$, then $\mathbb{T}^{x\vee y\vee z} = \mathbb{T}^{y\vee z}\mathbb{T}_z^x = \mathbb{T}[\bar{x}]\mathbb{T}[\bar{y}]$, the first equality being due to Theorem 11 and the second equality due to Lemma 7.

Conversely, by (21), we have $\mathbb{T}^{x\vee y\vee z} = \mathbb{T}^{x\vee z}\mathbb{T}^{y\vee z}\mathbb{T}_z = \mathbb{T}^{y\vee z}\mathbb{T}_z^x$. So $x \perp\!\!\!\perp y|z$ by Theorem 11. □

In the factorization Theorem if we let $z = \emptyset$, then we see that $x \perp\!\!\!\perp y$ holds if and only if $\mathbb{T}_y^x = \mathbb{T}[x]$ or $\mathbb{T}_x^y = \mathbb{T}[y]$, where $x, y$ are nontrivial and mutually exclusive.

*Example 4* Let $D = \{1, 2, 3\}$ and $\mathbb{L} = 2^D$. Any one of the following coin equations, $\mathbb{T}_3^{12} = \mathbb{T}_3^1\mathbb{T}_3^2$, $\mathbb{T}^{123} = \mathbb{T}^{13}\mathbb{T}_3^2$, $\mathbb{T}^{123} = \mathbb{T}^{23}\mathbb{T}_3^1$, $\mathbb{T}_{23}^1 = \mathbb{T}_3^1$, $\mathbb{T}_{13}^2 = \mathbb{T}_3^2$, gives a necessary and sufficient condition for $1 \perp\!\!\!\perp 2|3$.

## 4.2 The graphoid

Now we discuss a set of properties on conditional independence as defined in the previous section. The analogs of these properties for the classical PCI relations play important roles in statistical graphical models (Lauritzen 1996). The following property follows immediately from the M-law (25):

**Theorem 14** *(decomposition) If $x \wedge y = \emptyset$, then $x \perp\!\!\!\perp y|z \Rightarrow a \perp\!\!\!\perp b|z$ holds for any $a \leq x, b \leq y$.*

**Theorem 15** *(weak union) If $x, y, z, w$ are nontrivial and mutually exclusive, then $x \perp\!\!\!\perp (y \vee z)|w \Rightarrow x \perp\!\!\!\perp y|(z \vee w)$.*

*Proof* First, $x \perp\!\!\!\perp (y \vee z)|w$ implies $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{w}\mathbb{T}^{y\vee z\vee w}$. By the M-law, we then have $\mathbb{T}^{x}_{z\vee w} = \mathbb{T}^{x}_{w}$. These equations jointly imply $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{z\vee w}\mathbb{T}^{y\vee z\vee w}$, so $x \perp\!\!\!\perp y|(z \vee w)$. □

The following result is sometimes referred to as the *contraction* property of PCI (Pearl 2000, p.11).

**Theorem 16** *(contraction) If $x, y, z, w$ are nontrivial and mutually exclusive, then $x \perp\!\!\!\perp y|(z \vee w)$ and $x \perp\!\!\!\perp z|w$ hold if and only if $x \perp\!\!\!\perp (y \vee z)|w$.*

*Proof* "sufficiency." First, $x \perp\!\!\!\perp (y\vee z)|w$ implies, by the M-law, $x \perp\!\!\!\perp z|w$, which in turn implies $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{w}\mathbb{T}^{y\vee z\vee w} = \mathbb{T}^{x}_{z\vee w}\mathbb{T}^{y\vee z\vee w}$. So $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{z\vee w}\mathbb{T}^{y\vee z\vee w}$, showing $x \perp\!\!\!\perp y|(z \vee w)$.

"necessity". $x \perp\!\!\!\perp y|(z \vee w)$ and $x \perp\!\!\!\perp z|w$ imply $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x\vee z\vee w}\mathbb{T}^{y}_{z\vee w}$ and $\mathbb{T}^{x\vee z\vee w} = \mathbb{T}^{x\vee w}\mathbb{T}^{z}_{w}$. So $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x\vee w}\mathbb{T}^{z}_{w}\mathbb{T}^{y}_{z\vee w}$, which by M-law implies $\mathbb{T}^{y\vee z\vee w} = \mathbb{T}^{w}\mathbb{T}^{z}_{w}\mathbb{T}^{y}_{z\vee w}$, or equivalently, $\mathbb{T}^{y\vee z}_{w} = \mathbb{T}^{z}_{w}\mathbb{T}^{y}_{z\vee w}$. So $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x\vee w}\mathbb{T}^{y\vee z}_{w}$, showing $x \perp\!\!\!\perp (y \vee z)|w$. □

**Corollary 2** *If $x, y, z$ are nontrivial and mutually exclusive, then $x \perp\!\!\!\perp y|z$ and $x \perp\!\!\!\perp z$ hold if and only if $x \perp\!\!\!\perp (y \vee z)$.*

**Theorem 17** *(intersection) If $x, y, z, w$ are nontrivial and mutually exclusive, then $x \perp\!\!\!\perp y|(z \vee w)$ and $x \perp\!\!\!\perp z|(y \vee w)$ hold if and only if $x \perp\!\!\!\perp (y \vee z)|w$.*

*Proof* "necessity". The facts $x \perp\!\!\!\perp y|(z \vee w) \Rightarrow \mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{z\vee w}\mathbb{T}^{y\vee z\vee w}$ and $x \perp\!\!\!\perp z|(y \vee w) \Rightarrow \mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{y\vee w}\mathbb{T}^{y\vee z\vee w}$, jointly imply $\mathbb{T}^{x}_{z\vee w} = \mathbb{T}^{x}_{y\vee w}$. Since $\mathbb{T}^{x}_{y\vee w}$ is of the form $\mathbb{T}[\bar{z}]$, so by the N-law, we have $\mathbb{T}^{x}_{z\vee w} = \mathbb{T}^{x}_{w}$, which in turn implies that $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{w}\mathbb{T}^{y\vee z\vee w}$, or $x \perp\!\!\!\perp (y \vee z)|w$.

"sufficiency". By the M-law, $x \perp\!\!\!\perp (y \vee z)|w$ implies $x \perp\!\!\!\perp y|w$ and $x \perp\!\!\!\perp z|w$, or in coin equations, $\mathbb{T}^{x}_{y\vee w} = \mathbb{T}^{x}_{w}$ and $\mathbb{T}^{x}_{z\vee w} = \mathbb{T}^{x}_{w}$, respectively. But, $x \perp\!\!\!\perp (y \vee z)|w \Leftrightarrow \mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{w}\mathbb{T}^{y\vee z\vee w}$, implying $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{y\vee w}\mathbb{T}^{y\vee z\vee w}$ and $\mathbb{T}^{x\vee y\vee z\vee w} = \mathbb{T}^{x}_{z\vee w}\mathbb{T}^{y\vee z\vee w}$, or $x \perp\!\!\!\perp z|(y \vee w)$ and $x \perp\!\!\!\perp y|(z \vee w)$. □

**Corollary 3** *If $x, y, z$ are nontrivial and mutually exclusive, then $x \perp\!\!\!\perp y|z$ and $x \perp\!\!\!\perp z|y$ hold if and only if $x \perp\!\!\!\perp (y \vee z)$.*

**Definition 11** 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

| | |
|---|---|
| $G1 : x \amalg y|z \Rightarrow y \amalg x|z$ | (Symmetry) |
| $G2 : x \amalg (y \vee w)|z \Rightarrow x \amalg y|z$ | (Decomposition) |
| $G3 : x \amalg (y \vee z)|w \Rightarrow x \amalg y|(z \vee w)$ | (Weak union) |
| $G4 : x \amalg y|(z \vee w), x \amalg z|w \Rightarrow x \amalg (y \vee z)|w$ | (Contraction) |
| $G5 : x \amalg y|(z \vee w), x \amalg z|(y \vee w) \Rightarrow x \amalg (y \vee z)|w$ | (Intersection). |

In the finitary case, properties G1–G5 were discussed by Dawid (1979a) and Spohn (1980). The name of 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 (2000). Theorems 14–17 prove that

**Theorem 18** *We say that a ternary relation* $x \perp\!\!\!\perp y|z$ *holds for nontrivial elements* $x, y, z$ *in* $\mathbb{L}$ *if the coin identity* $\mathbb{T}_z^{x \vee y} = \mathbb{T}_z^x \mathbb{T}_z^y$ *holds in* $\mathfrak{C}$. *If* $\mathbb{L}$ *is a Boolean algebra, then* $\cdot \perp\!\!\!\perp \cdot | \cdot$ *is a graphoid.*

### 4.3 The separoid

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

**Definition 12** *(separoid)* Let $(S, \leq)$ be a join-semilattice. Let $\cdot \perp \cdot | \cdot$ be a ternary relation on $S$. Then $(S, \leq, \perp)$ is a *separoid* if the following hold:

$$\begin{array}{ll}
\text{P1: } x \perp y|x & \\
\text{P2: } x \perp y|z & \Rightarrow y \perp x|z \\
\text{P3: } x \perp y|z \ \&\ w \leq y & \Rightarrow x \perp w|z \\
\text{P4: } x \perp y|z \ \&\ w \leq y & \Rightarrow x \perp y|(z \vee w) \\
\text{P5: } x \perp y|z \ \&\ x \perp w|(y \vee z) & \Rightarrow x \perp (y \vee w)|z.
\end{array}$$

A separoid $(S, \leq, \perp)$ is said to be *strong* if $(S, \leq)$ is a lattice and the following additional property holds

$$\begin{array}{ll}
\text{P6: If} z \leq y \ \&\ w \leq y \quad \text{then} & \\
\quad x \perp y|z \ \&\ x \perp y|w & \Longrightarrow x \perp y|(z \wedge w).
\end{array}$$

Now we show that P1–P6 are satisfied by a cain conditional independence relation (see the Appendix for a proof).

**Theorem 19** *Let* $\mathfrak{C}$ *be a cain of a Boolean algebra* $(\mathbb{L}, \leq)$. *We say that a ternary relation* $x \perp\!\!\!\perp y|z$ *holds for nontrivial elements* $x, y, z$ *in* $\mathbb{L}$ *if the coin identity* $\mathbb{T}_z^{x \vee y} = \mathbb{T}_z^x \mathbb{T}_z^y$ *holds in* $\mathfrak{C}$. *Then* $(\mathbb{L}, \leq, \perp\!\!\!\perp)$ *is a strong separoid.*

## 5 Discussions

Inspired by the work of Dawid (2001) we have developed the theory of cain based on a possibly infinite lattice, although for usual statistical applications a finite Boolean lattice suffices. Although we did not explore graphical implications in this paper, Andersson and Perlman (1993) does give a useful definition of PCI, the lattice conditional independence, based on finite distributive lattices. Andersson et al. (1997) show that the lattice conditional independence models coincide with DAG models induced by transitive acyclic directed graphs. For further developments of Andersson and Perlman's lattice conditional independence, see Andersson et al. (1995), Andersson and Madsen (1998), Andersson and Perlman (1995a), Andersson and Perlman (1995b), and Massam and Neher (1998).

The algebraic framework proposed in this paper gives a new look at statistical models which reply on PCI relations. The algebraic approach opens possibilities for 'automatic' derivation and discovery of PCI relations using tools similar to the Gröbner basis theories (Cox et al. 1997). A coin is an algebraic analog of the classic joint probability density function (PDF). The relevant operations for PCI are products of the PDFs and integrations of the PDF (to get marginal density functions). This is the major motivation for only considering coin product in the cainoid. However, there is a natural homomorphism between the cain and certain polynomial domain, called the cain polynomial domain. In the cain polynomial domain, we consider summation of two cain polynomials, which corresponds to the product of two coins. Division of two cain polynomials are also defined, which corresponds to coin marginalization. Problems for deriving a PCI relation from a set of other PCI relations can then be solved through computation of the cain polynomials. We will publish the material on cain polynomials elsewhere.

One potential advantage of the cain over the existing axiomatic systems is that PCI relations can be derived 'automatically' by transforming the coin equations using the (equational) cain axioms. Although not explored in this paper, interesting properties on probability density functions other than PCI relations may also be studied in the cain algebra. For instance, as a referee pointed out, properties such as "no three-factor interaction" in a graphical model can be expressed by density factorization. Thus, the cain algebra may be suited for studying these kinds of properties.

## Appendix

*Proof of Theorem 1.* Let $\Pi = \Pi_{d_1}^{c_1} \cdots \Pi_{d_s}^{c_s}$ be an arbitray expression of $\Pi$ with $c_i \vee d_i > \emptyset$ (i.e., $\Pi_{d_i}^{c_i} \neq 1$). If $c_i > \emptyset$, then by C5 we have $\Pi_{d_i}^{c_i} = \Pi^{c_i \vee d_i} \Pi_{d_i} = \Pi^{c_i \vee d_i} (\Pi^{d_i})^{-1}$. If $c_i = \emptyset$ then $\Pi_{d_i}^{c_i} = (\Pi^{d_i})^{-1}$. So by rearranging the terms, $\Pi$ can be alternatively expressed as

$$\Pi = \left\{ \Pi^{e_1} (\Pi^{d_1})^{-1} \cdots \Pi^{e_m} (\Pi^{d_m})^{-1} \right\} \left\{ (\Pi^{d_{m+1}})^{-1} \cdots (\Pi^{d_s})^{-1} \right\} \tag{26}$$

where $e_i = c_i \vee d_i$ and $m \geq 0$. Let

$$\{f_1, \ldots, f_t\} = \{e_1, d_1, \ldots, e_m, d_m, d_{m+1}, \ldots, d_s\}$$

where $f_i \neq f_j, i \neq j = 1, \ldots, t$. In other words, $f_1, \ldots, f_t$ represent the distinct elements of $e_1, \ldots, e_m$ and $d_1, \ldots, d_s$. Let $u_i$ be the number of $e_j$ appearing in (26), and $v_i$ the number of $d_j$ appearing in (26), so that $e_j = d_j = f_i$, for $i = 1, \ldots, t$.

That is

$$u_i = \sharp\{e_j = f_i \mid j = 1, \ldots, m\}, \quad v_i = \sharp\{d_j = f_i \mid j = 1, \ldots, s\}.$$

By associativity we can then rearrange (26) to get

$$\begin{aligned}\mathbb{T} &= (\mathbb{T}^{f_1})^{u_1}(\mathbb{T}^{f_1})^{-v_1} \cdots (\mathbb{T}^{f_t})^{u_t}(\mathbb{T}^{f_t})^{-v_t} \\ &= (\mathbb{T}^{f_1})^{u_1 - v_1} \cdots (\mathbb{T}^{f_t})^{u_t - v_t} \\ &= (\mathbb{T}^{x_1})^{n_1} \cdots (\mathbb{T}^{x_r})^{n_r}\end{aligned}$$

where $0 \neq n_i = u_i - v_i, i = 1, \ldots, r$. That there exists at least one $n_i \neq 0$ follows from the assumption $\mathbb{T} \neq 1$. This completes the proof. □

*Proof of Theorem* 2. By Theorem 1, there exist nonzero integers $m_1, \ldots, m_s$ and mutually distinct coins $\mathbb{T}^{y_1}, \ldots, \mathbb{T}^{y_s}$ so that

$$\mathbb{T} = (\mathbb{T}^{y_1})^{m_1} \cdots (\mathbb{T}^{y_s})^{m_s}. \tag{27}$$

If $\mathbb{T}^{y_1}$ is not a prime coin, then there exist $y_{11}, \ldots, y_{1t_1}$ so that

$$\mathbb{T}^{y_1} = (\mathbb{T}^{y_{11}})^{m_{11}} \cdots (\mathbb{T}^{y_{1t_1}})^{m_{1t_1}}$$

where $m_{11}, \ldots, m_{1t_1}$ are non-zero integers, and $\mathbb{T}^{y_{11}}, \ldots, \mathbb{T}^{y_{1t_1}}$ are mutually distinct. If any one of these raising coins is not prime, then we can repeat the above decomposition one more time. Since $\mathbb{L}$ satisfies the DCC, this process will have to stop after a finite number of steps. Finally, we will arrive at the following expression:

$$\mathbb{T}^{y_1} = (\mathbb{T}^{z_{11}})^{p_{11}} \cdots (\mathbb{T}^{z_{1w_1}})^{p_{1w_1}}$$

where (a) each $p_{1j}$ is a non-zero integer, (b) each $\mathbb{T}^{z_{1j}}$ is prime, and (c) $\mathbb{T}^{z_{1i}}$ and $\mathbb{T}^{z_{1j}}$ are mutually distinct for $i \neq j$.

Repeating the above arguments for each $\mathbb{T}^{y_i}$ of (27) and putting these expressions into (27), we conclude that $\mathbb{T}$ can be written as

$$\mathbb{T} = (\mathbb{T}^{x_1})^{n_1} \cdots (\mathbb{T}^{x_r})^{n_r} \tag{28}$$

where (a) each $n_i$ is a non-zero integer, (b) each $\mathbb{T}^{x_i}$ is prime, and (c) $\mathbb{T}^{x_i}$ and $\mathbb{T}^{x_j}$ are mutually distinct for $i \neq j$. This proves the existence of the expression (8).

Now we show that the expression (8) is unique. Suppose that we have another expression

$$\mathbb{T} = (\mathbb{T}^{f_1})^{k_1} \cdots (\mathbb{T}^{f_q})^{k_q} \tag{29}$$

where (a) each $k_i$ is a non-zero integer, (b) each $\mathbb{T}^{f_i}$ is prime, and (c) $\mathbb{T}^{f_i}$ and $\mathbb{T}^{f_j}$ are mutually distinct for $i \neq j$. Comparing (28) and (29) we have

$$(\mathbb{T}^{x_1})^{n_1} = (\mathbb{T}^{x_2})^{-n_2} \cdots (\mathbb{T}^{x_r})^{-n_r}(\mathbb{T}^{f_1})^{k_1} \cdots (\mathbb{T}^{f_q})^{k_q}. \tag{30}$$

Since $\mathbb{T}^{x_1}$ is prime, we conclude that there must exist one $f_i$, say $f_1$, so that $x_1 = f_1, n_1 = k_1$. Then by (30) we have $1 = (\mathbb{T}^{x_2})^{-n_2} \cdots (\mathbb{T}^{x_r})^{-n_r} (\mathbb{T}^{f_2})^{k_2} \cdots (\mathbb{T}^{f_q})^{k_q}$, which leads to $(\mathbb{T}^{x_2})^{n_2} = (\mathbb{T}^{x_3})^{-n_3} \cdots (\mathbb{T}^{x_r})^{-n_r} (\mathbb{T}^{f_2})^{k_2} \cdots (\mathbb{T}^{f_q})^{k_q}$. Since $\mathbb{T}^{x_2}$ is prime so we have $x_2 = f_2$ and $n_2 = k_2$. Repeating this process we finally conclude that $q = r$, $n_i = k_i$, $x_i = f_i$ $(i = 1, \ldots, r)$, proving the uniqueness of (8). This completes the proof. □

*Proof of Theorem* 8 (Law of Normalization) For (18), since $x > \emptyset$, we have

$$\begin{aligned}
\mathbb{T}^{(x\vee y)\wedge \bar{z}} &= \int \mathbb{T}^{x\vee y\vee z}\, \mathrm{d}z && \text{(Theorem 6)}\\
&= \int \mathbb{T}^{x}_{y\vee z}\mathbb{T}^{y\vee z}\, \mathrm{d}z && \text{(R-law)}\\
&= \int \mathbb{T}[\bar{z}]\mathbb{T}^{y\vee z}\, \mathrm{d}z && \text{(Assumption)}\\
&= \mathbb{T}[\bar{z}]\int \mathbb{T}^{y\vee z}\, \mathrm{d}z && \text{(Theorem 4)}\\
&= \mathbb{T}[\bar{z}]\mathbb{T}^{y\wedge \bar{z}} && \text{(Theorem 6)}
\end{aligned}$$

implying $\mathbb{T}[\bar{z}] = \mathbb{T}^{(x\vee y)\wedge \bar{z}}\mathbb{T}_{y\wedge \bar{z}} = \mathbb{T}^{x\wedge \bar{z}}_{y\wedge \bar{z}}$ due to Lemma 1, proving (18).

For (19), let $a = \mathfrak{I}(\mathbb{T}[\bar{x}]) \le \bar{x}, b = \mathfrak{I}(\mathbb{T}[\bar{y}]) \le \bar{y}$. Then $\mathbb{T}^{x\vee y\vee z} = \mathbb{T}[\bar{y}]\mathbb{T}[\bar{x}]$ implies $x \vee y \vee z \le a \vee b$ by the sub-additivity of contexts. So $y \le a \vee b$. But $b \le \bar{y}$, so $y \le a$, showing that $\mathbb{T}[\bar{x}]$ is $y$-integrable. Similarly, $\mathbb{T}[\bar{y}]$ is $x$-integrable. Thus

$$\begin{aligned}
\mathbb{T}^{(x\vee z)\wedge \bar{y}} &= \int \mathbb{T}^{x\vee y\vee z}\, \mathrm{d}y && \text{(Theorem 6)}\\
&= \int \mathbb{T}[\bar{y}]\mathbb{T}[\bar{x}]\, \mathrm{d}y && \text{(Assumption)}\\
&= \mathbb{T}[\bar{y}]\int \mathbb{T}[\bar{x}]\, \mathrm{d}y && \text{(Theorem 4)}\\
&= \mathbb{T}[\bar{y}]\hat{\mathbb{T}}[\bar{x}\wedge \bar{y}] && \text{(Eq. (11))}
\end{aligned}$$

implying $\mathbb{T}[\bar{y}] = \mathbb{T}^{(x\vee z)\wedge \bar{y}}\hat{\mathbb{T}}^{-1}[\bar{x}\wedge \bar{y}]$. Similar arguments for $x$ will give $\mathbb{T}[\bar{x}] = \mathbb{T}^{(y\vee z)\wedge \bar{x}}\tilde{\mathbb{T}}^{-1}[\bar{x}\wedge \bar{y}]$. So $\mathbb{T}^{x\vee y\vee z} = \mathbb{T}^{(x\vee z)\wedge \bar{y}}\mathbb{T}^{(y\vee z)\wedge \bar{x}}\mathbb{T}[\bar{x}\wedge \bar{y}]$, where $\mathbb{T}[\bar{x}\wedge \bar{y}] = \hat{\mathbb{T}}^{-1}[\bar{x}\wedge \bar{y}]\tilde{\mathbb{T}}^{-1}[\bar{x}\wedge \bar{y}]$. Thus

$$\begin{aligned}
\mathbb{T}^{(x\vee z)\wedge \bar{y}} &= \int \mathbb{T}^{x\vee y\vee z}\, \mathrm{d}y\\
&= \int \mathbb{T}^{(x\vee z)\wedge \bar{y}}\mathbb{T}^{(y\vee z)\wedge \bar{x}}\mathbb{T}[\bar{x}\wedge \bar{y}]\, \mathrm{d}y
\end{aligned}$$

$$= \mathbb{T}^{(x\vee z)\wedge\bar{y}}\mathbb{T}[\bar{x}\wedge\bar{y}]\int \mathbb{T}^{(y\vee z)\wedge\bar{x}}\,\mathrm{d}y$$
$$= \mathbb{T}^{(x\vee z)\wedge\bar{y}}\mathbb{T}[\bar{x}\wedge\bar{y}]\mathbb{T}^{(y\vee z)\wedge\bar{x}\wedge\bar{y}}$$
$$= \mathbb{T}^{(x\vee z)\wedge\bar{y}}\mathbb{T}[\bar{x}\wedge\bar{y}]\mathbb{T}^{z\wedge\bar{x}\wedge\bar{y}}$$

implying $\mathbb{T}[\bar{x}\wedge\bar{y}] = \mathbb{T}_{z\wedge\bar{x}\wedge\bar{y}}$, proving (19). This completes the proof. □

*Proof of Theorem* 9 *(Law of Marginalization)* It suffices to show $\mathbb{T}_z^{x\vee(y\wedge b)} = \mathbb{T}_z^x\mathbb{T}_z^{y\wedge b}$. Since $x\wedge y=\emptyset$, so we have

$$\begin{aligned}\int \mathbb{T}_z^{x\vee y}\,\mathrm{d}(y\wedge\bar{b}\wedge\bar{z}) &= \mathbb{T}_z^{(x\vee y)\wedge\overline{y\wedge\bar{b}\wedge\bar{z}}} && \text{(Theorem 6)}\\ &= \mathbb{T}_z^{(x\vee y)\wedge(\bar{y}\vee b\vee z)}\\ &= \mathbb{T}_z^{(x\vee y)\wedge(\bar{y}\vee b)} && \text{(Lemma 1)}\\ &= \mathbb{T}_z^{x\vee(y\wedge b)} && \text{(Lemma 7)}\end{aligned}$$

where the third equality is due to $(x\vee y)\wedge(\bar{y}\vee b\vee z) = ((x\vee y)\wedge(\bar{y}\vee b))\vee((x\vee y)\wedge z)$, $(x\vee y)\wedge z\le z$, and Lemma 1. The fourth equality is due to $x = x\wedge(\bar{y}\vee b)$ by Lemma 7 and $(x\vee y)\wedge(\bar{y}\vee b) = (x\wedge(\bar{y}\vee b))\vee(y\wedge(\bar{y}\vee b)) = x\vee((y\wedge\bar{y})\vee(y\wedge b)) = x\vee(y\wedge b)$.

Similarly, since $(x\vee z)\wedge(y\wedge\bar{b}\wedge\bar{z})=\emptyset$ we have

$$\begin{aligned}\int \mathbb{T}_z^x\mathbb{T}_z^y\,\mathrm{d}(y\wedge\bar{b}\wedge\bar{z}) &= \mathbb{T}_z^x\int \mathbb{T}_z^y\,\mathrm{d}(y\wedge\bar{b}\wedge\bar{z}) && \text{(Theorem 4)}\\ &= \mathbb{T}_z^x\mathbb{T}_z^{y\wedge\overline{y\wedge\bar{b}\wedge\bar{z}}} && \text{(Theorem 6)}\\ &= \mathbb{T}_z^x\mathbb{T}_z^{y\wedge(b\vee z)}\\ &= \mathbb{T}_z^x\mathbb{T}_z^{y\wedge b}. && \text{(Lemma 1)}\end{aligned}$$

The two parts jointly prove (25). □

*Proof of Theorem* 19. "P1". The corresponding coin equation of P1, $x\perp\!\!\!\perp y|x$, is

$$\mathbb{T}_x^{x\vee y} = \mathbb{T}_x^x\mathbb{T}_x^y,$$

which holds because $\mathbb{T}_x^x = 1$ and $\mathbb{T}_x^y = \mathbb{T}_x^{x\vee y}$ by Lemma 1.

"P2". The coin algebraic counterpart of P2, $x\perp\!\!\!\perp y|z \Rightarrow y\perp\!\!\!\perp x|z$, is

$$\mathbb{T}_z^{x\vee y} = \mathbb{T}_z^x\mathbb{T}_z^y \Longrightarrow \mathbb{T}_z^{y\vee x} = \mathbb{T}_z^y\mathbb{T}_z^x,$$

which holds due to the commutativity of both the coin product and the join operation.

"P3". Since $(\mathbb{L}, \leq)$ is a lattice, P3 is equivalent to, for any $w \in \mathbb{L}$, the following form:

$$x \perp\!\!\!\perp y|z \Rightarrow x \perp\!\!\!\perp (w \wedge y)|z$$

or in terms of coins,

$$\top\!\!\!\top_z^{x\vee y} = \top\!\!\!\top_z^x \top\!\!\!\top_z^y \Longrightarrow \top\!\!\!\top_z^{x\vee(w\wedge y)} = \top\!\!\!\top_z^x \top\!\!\!\top_z^{w\wedge y} \tag{31}$$

which holds due to the fact $w \wedge y \leq y$ and the M-Law (Theorem 9).

"P4". This axiom is equivalent to, for any $w \in \mathbb{L}$,

$$x \perp\!\!\!\perp y|z \Rightarrow x \perp\!\!\!\perp y|(z \vee (w \wedge y))$$

the corresponding coin form being

$$\top\!\!\!\top_z^{x\vee y} = \top\!\!\!\top_z^x \top\!\!\!\top_z^y \Longrightarrow \top\!\!\!\top_{z\vee(w\wedge y)}^{x\vee y} = \top\!\!\!\top_{z\vee(w\wedge y)}^x \top\!\!\!\top_{z\vee(w\wedge y)}^y.$$

First, by (31), $\top\!\!\!\top_z^{x\vee y} = \top\!\!\!\top_z^x \top\!\!\!\top_z^y$ implies $\top\!\!\!\top_z^{x\vee(y\wedge w)} = \top\!\!\!\top_z^x \top\!\!\!\top_z^{(y\wedge w)}$, which, when multiplied by $\top\!\!\!\top^z \top\!\!\!\top_{z\vee(w\wedge y)}$, gives $\top\!\!\!\top_{z\vee(w\wedge y)}^x = \top\!\!\!\top_z^x$. On the other hand, multiplying $\top\!\!\!\top_z^{x\vee y} = \top\!\!\!\top_z^x \top\!\!\!\top_z^y$ by $\top\!\!\!\top^z \top\!\!\!\top_{(y\wedge w)\vee z}$ gives $\top\!\!\!\top^{x\vee y\vee z} \top\!\!\!\top_{(y\wedge w)\vee z} = \top\!\!\!\top_z^x \top\!\!\!\top^{y\vee z} \top\!\!\!\top_{(y\wedge w)\vee z}$. Since $x \vee y \vee z \geq (y \wedge w) \vee z$ and $y \vee z \geq (y \wedge w) \vee z$, by Lemma 1, we have

$$\top\!\!\!\top_{(y\wedge w)\vee z}^{x\vee y\vee z} = \top\!\!\!\top_z^x \top\!\!\!\top_{(y\wedge w)\vee z}^{y\vee z} = \top\!\!\!\top_{(y\wedge w)\vee z}^x \top\!\!\!\top_{(y\wedge w)\vee z}^{y\vee z},$$

which implies the required result by noting that $\top\!\!\!\top_{(y\wedge w)\vee z}^{x\vee y\vee z} = \top\!\!\!\top_{(y\wedge w)\vee z}^{x\vee y}$ and $\top\!\!\!\top_{(y\wedge w)\vee z}^{y\vee z} = \top\!\!\!\top_{(y\wedge w)\vee z}^y$ due to Lemma 1 again.

"P5". This is identical with the contraction property of the graphoid. For completeness we include a slightly different proof here. $x \perp\!\!\!\perp y|z, x \perp\!\!\!\perp w|(y \vee z) \Rightarrow x \perp\!\!\!\perp (y \vee w)|z$, corresponds to

$$\top\!\!\!\top_z^{x\vee y} = \top\!\!\!\top_z^x \top\!\!\!\top_z^y,\ \top\!\!\!\top_{y\vee z}^{x\vee w} = \top\!\!\!\top_{y\vee z}^x \top\!\!\!\top_{y\vee z}^w \Longrightarrow \top\!\!\!\top_z^{x\vee y\vee w} = \top\!\!\!\top_z^x \top\!\!\!\top_z^{y\vee w}. \tag{32}$$

The R- and L-Law applied to the first condition in (32) using $\top\!\!\!\top^z \top\!\!\!\top_{y\vee z}$ yields $\top\!\!\!\top_{y\vee z}^x = \top\!\!\!\top_z^x$. The R-Law applied to the second condition using $\top\!\!\!\top^{y\vee z}$ gives $\top\!\!\!\top^{x\vee w\vee y\vee z} = \top\!\!\!\top_{y\vee z}^x \top\!\!\!\top^{w\vee y\vee z}$. So $\top\!\!\!\top^{x\vee w\vee y\vee z} = \top\!\!\!\top_z^x \top\!\!\!\top^{w\vee y\vee z}$, which by the L-Law using $\top\!\!\!\top_z$ implies the last equation in (32).

"P6". Note that P6 is equivalent to, for any $z, w \in \mathbb{L}$,

$$x \perp\!\!\!\perp y|(y \wedge z),\quad x \perp\!\!\!\perp y|(y \wedge w) \Longrightarrow x \perp\!\!\!\perp y|(y \wedge z \wedge w)$$

the coin equational form being given by

$$\top\!\!\!\top_{y\wedge z}^{x\vee y} = \top\!\!\!\top_{y\wedge z}^x \top\!\!\!\top_{y\wedge z}^y,\quad \top\!\!\!\top_{y\wedge w}^{x\vee y} = \top\!\!\!\top_{y\wedge w}^x \top\!\!\!\top_{y\wedge w}^y \Longrightarrow \top\!\!\!\top_{y\wedge z\wedge w}^{x\vee y} = \top\!\!\!\top_{y\wedge z\wedge w}^x \top\!\!\!\top_{y\wedge z\wedge w}^y$$

which, by the R-Law, is equivalent to

$$\mathbb{T}^{x\vee y} = \mathbb{T}^{x}_{y\wedge z}\mathbb{T}^{y}, \quad \mathbb{T}^{x\vee y} = \mathbb{T}^{x}_{y\wedge w}\mathbb{T}^{y} \Longrightarrow \mathbb{T}^{x\vee y} = \mathbb{T}^{x}_{y\wedge z\wedge w}\mathbb{T}^{y}. \tag{33}$$

The conditions in (33) imply $\mathbb{T}^{x}_{y\wedge z} = \mathbb{T}^{x}_{y\wedge w}$, which, by Theorem 5, implies in turn $\mathbb{T}^{x}_{y\wedge z} = \mathbb{T}^{x}_{y\wedge z\wedge w}$, proving (33). □

**Acknowledgments** I wish to thank Professors Masafumi Akahira, Masaaki Shibuya and Akimichi Takemura for encouragements and thought-provoking comments during various stages of this work. The comments of two anonymous referees helped improve the readability of the paper. This work was partially supported by Grand-in-Aid 19300088 from the Japanese Ministry of Education, Science, Sports and Culture.

## References

Andersson, S. A., Madsen, J. (1998). Symmetry and lattice conditional independence in a multivariate normal distribution. *Annals of Statistics, 26*, 525–572.

Andersson, S. A., Perlman, M. D. (1993). Lattice models for conditional independence in a multivariate normal distribution. *Annals of Statistics, 21*, 1318–1358.

Andersson, S. A., Perlman, M. D. (1995a). Testing lattice conditional independence models. *Journal of Multivariate Analysis, 53*, 18–38.

Andersson, S. A., Perlman, M. D. (1995b). Unbiasedness for the likelihood ratio test for lattice conditional independence models. *Journal of Multivariate Analysis, 53*, 1–17.

Andersson, S. A., Madigan, D., Perlman, M. D., Triggs, C. M. (1995). On the relation between conditional independence models determined by finite distributive lattices and by directed acyclic graphs. *Journal of Statistical Planning and Inference, 48*, 25–46.

Andersson, S. A., Madigan, D., Perlman, M. D., Triggs, C. M. (1997). A graphical characterization of lattice conditional independence models. *Annals of Mathematics and Artificial Intelligence 21*, 27–50.

Birkhoff, G. (1967). *Lattice theory* (Vol. 25, 3rd ed.). New York: American Mathematical Society Colloquium Publications, American Mathematical Society.

Blyth, T. S. (2005). *Lattices and ordered algebraic structures*. London: Springer.

Cox, D., Little, J., O'Shea, D. (1997). *Ideals, varieties and algorithms* (2nd ed.). New York: Springer.

Davey, B. A., Priestley, H. A. (2002). *Introduction to lattices and order*. Cambridge: Cambridge University Press.

Dawid, A. P. (1979a). Conditional independence in statistical theory (with discussion). *Journal of the Royal Statistical Society: Series B, 41*, 1–31.

Dawid, A. P. (1979b). Some misleading arguments involving conditional independence. *Journal of the Royal Statistical Society: Series B, 41*, 249–252.

Dawid, A. P. (1980a). A Bayesian look at nuisance parameters. In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, A. F. M. Smith (Eds.), *Bayesian statistics* (pp. 167–184). Valencia: Valencia University Press.

Dawid, A. P. (1980b). Conditional independence for statistical operations. *Annals of Statistics, 8*, 598–617.

Dawid, A. P. (1985). Invariance and independence in multivariate distribution theory. *Journal of Multivariate Analysis, 17*, 304–315

Dawid, A. P. (1988). Conditional independence. In S. Kotz, C. B. Read, D. L. Banks (Eds.), *Encyclopedia of statistical science* (Vol. 2, pp. 146–155). New York: Wiley.

Dawid, A. P. (2001). Separoids: A mathematical framework for conditional independence and irrelevance. *Annals of Mathematics and Artificial Intelligence, 32*, 335–372.

Dawid, A. P., Evett, I. W. (1997). Using a graphical method to assist the evaluation of complicated patterns of evidence. *Journal of Forensic Sciences, 42*, 226–231.

Dawid, A. P., Lauritzen, S. L. (1993). Hyper Markov laws in the statistical analysis of decomposable graphical models. *Annals of Statistics, 21*, 1272–1317.

Dawid, A. P., Mortera, J. (1996). Coherent analysis of forensic identification evidence. *Journal of the Royal Statistical Society: Series B, 58*, 425–443.

Dawid, A. P., Studený, M. (1999). Conditional products: an alternative approach to conditional independence. In D. Heckerman, J. Whittaker (Eds.), *Artificial intelligence and statistics* (Vol. 99, pp. 32–40). San Francisco: Morgan Kaufmann.

Freese, R., Ježek, J., Nation, J. B. (1995). *Free lattices*. Mathematical surveys and monographs. Providence: American Mathematical Society.

Galles, D., Pearl, J. (1997). Axioms of causal relevance. *Artificial Intelligence, 97*, 9–43.

Halmos, P. R. (1974). *Lectures on boolean algebras*. New York: Springer.

Jipsen, P., Rose, H. (1992). *Varieties of lattices*. Lecture notes in mathematics (Vol. 1533). Springer: New York. http://www1.chapman.edu/~jipsen/JipsenRoseVoL.html.

Jónsson, B., Kiefer, J. (1962). Finite sublattices of a free lattice. *Canadian Journal of Mathematics, 14*, 487–497.

Lauritzen, S. L. (1982). *Lectures on contingency tables* (2nd ed.) Aalborg: University of Aalborg Press.

Lauritzen, S. L. (1996). *Graphical models*. Oxford: Clarendon.

Lauritzen, S. L., Dawid, A. P., Larsen, B. N., Leimer, H. G. (1990). Independence properties of directed Markov fields. *Networks, 20*, 491–505.

Massam, H., Neher, E. (1998). Estimation and testing for lattice conditional independence models on Euclidean Jordan algebras. *Annals of Statistics, 26*, 1051–1082.

Nation, J. B. *Notes on lattice*, unpublished course notes. http://www.math.hawaii.edu/~jb/lat1-6.pdf.

Paz, A., Pearl, J. (1994). Axiomatic characterization of directed graphs. Technical Report R-234. Los Angeles: Cognitive Systems Laboratory, Computer Science Department, University of California.

Paz, A., Pearl, J., Ur, S. (1996). A new characterization of graphs based on interception relations. *Journal of Graph Theory, 22*, 125–136.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Mateo: Morgan Kaufmann.

Pearl, J. (2000). *Causality—models, reasoning and inference*. New York: Cambridge University Press.

Pearl, J., Paz, A. (1987). Graphoids: a graph-based logic for reasoning about relevance relations. In D. Hogg, L. Steels (Eds.), *Advances in artificial intelligence* (pp. 357–363). Amsterdam: North-Holland.

Reichenbach, H. (1956). *The direction of time*. Berkeley: University of California Press.

Sagiv, Y., Walecka, S. F. (1982). Subset dependencies and completeness result for a subclass of embedded multivalued dependencies. *Journal of the Association for Computing Machinery, 29*, 103–117.

Salii, V. N. (1988). *Lattices with unique complements*. Translations of mathematical monographs. Providence: American Mathematical Society.

Schoken, S. S., Hummel, R. A. (1993). On the use of the dempster shafer model in information indexing and retrieval applications. *International Journal of Man–Machine Studies, 39*, 1–37.

Shafer, G. (1976). *A Mathematical theory of evidence*. Princeton: Princeton University Press.

Shenoy, P. P. (1994). Conditional independence in valuation-based systems. *International Journal of Approximate Reasoning, 10*, 203–234.

Spohn, W. (1980). Stochastic independence, causal independence, and shieldability. *Journal of Philosophical Logic, 9*, 73–99.

Spohn, W. (1988). Ordinal conditional functions: a dynamic theory of epistemic states. In W. L. Harper, B. Skyrms (Eds.), *Causation in decision, belief change, and statistics* (Vol. 2, pp. 105–134). Dordrecht: Kluwer.

Spohn, W. (1994). On the properties of conditional independence. In P. Humphreys (Ed.), *Patrick Suppes: scientific philosopher*. Probability and probabilistic causality (Vol. 1, pp. 173–194). Dordrecht: Kluwer.

Stanley, B. N., Sankappanavar, H. P. (1981). *A course in universal algebra*. New York: Springer. http://www.math.uwaterloo.ca/~snburris/htdocs/ualg.html.

Studený, M. (1993). Formal properties of conditional independence in different calculi of AI. In *Symbolic and quantitative approaches to reasoning and uncertainty*. Lecture notes in computer science (Vol. 747, pp. 341–348). Berlin: Springer.

Studený, M. (2005). *Probabilistic conditional independence structures* (Vol. 97). London: Springer.

Teixeira de Silva, W., Milidiu, R. L. (1993). Belief function model for information retrieval. *Journal of the American Society for Information Science, 44*, 10–18.

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems, 1*, 3–28.