数理解析研究所講究録
1506 巻 2006 年 177-196

The Coin Algebra for Conditional Independence[1]

千葉大学・大学院自然科学研究科　汪　金芳 (Jinfang Wang)
Graduate School of Science and Technology
Chiba University

1 Introduction

In this paper we introduce a universal algebraic structure termed the *coin*. The coin algebra is intended for formal manipulation of the conditional independence relation intrinsically associated with random variables. In the coin algebra, conditional independence is defined as a special coin equation. The major advantage of the universal algebraic definition of the coin algebra is that all properties on conditional independence can be derived by transforming one coin equation to another coin equation. Secondly, but equally importantly, the axioms of the coin algebra are developed from the basic properties of probability density functions. This is in contrast with other well-known axiomatic systems for conditional independence, which typically focus on a few "principal" properties of conditional independence widely useful for "probabilistic reasoning" involving the concept of irrelevance probabilistic or not. It is therefore not surprising that the properties on conditional indepdence derivable from the coin algebra proposed in this paper include, but not limited to, the axioms of a *strong separoid* of Dawid (2001), the graphoid of Pearl and Paz (1987) being a special but particularly important example.

2 The Coin

2.1 Definition

Let $\mathcal{D} = \{D|D \subset \mathbb{D}\}$ be the power set of $\mathbb{D}$, which include the empty set $\emptyset$. Let $\mathcal{D} \otimes \mathcal{D}$ be the *exclusive direct product* of $\mathcal{D}$ and $\mathcal{D}$, that is

$$\mathcal{D} \otimes \mathcal{D} = \{(R, L) \mid R, L \in \mathcal{D} \text{ and } R \cap L = \emptyset\}$$

Note that (R, L) are (L, R) are regarded as different if $R \neq L$. Note also that $(\emptyset, \emptyset) \in \mathcal{D} \otimes \mathcal{D}$.

DEFINITION 2.1 (COIN OPERATOR). *The coin*[2] *operator, denoted by* $\mathbb{T}$, *is a binary operator defined on the exclusive direct product space* $\mathcal{D} \otimes \mathcal{D}$ *to the posive real line,* $\mathbb{T} : \mathcal{D} \otimes \mathcal{D} \longrightarrow \mathbb{R}^+$. *For* $(R, L) \in \mathcal{D} \otimes \mathcal{D}$, *we shall write* $\mathbb{T}_L^R$ *(reads as* `coin-R-L`*) instead of* $\mathbb{T}(R, L)$ *to denote the image of* (R, L) *by* $\mathbb{T}$.

[1]Partially supported by Grand-in-Aid 15500179 of the Japanese Ministry of Education, Science, Sports and Culture. The author thanks Professor A.P. Dawid for pointing out the work of Dawid (2001), and Professor M. Akahira for discussions on conditional independece and statistical sufficiency.

[2]*Coin* stands for *co*nditional *in*dependence.

The coin operator Π satisfies the following axioms.

$$\text{Normalization Axiom:} \quad \Pi_{\emptyset}^{\emptyset} = 1 \tag{2.1}$$

$$\text{Coditionality Axiom:} \quad \Pi_{L}^{R} = \Pi^{RL}\Pi_{L} \qquad R \neq \emptyset \tag{2.2}$$

$$\text{Inversion Axiom:} \quad \Pi_{L}\Pi^{L} = 1 \tag{2.3}$$

where $\Pi^{R} \equiv \Pi_{\emptyset}^{R}$ *and* $\Pi_{L} \equiv \Pi_{L}^{\emptyset}$ *and* $RL = R \cup L$.

Note that by (2.1), we have $\Pi^{\emptyset} = \Pi_{\emptyset} = 1$.

DEFINITION 2.2 (ATOM COINS). *For any* $(R, L) \in \mathcal{D}\otimes\mathcal{D}$, *we shall call* Π_{L}^{R} *(reads as* `coin-R-L`*) the atom coin with* raising index R *and* lowering index L.

Note that when $R_1 = R_2, L_1 = L_2$ we have $\Pi_{L_1}^{R_1} = \Pi_{L_2}^{R_2}$. But the reverse, that is, $\Pi_{L_1}^{R_1} = \Pi_{L_2}^{R_2}$ implying $R_1 = R_2, L_1 = L_2$, can not be derived from the definition above.

DEFINITION 2.3 (RAISING, LOWERING, MIXED COINS). *We classify the atom coins into three types.*

(i) Raising coin: Π^{R} *is called a raising coin with raising index R.*

(ii) Lowering coin: Π_{L} *is called a lowering coin with lowering index L.*

(iii) Mixed coin: Π_{L}^{R} *is called a mixed coin with raising index R and lowering index L.*

In Definition 2.1 for the coin operator Π, multiplications of the coins appear both in (2.2) and in (2.3). We have implicitely assumed that these products are carried out with respect to the usual multiplication of real numbers. This shall be assumed throughout the paper.

We know that joint probability density functions and conditional density functions can be multiplied to give other joint (conditional) density functions. For instance, the Bayes theorem states that

$$f(\omega_R|\omega_L) = f(\omega_L|\omega_R)f(\omega_R)f^{-1}(\omega_L)$$

This can be translated in terms of coins as follows

THEOREM 2.1 (BAYES THEOREM). *if* $R \neq \emptyset, L \neq \emptyset$ *and* $R \cap L = \emptyset$, *then*

$$\Pi_{L}^{R} = \Pi_{R}^{L}\Pi^{R}\Pi_{L} \tag{2.4}$$

Raising and lowering the indices of the mixed coin Π_{R}^{L} on r.h.s. of (2.4) using the raising coin Π^{R} and the lowering coin Π_{L}, the r.h.s. can be 'transformed' to the l.h.s. as follows.

$$\Pi_{R}^{L}\Pi^{R}\Pi_{L} = \left(\Pi_{R}^{L}\Pi^{R}\right)\Pi_{L} \Rightarrow \Pi^{RL}\Pi_{L} \Rightarrow \Pi_{L}^{R}$$

This 'proves' the Bayes theorem. A formal proof using the axioms of coins goes as follows.

Proof. Repeatedly applying (2.2) to appropriate mixed coins we have

$$\begin{aligned} \mathbb{T}_R^L \mathbb{T}^R \mathbb{T}_L &= \left(\mathbb{T}^{RL}\mathbb{T}_R\right)\left(\mathbb{T}^R\mathbb{T}_L\right) \\ &= \mathbb{T}^{RL}\left(\mathbb{T}_R\mathbb{T}^R\right)\mathbb{T}_L \\ &= \mathbb{T}^{RL}\mathbb{T}_L \\ &= \mathbb{T}_L^R. \end{aligned}$$

The first equality uses the C-Axiom to $\mathbb{T}_R^L$ and the last equality uses the C-Axiom to $\mathbb{T}_L^R$. This completes the proof. □

For probability density functions, when ω_R and ω_L are independent, the product of the density functions $f(\omega_A)f(\omega_B) = f(\omega_{A\cup B})$ gives the joint density function of $\omega_{A\cup B}$. When this assumption is not available, that is when ω_A and ω_B are correlated, there may exist no subsets $C, D \subset \mathbb{D}$ so that $f(\omega_C|\omega_D) = f(\omega_A)f(\omega_B)$. The same is true for atom coins. In general, the product of two atom coins is not necessarily an atom coin. That is, the set of all atom coins is not closed under the usual multiplication.

DEFINITION 2.4 (COIN). *A* coin, $\mathbb{T}$, *is a product of an arbitrary finite sequence of the atom coins. That is, there exist* $(R_i, L_i) \in \mathcal{D} \otimes \mathcal{D}, i = 1, \cdots, r$ *such that*

$$\mathbb{T} = \mathbb{T}_{L_1}^{R_1}\mathbb{T}_{L_2}^{R_2}\cdots\mathbb{T}_{L_r}^{R_r} \tag{2.5}$$

The set of all coins will be denoted by $\mathbb{I}(D)$ *or simply by* $\mathbb{I}$ *when* D *is clear from the context.*

We have used the same symbol $\mathbb{T}$ to denote both the coin operator and a general coin. At present there is no ambiguity anticipated with this abuse of notation. We shall also use the genetic symbols $\mathbb{T}$, $\tilde{\mathbb{T}}$, $\hat{\mathbb{T}}$, etc., to denote arbitrary coins in $\mathbb{I}$.

Note that since the C-Axiom postulates that $\mathbb{T}_L^R = \mathbb{T}^{RL}\mathbb{T}_L$, it is then evident that the set of all coins $\mathbb{I}$ is generated by all raising and lowering coins.

EXAMPLE 2.1. *Let* $\mathbb{D} = \{1,2\}$. *Let* $\mathbb{I}$ *be the set of all coins. Let* $\mathbb{T} = \mathbb{T}^R$ *be a raising coin. If* n *is a positive integer then we denote by* $(\mathbb{T})^n$ *the product of* n *copies of* $\mathbb{T}$. *If* n *is a negative integer then we denote by* $(\mathbb{T})^n$ *the product of* $|n|$ *copies of* $\mathbb{T}_R$. *When* $n = 0$, *we let* $\mathbb{T}^n = 1$. *Then* $\mathbb{I}$ *can be written as*

$$\mathbb{I} = \left\{(\mathbb{T}^1)^{n_1}(\mathbb{T}^2)^{n_2}(\mathbb{T}^{12})^{n_3} \mid n_1, n_2, n_3 = 0, \pm1, \pm2, \cdots\right\}$$

THEOREM 2.2 (COIN GROUP). *The set of all coins* $\mathbb{I}$ *forms a commutative (Abelian) group with respect to the usual multiplication of real numbers. We call* $\mathbb{I}$ *the* coin group.

NOTATION 2.1. *Since the inverse of any element of a group is unique we shall use the symbol* $(\mathbb{T})^{-1}$ *or simply* $\mathbb{T}^{-1}$ *to denote the inverses of* $\mathbb{T}$.

2.2 The raising-up and lowering-down laws

Now we justify the terminologies that $\mathbb{T}^R$ are raising coins and $\mathbb{T}_L$ are lowering coins.

THEOREM 2.3 (RAISING-UP LAW). *(i) For any* $(R, L) \in \mathcal{D} \otimes \mathcal{D}$, *we have*

$$\mathbb{T}^{RL} = \mathbb{T}_L^R \mathbb{T}^L \qquad (R \neq \emptyset) \tag{2.6}$$

(ii) if A, B, C *are mutually exclusive then we have*

$$\mathbb{T}_C^{AB} = \mathbb{T}_{BC}^A \mathbb{T}^B \Leftrightarrow \mathbb{T}^{BC} = \mathbb{T}^B \mathbb{T}^C \qquad (A \neq \emptyset) \tag{2.7}$$

The atom coin $\mathbb{T}^{RL}$ on the left hand side of (2.6) is obtained by *raising* the subscript L of the mixed coin $\mathbb{T}_L^R$ by the atom coin $\mathbb{T}^L$. Similarly, the atom coin $\mathbb{T}_C^{AB}$ on the left hand side of (2.7) is obtained by *detaching* the subscript B from BC of the mixed coin $\mathbb{T}_{BC}^A$ and *raising* B by the atom coin $\mathbb{T}^B$. To validate the detachment of B from BC, we need the condition $\mathbb{T}^{BC} = \mathbb{T}^B \mathbb{T}^C$, a condition, we shall see later, says that 'B is independent of C'. Note that for any $L \subset \mathbb{D}$ we have $\mathbb{T}^{L\emptyset} = \mathbb{T}^L \mathbb{T}^\emptyset = \mathbb{T}^L$. That is, '$L$ is independent of $\emptyset$'. It follows that (2.6) may be regarded as a special case of (2.7).

We shall refer to the relations (2.6) and (2.7) as the *Raising-up Law*, or the *R-Law* for short.

THEOREM 2.4 (LOWERING-DOWN LAW). *(i) For any* $(R, L) \in \mathcal{D} \otimes \mathcal{D}$, *we have*

$$\mathbb{T}_L^R = \mathbb{T}^{RL} \mathbb{T}_L \qquad (R \neq \emptyset) \tag{2.8}$$

(ii) if A, B, C *are mutually exclusive then we have*

$$\mathbb{T}_{BC}^A = \mathbb{T}_C^{AB} \mathbb{T}_B \Leftrightarrow \mathbb{T}^{BC} = \mathbb{T}^B \mathbb{T}^C \qquad (A \neq \emptyset) \tag{2.9}$$

The mixed coin $\mathbb{T}_L^R$ on the left hand side of (2.8) is obtained by *lowering* the subscript L of the raising coin $\mathbb{T}^{RL}$ by the coin $\mathbb{T}_L$. Similarly, the mixed coin $\mathbb{T}_{BC}^A$ on the left hand side of (2.9) is obtained by *lowering* the subscript B of the mixed coin $\mathbb{T}_C^{AB}$ by the atom coin $\mathbb{T}_B$, and *emerging* B with the already-existing subscript C. To ensure the validity of emerging of B with C, we need the condition $\mathbb{T}^{BC} = \mathbb{T}^B \mathbb{T}^C$. Again since $\mathbb{T}^{B\emptyset} = \mathbb{T}^B \mathbb{T}^\emptyset$, it follows that (2.9) reduces to (2.8) by letting $C = \emptyset$ in (2.9).

We shall refer to both (2.8) and (2.9) as the *Lowering-down Law*, or the *L-Law* for short. Both the R-Law and the L-Law can be very convenient for combining or decomposing various coins without referring to the formal properties of coins. We shall see later that these laws can be so powerful that they can even be helpful for 'discovering' necessary and sufficient conditions for independence and conditional indpendence relations among a set of variables.

The following Lemma is useful for transforming a given coin equation to another coin equation. For instance, it can be used to show the *chaining rule* (Lauritzen, 1982) of conditional independence.

LEMMA 2.1. *Let* A, B, C, D *be mutually disjoint subsets of* $\mathbb{D}$, *then*

$$\mathbb{T}_A^{BC} = \mathbb{T}_A^{BD} \mathbb{T}[\mathbb{D}] \Longleftrightarrow \mathbb{T}_B^{AC} = \mathbb{T}_B^{AD} \mathbb{T}[\mathbb{D}] \tag{2.10}$$

where $\mathbb{T}[\mathbb{D}]$ *is an arbitrary coin.*

2.3 Canonical expressions and the null model

Let $\mathbb{T} \in \mathbb{I}$ be an arbitrary coin. By definition there exists $(A_i, B_i) \in \mathcal{D} \otimes \mathcal{D}, i = 1, \cdots, r$ so that

$$\mathbb{T} = \mathbb{T}_{B_1}^{A_1} \times \cdots \times \mathbb{T}_{B_r}^{A_r} \tag{2.11}$$

When $\mathbb{T}$ is written in (2.11), we call the right hand side of (2.11) an *expression* of $\mathbb{T}$ with length r.

However, given any prob $\mathbb{T}$, there are infinitely many expressions which are equal to one another. To see this, we note that for any $A \subset \mathbb{D}$ and any integer n, we have $1 = (\mathbb{T}^A)^n (\mathbb{T}_A)^n$. It follows then $\mathbb{T} = \mathbb{T}\{(\mathbb{T}^A)^n (\mathbb{T}_A)^n\}$. For less trivial examples, suppose that $\mathbb{D} = A_1 \sqcup A_2 = B_1 \sqcup B_2 \sqcup B_3$. Then we can write $\mathbb{T}^{\mathbb{D}}$ in many different ways such as

$$\begin{aligned}\mathbb{T}^{\mathbb{D}} &= \mathbb{T}^{A_1}\mathbb{T}^{A_2}_{A_1} = \mathbb{T}^{A_2}\mathbb{T}^{A_1}_{A_2}\\ \mathbb{T}^{\mathbb{D}} &= \mathbb{T}^{B_1}\mathbb{T}^{B_2}_{B_1}\mathbb{T}^{B_3}_{B_1B_2} = \mathbb{T}^{B_2}\mathbb{T}^{B_1}_{B_2}\mathbb{T}^{B_3}_{B_1B_2}\end{aligned}$$

These different expressions may be useful from an *inferential* viewpoint, but they are equivalent in the sense that they impose no additional restriction, whatever on the the coin group $\mathbb{I}$.

The fact that any positive integer n can be uniquely written as $n = p_1^{n_1} \cdots p_r^{n_r}$ where $n_1, \cdots, n_r$ are positive integers and $p_1, \cdots, p_r$ are prime numbers plays an important role in number theory. It will be conveninent to do the same for coins.

DEFINITION 2.5 (MUTUALLY PRIME COINS). *Two raising coins $\mathbb{T}^A$ and $\mathbb{T}^B$ with $A \neq \emptyset, B \neq \emptyset$ are said to be* mutually prime, *if $A \neq B$.*

EXAMPLE 2.2. *Let $\mathbb{D} = A \sqcup B \sqcup C$ and none of A, B, C is empty. Then*

$$\mathbb{T}^{\mathbb{D}}, \mathbb{T}^A, \mathbb{T}^{AB}, \mathbb{T}^C, \mathbb{T}^{AC}$$

are mutually prime coins.

THEOREM 2.5. *Let $1 \neq \mathbb{T} \in \mathbb{I}$ be an arbitrary coin. Then there exists nonzero integers $n_1, \cdots, n_r$ and mutually prime coins $\mathbb{T}^{A_i}, i = 1, \cdots, r$ such that the following holds*

$$\mathbb{T} = (\mathbb{T}^{A_1})^{n_1} \times \cdots \times (\mathbb{T}^{A_r})^{n_r} \tag{2.12}$$

DEFINITION 2.6 (PRIME COIN). *A raising coin $\mathbb{T}^A$ is called a* prime coin *if there does not exist an expression*

$$\mathbb{T}^A = (\mathbb{T}^{A_1})^{n_1} \times \cdots \times (\mathbb{T}^{A_r})^{n_r}$$

so that each $A_i, i = 1, \cdots, r$ is a proper *subset of A.*

DEFINITION 2.7 (NULL MODEL). *A coin group $\mathbb{I} = \mathbb{I}(\mathbb{D})$ is called a* null model *if every raising coin $\mathbb{T}^A$ is a prime coin.*

In a null model $\mathbb{I}(\mathbb{D})$ (a coin group), there are no additional assumptions on relations among coins of $\mathbb{I}$, other than those stated in the definition of coins. This corresponds to the situation that the joint probability density function $f(\omega_A)$ cannot not be decomposed by using marginal density functions or conditional density functions for *any* $A \subset \mathbb{D}$.

THEOREM 2.6. *Every coin* $\mathbb{T} \in \mathbb{I}$ *has a unique expression*

$$\mathbb{T} = (\mathbb{T}^{A_1})^{n_1} \times \cdots \times (\mathbb{T}^{A_r})^{n_r} \tag{2.13}$$

where $n_1, \cdots, n_r$ *are nonzero integers and*

(i) $\mathbb{T}^{A_1}, \cdots, \mathbb{T}^{A_r}$ *are prime coins; and*

(ii) $\mathbb{T}^{A_1}, \cdots, \mathbb{T}^{A_r}$ *are mutually prime.*

DEFINITION 2.8 (CANONICAL EXPRESSION). *The unique expression of a coin* $\mathbb{T}$ *given by (2.12) is called a* canonical expression *of* $\mathbb{T}$. *And we call*

(i) r *the order of* $\mathbb{T}$, *and write* $|\mathbb{T}| = r$; *and*

(ii) $A = A_1 \cup \cdots \cup A_r$ *the* index set, *or simply the* index, *of* $\mathbb{T}$, *and write* $\mathfrak{I}(\mathbb{T}) = A$.

THEOREM 2.7. *The index* $\mathfrak{I}(\cdot)$ *has the following properties.*

(i) $\mathfrak{I}(\mathbb{T}) = \mathfrak{I}(\mathbb{T}^{-1})$;

(ii) sub-additivity: $\mathfrak{I}(\mathbb{T}\hat{\mathbb{T}}) \subset \mathfrak{I}(\mathbb{T}) \cup \mathfrak{I}(\hat{\mathbb{T}})$.

3 Coin Integration

3.1 The Integrands

It is useful to introduce the following notations to distinguish three different types of coins.

NOTATION 3.1. *Let* A *be a subset of* $\mathbb{D}$.

(i) $\mathbb{T}^A$ *denotes the raising coin in* $\mathbb{I}$ *with raising index* A.

(ii) $\mathbb{T}[A]$ *denotes an arbitrary coin restricted to* $\mathbb{I}_A$, *the coin group with respect to* A.

(iii) $\mathbb{T}\{A\}$ *denotes an arbitrary coin in* $\mathbb{I}_A$ *with index* $\mathfrak{I}(\mathbb{T}\{A\})$ *equal to* A.
Note that $\mathfrak{I}(\mathbb{T}^A) = \mathfrak{I}(\mathbb{T}\{A\}) = A$ *and* $\mathfrak{I}(\mathbb{T}[A]) \subset A$.

DEFINITION 3.1 (INTEGRAND). *Let* $\mathcal{D}$ *be the power set of* $\mathbb{D}$. *Let* $A \in \mathcal{D}$. *We denote by* $\mathbb{I}(A)$ *the set of all coins* $\mathbb{T}\{B\}$ *of* $\mathbb{I}$ *such that* A *is a subset of* B, *that is*

$$\mathbb{I}(A) = \{\mathbb{T}\{B\} \mid A \subset B \in \mathcal{D}\}$$

We shall call any $\mathbb{T}\{B\} \in \mathbb{I}(A)$ *an* integrand with respect to A, *or simply an* A-integrand. *The set of coins* $\mathbb{I}(A)$ *will be referred to as the* A-integrand set.

Note that the $\emptyset$-integrand set contains the unit coin 1 only. On the other hand, the $\mathbb{D}$-integrand set consists of all coins with the index set $\mathbb{D}$, the largest possible set. We also note that if $\mathbb{T} \in \mathbb{I}(A)$ then $\mathbb{T}^{-1} \in \mathbb{I}(A)$. To see this suppose that $\mathbb{T}$ has the following unique canonical expression

$$\mathbb{T} = (\mathbb{T}^{A_1})^{m_1} \times \cdots \times (\mathbb{T}^{A_r})^{m_r}$$

then $\mathbb{T}^{-1}$ is given by

$$\mathbb{T}^{-1} = (\mathbb{T}^{A_1})^{-m_1} \times \cdots \times (\mathbb{T}^{A_r})^{-m_r}$$

which is also canonical. Hence

$$\mathfrak{I}(\mathbb{T}^{-1}) = \cup_{i=1}^{r} A_i = \mathfrak{I}(\mathbb{T}) \supset A$$

showing that $\mathbb{T}^{-1} \in \mathbb{I}(A)$.

However, when $\mathbb{T}, \hat{\mathbb{T}} \in \mathbb{I}(A)$ it is not true that $\mathbb{T}\hat{\mathbb{T}} \in \mathbb{I}(A)$. For instance, if $A = A_1 \sqcup A_2$, $\mathbb{T} = \mathbb{T}^{A_1}\mathbb{T}^{A_2}$, $\hat{\mathbb{T}} = \mathbb{T}_{A_1}\mathbb{T}^{A_2}$ then $\mathbb{T}\hat{\mathbb{T}} = (\mathbb{T}^{A_2})^2 \notin \mathbb{I}(A)$. So $\mathbb{I}(A)$ is not closed under mutiplication. In other words, $\mathbb{I}(A)$ does not form a subgroup.

3.2 Definition and Properties

Now we are in a position to define the *integration cf coins.*

DEFINITION 3.2 (INTEGRATION). *Let $A \in \mathcal{D}$ and $\mathbb{I}(A)$ be the A-integrand set. We define the* A-integration, *or simply* integration, *as a function, denoted by $\int_A$, from $\mathbb{I}(A)$ into $\mathbb{I}$,*

$$\int_A : \mathbb{I}(A) \longrightarrow \mathbb{I}$$

so that for any A-integrand $\mathbb{T}\{B\} \in \mathbb{I}(A)$, there is a unique coin $\mathbb{T}\{B \setminus A\} \in \mathbb{I}$ such that

$$\int_A (\mathbb{T}\{B\}) = \mathbb{T}\{B \setminus A\} \tag{3.1}$$

The following properties hold for the integration.

(i) if the raising coin $\mathbb{T}^B$ is an A-integrand then we have

$$\int_A (\mathbb{T}^B) = \mathbb{T}^{B\setminus A} \tag{3.2}$$

(ii) Let $A = A_1 \sqcup A_2$. Let $\mathbb{T}\{B\} = \mathbb{T}\{B_1\}\mathbb{T}\{B_2\}$ be an A-integrand, where $\mathbb{T}\{B_1\}$ is an A_1-integrand and $\mathbb{T}\{B_2\}$ is an A_2-integrand. Further assume that $A_1 \cap B_2 = A_2 \cap B_1 = \emptyset$. Then we have

$$\begin{aligned} \int_A (\mathbb{T}\{B\}) &= \int_{A_1 \sqcup A_2} (\mathbb{T}\{B_1\}\mathbb{T}\{B_2\}) \\ &= \int_{A_1} (\mathbb{T}\{B_1\}) \int_{A_2} (\mathbb{T}\{B_2\}) \end{aligned} \tag{3.3}$$

(iii) Finally, for any coin $\mathbb{T} \in \mathbb{I}$ *we have*

$$\int_{\emptyset}(\mathbb{T}) = \mathbb{T} \tag{3.4}$$

REMARK 3.1. *Note that for any coin* $\mathbb{T} \in \mathbb{I}$ *if* B *is the index set of* $\mathbb{T}$*, then we can integrate* $\mathbb{T}$ *by applying* $\int_A$ *for any* $A \subset B$*. That is,* $\int_A(\mathbb{T})$ *is well defined for any* $A \subset B$*.*

NOTATION 3.2. *To mimic the conventional notation for integration, we shall use the following notation for coin integration*

$$\int_A(\mathbb{T}) = \int \mathbb{T}\, dA$$

Using this notation, the defining properties of the coin integration (3.2)-(3.4) can be reexpressed as

$$\int \mathbb{T}^B\, dA = \mathbb{T}^{B\setminus A} \tag{3.5}$$

$$\int (\mathbb{T}\{B_1\}\mathbb{T}\{B_2\})\, dA_1 \sqcup A_2 = \int \mathbb{T}\{B_1\} dA_1 \int \mathbb{T}\{B_2\} dA_2 \tag{3.6}$$

$$\int \mathbb{T} d\emptyset = \mathbb{T} \tag{3.7}$$

Note that (3.5) corresponds to the definition of marginal probability density functions. Note also that if we let $A = \emptyset$ in (3.5) then we have $\int \mathbb{T}^B\, d\emptyset = \mathbb{T}^B$, which is a special case of (3.7). The requirment (3.6) corresponds to the following basic property of the usual integration

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

The requirment (3.7) is used to show (3.8), a property corresponds to the following basic property of the usual integration

$$\int cf(x,y)\, dx = c \int f(x,y)\, dx$$

where c is a constant functionally independent of both x and y.

Now we discuss consequences of the definiton of the integration. Using (3.6) and (3.7) we get the following property for the coin integration.

THEOREM 3.1. *Let* $\mathbb{T}\{B\}$ *be an A-integrand. Suppose that* $A \cap C = \emptyset$*. The we have*

$$\int \mathbb{T}[C]\mathbb{T}\{B\}\, dA = \mathbb{T}[C] \int \mathbb{T}\{B\}\, dA \tag{3.8}$$

The following two theorems are anologies of the facts that both the probability density functions and the conditional probability density functions are normed positive functions. They form the basis for the *normalization laws* for coin identities to be studied shortly.

THEOREM 3.2. *For any $R \subset \mathbb{D}$ we have*

$$\int \mathbb{T}^R \, dR = 1 \tag{3.9}$$

THEOREM 3.3. *For any $(R, L) \in \mathcal{D} \otimes \mathcal{D}$ with $R \neq \emptyset$ we have*

$$\int \mathbb{T}_L^R \, dR = 1 \qquad (R \neq \emptyset) \tag{3.10}$$

Note that Theorem 3.2 is a special case of Theorem 3.3. Putting $L = \emptyset$ into (3.10) we get (3.9). We state both results as separate theorems due to their importance.

The results contained in the following theorem will be frequently used in the sequel.

THEOREM 3.4. *If A, B, C, D are exclusive subsets of $\mathbb{D}$, then we have*

$$\int \mathbb{T}[A]\mathbb{T}^B \, dB = \mathbb{T}[A] \tag{3.11}$$

$$\int \mathbb{T}[A]\mathbb{T}^{BC} \, dC = \mathbb{T}[A]\mathbb{T}^B \tag{3.12}$$

$$\int \mathbb{T}[AD]\,\mathbb{T}^{ABC} \, dC = \mathbb{T}[AD]\mathbb{T}^{AB} \tag{3.13}$$

$$\int \mathbb{T}_C^{AB} \, dB = \mathbb{T}_C^A \qquad (A \neq \emptyset) \tag{3.14}$$

3.3 The N-Law and the M-Law

Now we prove two important laws concerning certian types of coin itentities. These laws allow formal logic deduction from one coin equation (an equation relating various coins) to another coin equation. Since indepednece and conditional independence will be defined in terms of coin equations, these laws are thus of fundamental interest for formal reasoning about conditional independence.

THEOREM 3.5 (LAW OF NORMALIZATION). *(i) Let $(A, B) \in \mathcal{D} \otimes \mathcal{D}$ with $A \neq \emptyset$. Let $\bar{B} = \mathbb{D} \setminus B$ be the complementary set of B. Denote by $\mathbb{T}[\bar{B}] \in \mathbb{I}_{\bar{B}}$ an arbitrary coin free of B, then we have*

$$\mathbb{T}_B^A = \mathbb{T}[\bar{B}] \Rightarrow \mathbb{T}_B^A = \mathbb{T}^A \qquad (A \neq \emptyset) \tag{3.15}$$

(ii) Suppose that $A \neq \emptyset$, and A, B, C are mutually exclusive. Let $\mathbb{T}[\bar{C}] \in \mathbb{I}_{\bar{C}}$ be an arbitrary coin free of CB. Then we have

$$\mathbb{T}_{BC}^A = \mathbb{T}[\bar{C}] \Rightarrow \mathbb{T}_{BC}^A = \mathbb{T}_B^A \qquad (A \neq \emptyset) \tag{3.16}$$

(iii) Suppose that A, B, C are mutually exclusive then we have

$$\mathbb{T}^{ABC} = \mathbb{T}[\bar{B}]\mathbb{T}[\bar{A}] \Rightarrow \mathbb{T}_C^{AB} = \mathbb{T}_C^A\mathbb{T}_C^B \tag{3.17}$$

In particular when $C = \emptyset$ in (3.17) we get

$$\mathbb{T}^{AB} = \mathbb{T}[\bar{B}]\mathbb{T}[\bar{A}] \Rightarrow \mathbb{T}^{AB} = \mathbb{T}^A\mathbb{T}^B .$$

For convenience, we shall refer to the *Law of Normalization* as the *N-Law*. Using the N-Law we can pass from an 'ambiguous' coin equation to an 'exact' coin equation. This is useful, for instance, when a lot of atom coins are entering into a coin equation but we are only interested in relations concerning coins with indices being small subsets of $\mathbb{D}$. Those nuisance coins can be treated in a similar way as a proportionality ot normalizing constant in operations with integrating probability density functions.

The next *Law of Marginalization*, or the *M-Law*, is useful, in a systematical manner, to deduce sets of marginal coin relations from a larger coin identity.

THEOREM 3.6 (LAW OF MARGINALIZATION). *If A, B, C are exclusive subsets of $\mathbb{D}$. Then we have*

$$\pi_C^{AB} = \pi_C^A \pi_C^B \Rightarrow \pi_C^{ab} = \pi_C^a \pi_C^b, \quad \forall a \subset A, \forall b \subset B \tag{3.18}$$

In particular, we have

$$\pi^{AB} = \pi^A \pi^B \Rightarrow \pi^{ab} = \pi^a \pi^b, \quad \forall a \subset A, \forall b \subset B \tag{3.19}$$
$$\pi_C^{AB} = \pi_C^A \pi_C^B \Rightarrow \pi_C^{ij} = \pi_C^i \pi_C^j, \quad \forall i \in A, \forall j \in B \tag{3.20}$$

The following result will also be very useful in manipulating the relations on conditional independence involving a set of random variables.

THEOREM 3.7. *If A, B, C are mutually exclusive, and $\pi[\bar{C}]$ denote a coin independent of C, then*

$$\pi^{ABC} = \pi[\bar{C}]\pi^{BC} \Rightarrow \pi^{AB} = \pi[\bar{C}]\pi^B \tag{3.21}$$

4 Independence

4.1 Coin Equivalence

We first give a formal definition for coins restricted to a subset $A \subset \mathbb{D}$.

DEFINITION 4.1 (MARGINAL COINS). *We call π_L^R a* marginal atom coin *of A if $R \subset A, L \subset A$.*

A coin π is said a marginal coin *of A if π is the product of some finite sequence of marginal atom coins of A. The set of all marginal coins of A is denoted by Π_A.*

THEOREM 4.1 (MARGINAL COIN GROUP). *Let $A \subset \mathbb{D}$. Then Π_A is a subgroup of Π. We shall refer to Π_A as the* marginal group *of A.*

REMARK 4.1. *Note that $\Pi_\emptyset = \{1\}$ is also a subgroup of Π.*

The marginal group $\Pi_A \in \Pi$ introduces a natural *equivalence relation* in Π.

DEFINITION 4.2 (EQUIVALENT COINS). *Let Π_A be the marginal coin group of A. Two coins $\pi, \hat{\pi} \in \Pi$ are said to be equivalent with respect to A, written as $\pi \overset{A}{\sim} \hat{\pi}$, if $\pi\hat{\pi}^{-1} \in \Pi_A$. That is*

$$\pi \overset{A}{\sim} \hat{\pi} \Leftrightarrow \pi\hat{\pi}^{-1} \in \Pi_A \tag{4.1}$$

If we let $\mathbb{T}[A]$ to denote an appropriate marginal coin in $\mathbb{I}_A$, then (4.1) can be written alternatively as

$$\mathbb{T} \overset{A}{\sim} \hat{\mathbb{T}} \Leftrightarrow \mathbb{T} = \mathbb{T}[A]\hat{\mathbb{T}} \tag{4.2}$$

The set of all coins equivalent to $\hat{\mathbb{T}}$ with respect to A is called, using standard group terminology, the *coset* of $\hat{\mathbb{T}}$ with respect to $\mathbb{I}_A$. Thus, using (4.2), the coset of equivalent coins of $\hat{\mathbb{T}}$ is given by $\hat{\mathbb{T}}\mathbb{I}_A \equiv \{\mathbb{T}[A]\hat{\mathbb{T}} \mid \mathbb{T}[A] \in \mathbb{I}_A\}$. The coset $\hat{\mathbb{T}}\mathbb{I}_A$ is also sometimes referred to as the *orbit* of $\hat{\mathbb{T}}$ *caused* by group $\mathbb{I}_A$. Note that since $(\mathbb{T}[A])^{-1} \in \mathbb{I}_A$, the condition in (4.2) can be equivalently written as $\hat{\mathbb{T}} = \mathbb{T}[A]\mathbb{T}$.

4.2 Independence

THEOREM 4.2. *if $A \cap B = \emptyset$ and $A \neq \emptyset, B \neq \emptyset$, then*

$$\mathbb{T}^{AB} = \mathbb{T}^A\mathbb{T}^B \quad \Leftrightarrow \quad \mathbb{T}^{AB} \overset{A}{\sim} \mathbb{T}^B \tag{4.3}$$

$$\Leftrightarrow \quad \mathbb{T}^{AB} \overset{B}{\sim} \mathbb{T}^A \tag{4.4}$$

Note that the right hand side of (4.3) corresponds to the condition $f(\omega_A, \omega_B) = f(\omega_A)f(\omega_B)$, a condition saying that ω_A is independent of ω_B. Thus it is natural to make the following definition.

DEFINITION 4.3 (INDEPENDENCE). *Let $A \cap B = \emptyset$ and $A \neq \emptyset, B \neq \emptyset$. We say that $\mathbb{T}^A$ is independent of $\mathbb{T}^B$ if and only if $\mathbb{T}^{AB} \overset{A}{\sim} \mathbb{T}^B$, or $\mathbb{T}^{AB} \overset{B}{\sim} \mathbb{T}^A$. That is, using Dawid's notation*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \quad \Leftrightarrow \quad \mathbb{T}^{AB} \overset{A}{\sim} \mathbb{T}^B \tag{4.5}$$

$$\Leftrightarrow \quad \mathbb{T}^{AB} \overset{B}{\sim} \mathbb{T}^A \tag{4.6}$$

Definition 4.3 says that $\mathbb{T}^A$ is independent of $\mathbb{T}^B$ if $\mathbb{T}^{AB}$ is in the orbit of the marginal coin $\mathbb{T}^A$ caused by the marginal group $\mathbb{I}_B$, or if $\mathbb{T}^{AB}$ is in the orbit of $\mathbb{T}^B$ caused by $\mathbb{I}_A$. It is convenient to use Definition 4.3 to investigate algebraic structures brought about by various independence relations among raising coins of $\mathbb{I}$.

Operationally, however, the results of Theorem 4.2 are of direct use. We thus give an alternative definition of independence using these results.

DEFINITION 4.4 (INDEPENDENCE). *Let $A \cap B = \emptyset$ and $A \neq \emptyset, B \neq \emptyset$. We say that $\mathbb{T}^A$ is independent of $\mathbb{T}^B$ if and only if $\mathbb{T}^{AB} = \mathbb{T}^A\mathbb{T}^B$, that is*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \Leftrightarrow \mathbb{T}^{AB} = \mathbb{T}^A\mathbb{T}^B \tag{4.7}$$

4.3 Properties of independence

THEOREM 4.3. *Let R and L be exclusive nonempty subsets of $\mathbb{D}$. The following equations are equivalent to one another.*

$$\mathbb{T}^{RL} = \mathbb{T}^R\mathbb{T}^L \tag{4.8}$$

$$\mathbb{T}^R_L = \mathbb{T}^R \tag{4.9}$$

$$\mathbb{T}^L_R = \mathbb{T}^L \tag{4.10}$$

Since (4.8) is the condition for independence $\mathbb{T}^R \perp\!\!\!\perp \mathbb{T}^L$, we thus may use any one of the equations (4.8)-(4.10) to check if $\mathbb{T}^R \perp\!\!\!\perp \mathbb{T}^L$. By the N-Law, $\mathbb{T}^R_L = \mathbb{T}[R]$ implies $\mathbb{T}^R_L = \mathbb{T}^R$. Similarly, $\mathbb{T}^L_R = \mathbb{T}[L]$ implies $\mathbb{T}^L_R = \mathbb{T}^L$. Thus, the sufficient and necessary conditions (4.9) and (4.10) for $\mathbb{T}^R \perp\!\!\!\perp \mathbb{T}^L$ can be weakened as follows.

THEOREM 4.4. *Let R and L be exclusive nonempty subsets of $\mathbb{D}$. Then*

$$\mathbb{T}^R \perp\!\!\!\perp \mathbb{T}^L \Leftrightarrow \mathbb{T}^R_L = \mathbb{T}[R] \tag{4.11}$$

$$\Leftrightarrow \mathbb{T}^L_R = \mathbb{T}[L] \tag{4.12}$$

where $\mathbb{T}[R] \in \mathbb{I}_R$ and $\mathbb{T}[L] \in \mathbb{I}_L$ denote some coins depending only on R and L respectively.

Note that multiplying both sides of $\mathbb{T}^R_L = \mathbb{T}[R]$ by $\mathbb{T}^L$, we get $\mathbb{T}^{RL} = \mathbb{T}^L\mathbb{T}[R]$, showing that $\mathbb{T}^{RL} \overset{R}{\sim} \mathbb{T}^L$. So $\mathbb{T}^R \perp\!\!\!\perp \mathbb{T}^L$. From this 'new' proof we can see that the definition for independence given in Definition 4.3 is more flexible than the definition given in Definition 4.4.

As a direct consequence of the M-Law, we obtain the following important result on independence.

THEOREM 4.5 (MARGINALIZATION). *if $A \cap B = \emptyset$, then*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \Rightarrow \mathbb{T}^{A_1} \perp\!\!\!\perp \mathbb{T}^{B_1} \quad (\forall A_1 \subset A, \forall B_1 \subset B) \tag{4.13}$$

In particular,

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \Rightarrow \mathbb{T}^a \perp\!\!\!\perp \mathbb{T}^b \quad (\forall a \in A, \forall b \in B) \tag{4.14}$$

Theorem 4.5 says that *join independence implies marginal independence*. The reverse of this theorem is however not true. That is, the conditions $\mathbb{T}^{A_1} \perp\!\!\!\perp \mathbb{T}^{B_1}$ for any subsets $A_1 \subset A$ and $B_1 \subset B$ do *not* imply the *join independence* $\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B$.

5 Conditional independence

5.1 Definition

DEFINITION 5.1 (CONDITIONAL INDEPENDENCE). *Let A, B, C be mutually disjoint subsets of $\mathbb{D}$. Let A, B be nonempty. Then $\mathbb{T}^A$ is said* conditionally independent *of $\mathbb{T}^B$ given $\mathbb{T}^C$ if and only if $\mathbb{T}^{AB}_C$ is equivalent to $\mathbb{T}^B_C$ with respect to $\mathbb{I}_{AC}$, that is*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^C \Leftrightarrow \mathbb{T}^{AB}_C \overset{AC}{\sim} \mathbb{T}^B_C \tag{5.1}$$

By the N-Law we immediately have

THEOREM 5.1. *Let A, B, C be mutually disjoint nonempty subsets of $\mathbb{D}$. Let A, B be nonempty. Then*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^C \Leftrightarrow \mathbb{T}^{AB}_C \overset{BC}{\sim} \mathbb{T}^A_C \tag{5.2}$$

We have defined the concept of conditional independece using the concept of coin equivalence. For three mutually disjoint and nonempty subsets A, B, C of $\mathbb{D}$, both $\mathbb{I}_{AC}$ and $\mathbb{I}_{BC}$ define a marginal subgroup in $\mathbb{I}$. These subgroups introduce different orbits in $\mathbb{I}$. We have that $\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^C$ if and only if the mixed coin $\mathbb{T}_C^{AB}$ is in the orbit of the mixed coin $\mathbb{T}_C^B$ caused by $\mathbb{I}_{AC}$, or $\mathbb{T}_C^{AB}$ is in the orbit of $\mathbb{T}_C^A$ caused by $\mathbb{I}_{BC}$.

Note that in both (5.1) and (5.2) we allow the possibility that $C = \emptyset$. When $C = \emptyset$ the necessary and sufficient conditions for conditional independece reduce to the corresponding conditions for independence. An obvious advantage of the coin algebra introduced in this paper is that we can study both concepts in the same framework. To emphasize this viewpoint we give a new definition for the coin independence.

DEFINITION 5.2 (INDEPENDENCE). *Let A, B be mutually disjoint nonempty subsets of $\mathbb{D}$. Then $\mathbb{T}^A$ is said to be independent of $\mathbb{T}^B$, written as $\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B$, if and only if $\mathbb{T}^A$ is conditionally independent of $\mathbb{T}^B$ given $\mathbb{T}^\emptyset \equiv 1$. That is,*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \Leftrightarrow \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^\emptyset \tag{5.3}$$

5.2 Properties

Now we give operationally convenient conditions for conditional independence. All these conditions may be regarded as coin equations. One condition, namely coin equation, can be transformed to another condition in a relatively mechnical way by using the properties of the coins, such as the R-Law, the L-Law, the N-Law, the M-Law, and so on. Theorem 5.2 gives equivalent necessary and sufficient conditions for $\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^C$.

THEOREM 5.2. *Let A, B, C be disjoint subsets of $\mathbb{D}$. Then the following coin equations are equivalent to one another.*

$$\mathbb{T}_C^{AB} = \mathbb{T}_C^A \mathbb{T}_C^B \tag{5.4}$$
$$\mathbb{T}^{ABC} = \mathbb{T}^{AC} \mathbb{T}_C^B \tag{5.5}$$
$$\mathbb{T}^{ABC} = \mathbb{T}^{BC} \mathbb{T}_C^A \tag{5.6}$$
$$\mathbb{T}_{BC}^A = \mathbb{T}_C^A \tag{5.7}$$
$$\mathbb{T}_{AC}^B = \mathbb{T}_C^B \tag{5.8}$$

From (5.4)-(5.8) we can also derive other equivalent conditions. For instance, multiplying both sides of (5.5) by $\mathbb{T}_B$ results in $\mathbb{T}_B^{AC} = \mathbb{T}^{AC} \mathbb{T}_C^B \mathbb{T}_B$, and so on. However, the expressions (5.4)-(5.8) are the most frequently used ones. From the above proof we can also see that if we have the condition (5.4) then all other conditions can be derived in a straightforward manner using algebraic properties of the coins.

The following theorem gives a seemingly weaker condition for conditional independence.

THEOREM 5.3 (FACTORIZATION). *Let $A \neq \emptyset, B \neq \emptyset, C$ be disjoint subsets of $\mathbb{D}$, then*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^C \Leftrightarrow \mathbb{T}_C^{AB} = \mathbb{T}[\bar{B}]\mathbb{T}[\bar{A}] \tag{5.9}$$
$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^C \Leftrightarrow \mathbb{T}^{ABC} = \mathbb{T}[\bar{B}]\mathbb{T}[\bar{A}] \tag{5.10}$$

where $\mathbb{T}[\bar{B}] \in \mathbb{I}_{\bar{B}}, \mathbb{T}[\bar{A}] \in \mathbb{I}_{\bar{A}}$.

The condition (5.10) says that $\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^C$ holds if and only if the coin $\mathbb{T}^{ABC}$ admit expressions which can factorize into two subexpressions, one being in $\mathbb{I}_{AC}$ and the other in $\mathbb{I}_{BC}$. In other words, *any* atom coin $\mathbb{T}$ in $\mathbb{I}_{ABC}$ simultaneously invloving a subset of A and a subset of B cannot be a prime coin. More formally, we have the following theorem, which is an extension of Thorem 4.5 on independence.

THEOREM 5.4 (MARGINALIZATION). *Let $A \neq \emptyset, B \neq \emptyset, C$ be disjoint subset cf $\mathbb{D}$, then*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^C \Rightarrow \mathbb{T}^{A_1} \perp\!\!\!\perp \mathbb{T}^{B_1} | \mathbb{T}^C \quad (\forall A_1 \subset A, \forall B_1 \subset B) \tag{5.11}$$

By the marginalization theorem we know that joint independence $\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^{BC}$ implies marginal independence $\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B$ and $\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^C$. The reverse is not true. The following Theorem says that pair-wise conditional independence implies joint independence and vise versa. In a sense this theorem complements the Simpson's Paradox.

THEOREM 5.5 (INTERSECTION). *if A, B, C, D are exclusive, then we have*

$$\left.\begin{array}{c} \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \mid \mathbb{T}^{CD} \\ \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^C \mid \mathbb{T}^{BD} \end{array}\right\} \Leftrightarrow \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^{BC} \mid \mathbb{T}^D \tag{5.12}$$

COROLLARY 5.1. *if A, B, C are mutually exclusive and nonempty, then we have*

$$\left.\begin{array}{c} \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \mid \mathbb{T}^C \\ \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^C \mid \mathbb{T}^B \end{array}\right\} \Leftrightarrow \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^{BC} \tag{5.13}$$

The following theorem gives seemingly weaker sufficient and necessary conditions of the joint conditional independence $\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^{BC} | \mathbb{T}^D$. This theorem is sometimes referred to as the *contraction* property of conditional independence (Pear (2000), p.11).

THEOREM 5.6 (CONTRACTION). *if A, B, C, D are mutually exclusive and nonempty, then we have*

$$\left.\begin{array}{c} \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \mid \mathbb{T}^{CD} \\ \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^C \mid \mathbb{T}^D \end{array}\right\} \Leftrightarrow \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^{BC} \mid \mathbb{T}^D \tag{5.14}$$

COROLLARY 5.2. *if A, B, C are mutually exclusive and nonempty, then we have*

$$\left.\begin{array}{c} \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \mid \mathbb{T}^C \\ \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^C \end{array}\right\} \Leftrightarrow \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^{BC} \tag{5.15}$$

THEOREM 5.7 (WEAK UNION). *if A, B, C, D are mutually exclusive, then*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^{BC} | \mathbb{T}^D \quad \Rightarrow \quad \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B | \mathbb{T}^{CD} \tag{5.16}$$

THEOREM 5.8 (MIXING). *If A, B, C, D are mutually exclusive, then*

$$\left\{\begin{array}{l}\pi^A \perp\!\!\!\perp \pi^{BD} \mid \pi^C \\ \pi^B \perp\!\!\!\perp \pi^D \mid \pi^C\end{array}\right. \Longrightarrow \pi^{AD} \perp\!\!\!\perp \pi^B \mid \pi^C \tag{5.17}$$

The property of mixing rule was given by Dawid (1979). Examing the proof of Theorem 5.8 we immediately have

THEOREM 5.9 (STRONG-MIXING). *If A, B, C, D are mutually exclusive, then*

$$\left\{\begin{array}{l}\pi^A \perp\!\!\!\perp \pi^{BD} \mid \pi^C \\ \pi^B \perp\!\!\!\perp \pi^D \mid \pi^C\end{array}\right. \Longleftrightarrow \left\{\begin{array}{l}\pi^{AD} \perp\!\!\!\perp \pi^B \mid \pi^C \\ \pi^A \perp\!\!\!\perp \pi^D \mid \pi^C\end{array}\right. \tag{5.18}$$

The following property is known as the *chaining rule* (Lauritzen, 1982).

THEOREM 5.10 (CHAINING RULE). *If A, B, C, D are mutually exclusive, then*

$$\left\{\begin{array}{l}\pi^A \perp\!\!\!\perp \pi^B \mid \pi^C \\ \pi^{AC} \perp\!\!\!\perp \pi^D \mid \pi^B\end{array}\right. \Longrightarrow \pi^A \perp\!\!\!\perp \pi^D \mid \pi^C \tag{5.19}$$

Repeated use of the intersection theorem 5.5 gives the following result, which is of particulaly use in graphical modelling.

THEOREM 5.11 (SEPERATION THEOREM). *If A, B, C, D, S are mutually exclusive, then we have*

$$\left.\begin{array}{l}\pi^A \perp\!\!\!\perp \pi^C \mid \pi^{BDS} \\ \pi^A \perp\!\!\!\perp \pi^D \mid \pi^{BCS} \\ \pi^B \perp\!\!\!\perp \pi^C \mid \pi^{ADS} \\ \pi^B \perp\!\!\!\perp \pi^D \mid \pi^{ACS}\end{array}\right\} \Leftrightarrow \pi^{AB} \perp\!\!\!\perp \pi^{CD} \mid \pi^S \tag{5.20}$$

Using the M-Law and the seperation theorem we have the following result.

COROLLARY 5.3. *Suppose that A, B, C, D, S are exclusive, and the following hold*

$$\begin{array}{l}\pi^A \perp\!\!\!\perp \pi^C \mid \pi^{BDS} \\ \pi^A \perp\!\!\!\perp \pi^D \mid \pi^{BCS} \\ \pi^B \perp\!\!\!\perp \pi^C \mid \pi^{ADS} \\ \pi^B \perp\!\!\!\perp \pi^D \mid \pi^{ACS}\end{array} \tag{5.21}$$

Then for any subsets $E \subset A \cup B, F \subset C \cup D$, we have

$$\pi^E \perp\!\!\!\perp \pi^F \mid \pi^S \tag{5.22}$$

In particular, we have

$$\begin{array}{l}\pi^A \perp\!\!\!\perp \pi^C \mid \pi^S \\ \pi^A \perp\!\!\!\perp \pi^D \mid \pi^S \\ \pi^B \perp\!\!\!\perp \pi^C \mid \pi^S \\ \pi^B \perp\!\!\!\perp \pi^D \mid \pi^S\end{array} \tag{5.23}$$

DEFINITION 5.3 (MUTUAL INDEPENDENCE). *Let $A_1, \cdots, A_n$ be mutually exclusive nonempty subsets of $\mathbb{D}$. Then coins $\mathbb{T}^{A_1}, \cdots, \mathbb{T}^{A_n}$ are said mutually independent of each other, written as $\mathbb{T}^{A_1} \perp\!\!\!\perp \cdots \perp\!\!\!\perp \mathbb{T}^{A_n}$, if*

$$\mathbb{T}^{A_1 \cdots A_n} = \mathbb{T}^{A_1} \cdots \mathbb{T}^{A_n} \tag{5.24}$$

The following theorem gives sufficient and necessary conditions for mutual independence. Although it is stated in terms of three subsets, we can treat any number of subsets by recursively using it.

THEOREM 5.12. *If A, B, C are mutually exclusive nonempty subsetd of $\mathbb{D}$, then*

$$\mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \perp\!\!\!\perp \mathbb{T}^C \Leftrightarrow \begin{array}{c} \mathbb{T}^B \perp\!\!\!\perp \mathbb{T}^C \mid \mathbb{T}^A \\ \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^C \mid \mathbb{T}^B \\ \mathbb{T}^A \perp\!\!\!\perp \mathbb{T}^B \mid \mathbb{T}^C \end{array} \tag{5.25}$$

DEFINITION 5.4 (INDEPENDENT MODEL). *Let $\mathbb{D} = \{1, 2, \cdots, d\}$. The coin group $\mathbb{I}(\mathbb{D})$ is called an* independent model *if*

$$\mathbb{T}^{12\cdots d} = \mathbb{T}^1 \mathbb{T}^2 \cdots \mathbb{T}^d \tag{5.26}$$

The following theorem gives an important characterization of the independent model.

THEOREM 5.13 (CHARACTERIZATION OF INDEPENDENT MODEL). *A coin group $\mathbb{I}$ is an independent model if and only if there exists no prime coin.*

A model is a characterized coin group. Since in a null model every raising coin is a prime coin, we see that the null model and the independent model consist of two extremes in the *model space*. The properties of the model space will be formally studied in later papers.

6 Separoid

In this section we show that the coin algebra satisfies the defining axioms of a separoid of Dawid (2001). The separoid includes several axiomatic systems, such as the orthgoids and graphoid, relevent for formal reasoning essentially involving the concept of *irrelevence* of information.

The following definition was invented by Dawid (2001).

DEFINITION 6.1 (SEPAROID). *Let $(S, \leq)$ be a join-semilattice. Let $\cdot \perp\!\!\!\perp \cdot \mid \cdot$ be a ternary relation on S. Then $(S, \leq, \perp\!\!\!\perp)$ is a* separoid *if*

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

REMARK 6.1. *The above definition is slightly stronger than that given by Dawid (2001), who did not require the partial order of the semilattice to be anti-symmetric. Such an order is called a* quasiorder, *and when the anti-symmetry does hold for x and y we say that x and y are equivalent (instead of equal).*

REMARK 6.2. *None of the axioms P1-P5 in Definition 6.1 is* equational, *thus the description of a separoid $(S, \leq, \perp\!\!\!\perp)$ as given in Definition 6.1 does not constitute a universal algebra. It is possible to use the language of coins to redefine the separoid as a universal algebra so that every axiom is in equational form(conjecture).*

DEFINITION 6.2 (STRONG SEPAROID). *A separoid $(S, \leq, \perp\!\!\!\perp)$ is said to be a* strong separoid *if $(S, \leq)$ is a lattice and the following additional property holds*

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

Now we show that the relation of conditional independence derived from the coin algebra satisfies the axioms of a strong separoid. Let $\mathcal{D} = 2^{\mathbb{D}}$ be the power set of $\mathbb{D}$. Let $\leq$ be the usual set inclusion $\subseteq$. Then $(\mathcal{D}, \leq)$ froms a Boolean lattice. Before stating the theorem, we first note that the C-Axiom of the binary coin operation

$$\mathbb{T}_L^R = \mathbb{T}^{RL}\mathbb{T}_L \qquad (R \neq \emptyset)$$

is well defined for any R and L in $\mathcal{D}$ which may not be exclusive. For instance, since $AA = A$, we thus have

$$\mathbb{T}_A^A = \mathbb{T}^A\mathbb{T}_A = 1 \tag{6.27}$$

When $R \leq L$, we have

$$\mathbb{T}_L^R = \mathbb{T}^{RL}\mathbb{T}_L = \mathbb{T}^L\mathbb{T}_L = 1 \tag{6.28}$$

And, when $L \leq R$, we have

$$\mathbb{T}_L^R = \mathbb{T}^{RL}\mathbb{T}_L = \mathbb{T}^R\mathbb{T}_L \tag{6.29}$$

With this broader interpretation of the coin operation we now show that

THEOREM 6.1. *Let $(\mathcal{D}, \leq)$ be the Boolean lattice. Define the (partial) ternary relation $x \perp\!\!\!\perp y|z$ in $\mathcal{D}$ if the coin equation $\mathbb{T}_z^{xy} = \mathbb{T}_z^x\mathbb{T}_z^y$ holds, where $x \wedge y = \emptyset$. Then $(\mathcal{D}, \leq, \perp\!\!\!\perp)$ is a strong separoid, that is, P1-P6 hold.*

Proof. For P1, we want to show that

$$x \perp\!\!\!\perp y|x$$

or, in terms of coins

$$\mathbb{T}_x^{xy} = \mathbb{T}_x^x\mathbb{T}_x^y \tag{6.30}$$

Since $\mathbb{T}_x^{xy} = \mathbb{T}^{xyx}\mathbb{T}_x = \mathbb{T}_x^y = \mathbb{T}^{xy}\mathbb{T}_x = \mathbb{T}_x^y$ by (6.29), and $\mathbb{T}_x^x = 1$ by (6.27), So

$$\mathbb{T}_x^{xy} = \mathbb{T}_x^y = \mathbb{T}_x^x\mathbb{T}_x^y$$

proving (6.30)

For P2, we want to show that

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

or, in terms of coins

$$\mathbb{T}_z^{xy} = \mathbb{T}_z^x \mathbb{T}_z^y \implies \mathbb{T}_z^{yx} = \mathbb{T}_z^y \mathbb{T}_z^x \tag{6.31}$$

Since $\mathbb{T}_z^{xy} = \mathbb{T}_z^{yx}$, so (6.31) follows from the commutativity of coin multiplication.

For P3, we want to show that

$$x \perp\!\!\!\perp y|z, w \leq y \implies x \perp\!\!\!\perp w|z$$

where $x \wedge y = x \cap y = \emptyset$. Since $w \leq y$, there exists a unique $\bar{w}$ so that $y = w \vee \bar{w} = w\bar{w}$, where $\bar{w} = y \setminus w$. So what we need to show is that

$$x \perp\!\!\!\perp w\bar{w}|z \implies x \perp\!\!\!\perp w|z$$

or, in terms of coins

$$\mathbb{T}_z^{xw\bar{w}} = \mathbb{T}_z^x \mathbb{T}_z^{w\bar{w}} \implies \mathbb{T}_z^{xw} = \mathbb{T}_z^x \mathbb{T}_z^w \tag{6.32}$$

Since $xw\bar{w} = \emptyset$ so we have

$$\begin{aligned}
\mathbb{T}_z^{xw} &= \int \mathbb{T}_z^{xw\bar{w}}\, d\bar{w} \\
&= \int \mathbb{T}_z^x \mathbb{T}_z^{w\bar{w}}\, d\bar{w} \\
&= \mathbb{T}_z^x \int \mathbb{T}_z^{w\bar{w}}\, d\bar{w} \\
&= \mathbb{T}_z^x \mathbb{T}_z^w
\end{aligned}$$

proving (6.32).

For P4, we want to show that

$$x \perp\!\!\!\perp y|z, w \leq y \implies x \perp\!\!\!\perp y|(z \vee w)$$

where $x \wedge y = x \cap y = \emptyset$. Since $w \leq y$ iff $y = w \vee \bar{w} = w\bar{w}$, and $z \vee w = zw$, what we want to show is that

$$x \perp\!\!\!\perp w\bar{w}|z \implies x \perp\!\!\!\perp w\bar{w}|(wz)$$

or, in terms of coins

$$\mathbb{T}_z^{xw\bar{w}} = \mathbb{T}_z^x \mathbb{T}_z^{w\bar{w}} \implies \mathbb{T}_{wz}^{xw\bar{w}} = \mathbb{T}_{wz}^x \mathbb{T}_{wz}^{w\bar{w}} \tag{6.33}$$

Now assume that $\mathbb{T}_z^{xw\bar{w}} = \mathbb{T}_z^x \mathbb{T}_z^{w\bar{w}}$ hold. Since $x \wedge y = \emptyset$, in proving P3 we have had $\mathbb{T}_z^{xw} = \mathbb{T}_z^x \mathbb{T}_z^w$, which is equivalent to $\mathbb{T}_{wz}^x = \mathbb{T}_z^x$ by acting on $\mathbb{T}^x \mathbb{T}_{wz}$ (that is, multiplying both sides of $\mathbb{T}_z^{xw} = \mathbb{T}_z^x \mathbb{T}_z^w$ by the coin $\mathbb{T}^x \mathbb{T}_{wz}$)

On the other hand, $\Pi_z^{xw\bar{w}} = \Pi_z^x \Pi_z^{w\bar{w}}$ can be equivalently transformed as follows

$$\begin{aligned} & \Pi_z^{xw\bar{w}} = \Pi_z^x \Pi_z^{w\bar{w}} \\ \overset{\Pi^z}{\Longleftrightarrow} \quad & \Pi^{xw\bar{w}z} = \Pi_z^x \Pi^{w\bar{w}z} \\ \overset{\Pi_{wz}}{\Longleftrightarrow} \quad & \Pi_{wz}^{x\bar{w}} = \Pi_z^x \Pi_{wz}^{\bar{w}} \\ \Longleftrightarrow \quad & \Pi_{wz}^{x\bar{w}} = \Pi_{wz}^x \Pi_{wz}^{\bar{w}} \end{aligned}$$

Since we have

$$\Pi_{wz}^{w\bar{w}} = \Pi^{w\bar{w}wz}\Pi_{wz} = \Pi^{wz\bar{w}}\Pi_{wz} = \Pi_{wz}^{\bar{w}}$$

and similarly $\Pi_{wz}^{xw\bar{w}} = \Pi_{wz}^{x\bar{w}}$, so $\Pi_{wz}^{xw\bar{w}} = \Pi_{wz}^x \Pi_{wz}^{w\bar{w}}$ holds, proving (6.33).

For P5, we want to show that

$$x \perp\!\!\!\perp y|z\,, x \perp\!\!\!\perp w|(y \vee z) \implies x \perp\!\!\!\perp (y \vee w)|z$$

or, in terms od coin indeities

$$\Pi_z^{xy} = \Pi_z^x \Pi_z^y\,, \Pi_{yz}^{xw} = \Pi_{yz}^x \Pi_{yz}^w \implies \Pi_z^{xyw} = \Pi_z^x \Pi_z^{yw} \tag{6.34}$$

Acting Π^{yz} on $\Pi_{yz}^{xw} = \Pi_{yz}^x \Pi_{yz}^w$ gives $\Pi^{xwyz} = \Pi_{yz}^x \Pi^{wyz}$, and acting $\Pi^z \Pi_{yz}$ on $\Pi_z^{xy} = \Pi_z^x \Pi_z^y$ gives $\Pi_{yz}^x = \Pi_z^x$. So we have $\Pi^{xwyz} = \Pi_z^x \Pi^{wyz}$, which when acted on by Π_z gives $\Pi_z^{xyw} = \Pi_z^x \Pi_z^{yw}$, proving (6.34).

P6 is a generalization of the Intersection Axiom of the graphoids. We give a detailed proof here using the coin axioms. The essential property used is the N-Law. What we want to prove is that

$$\begin{matrix} x \perp\!\!\!\perp y|z\,, & z \le y \\ x \perp\!\!\!\perp y|w\,, & w \le y \end{matrix} \implies x \perp\!\!\!\perp y|(z \wedge w)$$

Note that since $z \le y$ and $w \le y$ so $z \wedge w \le y$. The first step is to uniquely decompose y as follows

$$y = qz_1cw_1 = q \vee z_1 \vee c \vee w_1$$

where

$$q = y \setminus (z \vee w)\,, c = z\wedge\,, z_1 = z \setminus c\,, w_1 = w \setminus c$$

So doing, what we want to prove is that

$$x \perp\!\!\!\perp qz_1cw_1|z_1c \;\&\; x \perp\!\!\!\perp qz_1cw_1|w_1c \implies x \perp\!\!\!\perp qz_1cw_1|c$$

or, in terms of coin identities

$$\begin{matrix} \Pi_{z_1c}^{xqz_1cw_1} = \Pi_{z_1c}^x \Pi_{z_1c}^{qz_1cw_1} \\ \Pi_{w_1c}^{xqz_1cw_1} = \Pi_{w_1c}^x \Pi_{w_1c}^{qz_1cw_1} \end{matrix} \implies \Pi_c^{xqz_1cw_1} = \Pi_c^x \Pi_c^{qz_1cw_1} \tag{6.35}$$

To show (6.35), by the absorption law,

$$\Pi_x^{xy} = \Pi_x^y$$

we only need to show

$$\left.\begin{array}{l} \Pi^{xqw_1}_{z_1c} = \Pi^x_{z_1c}\Pi^{qw_1}_{z_1c} \\ \Pi^{xqz_1}_{w_1c} = \Pi^x_{w_1c}\Pi^{qz_1}_{w_1c} \end{array}\right. \Longrightarrow \Pi^{xqz_1w_1}_c = \Pi^x_c\Pi^{qz_1w_1}_c \tag{6.36}$$

Acting Π^{z_1c} on $\Pi^{xqw_1}_{z_1c} = \Pi^x_{z_1c}\Pi^{qw_1}_{z_1c}$ and Π^{w_1c} on $\Pi^{xqz_1}_{w_1c} = \Pi^x_{w_1c}\Pi^{qz_1}_{w_1c}$ gives

$$\Pi^{xqw_1z_1c} = \Pi^x_{z_1c}\Pi^{qw_1z_1c}, \quad \Pi^{xqz_1w_1c} = \Pi^x_{w_1c}\Pi^{qz_1w_1c}$$

Comparing the two equations we must have

$$\Pi^x_{z_1c} = \Pi^x_{w_1c}$$

implying that $\Pi^x_{z_1c}$ is a coin free of both z_1 and w_1. So we may write

$$\Pi^x_{z_1c} = \Pi^x_{w_1c} = \Pi[\bar{z}_1 \wedge \bar{w}_1] \tag{6.37}$$

where $\bar{z}_1 = \mathbb{D} \setminus z_1$ and $\bar{w}_1 = \mathbb{D} \setminus w_1$. Applying the N-Law to (6.37), we have

$$\Pi^x_{z_1c} = \Pi^x_c$$

which implies that

$$\Pi^{xqw_1z_1c} = \Pi^x_c\Pi^{qw_1z_1c}$$

Acting Π_c upon this identity then gives the r.h.s. of (6.36), proving P6. □

References

[1] K. Takeuchi, M. Akahira (1975), Characterizations of prediction sufficiency (adequacy) in terms of risk functions, *The Annals of Statistics*, 1018–1024.

[2] A.P. Dawid (1979), Conditional independence in statistical theory (with Discussion), *Journal of the Royal Statistical Society, Series B* **41**, 1–31.

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

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

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

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

[7] J. Pearl (1988), *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Mateo, CA.

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

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