Annals of Mathematics and Artificial Intelligence **32:** 335–372, 2001.

Separoids: A mathematical framework for conditional independence and irrelevance

A.P. Dawid

Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK
E-mail: dawid@stats.ucl.ac.uk

We introduce an axiomatic definition of a mathematical structure that we term a *separoid*. We develop some general mathematical properties of separoids and related axiom systems, as well as connections with other mathematical structures, such as distributive lattices, Hilbert spaces, and graphs. And we show, by means of a detailed account of a number of models of the separoid axioms, how the concept of separoid unifies a variety of notions of 'irrelevance' arising out of different formalisms for representing uncertainty in Probability, Statistics, Artificial Intelligence, and other fields.

Keywords: basic separoid, belief function, belief independence, covariance independence, directed Markov, distributive, graphoid, hyper independence, irrelevance, join-orthogonal, lattice, linear independence, Markov, meta independence, modular, natural independence, orthogonal independence, orthogonoid, possibility function, probabilistic independence, qualitative independence, semi-graphoid, semilattice, separoid, statistical independence, strong orthogonoid, strong separoid, submodular, undirected Markov, variation independence

Introduction

The axiomatic treatment of probabilistic conditional independence was introduced by Dawid [16], following informal use in [15,23]; and developed and applied in a series of papers [17–20,24–27,40]. Spohn [58] developed essentially the same axiom system in a distinct logical context, while a closely related system intended for both probabilistic and graphical applications was constructed by Pearl and Paz [48]. Since then similar axiomatic systems, generally intended to explicate some concept of 'irrelevance', have been developed for a variety of different uncertainty formalisms. For an overview, see [21].

In this paper we take a step back, and attempt to abstract a still more fundamental mathematical structure underlying all these concepts and applications. We identify this as the *separoid*, a three-place relation on a join semilattice, subject to a system of five axioms.

The paper is divided into three somewhat distinct parts. Part I sets out the abstract mathematical definitions of separoid and strong separoid, and some related structures such as orthogonoids and semi-graphoids. It also contains some general theory relating to the concept of 'basic separoid': one in which the separoid relation is consistent with the underlying order.

Part II shows how separoids arise naturally in a number of purely mathematical contexts. In particular, concepts such as distributivity and modularity of a lattice have simple expressions in terms of separoid properties.

Finally, part III surveys a variety of specific examples of separoids that have been used to express irrelevance in different formal frameworks for representing uncertainty, including Probability Theory, Statistics, Possibility Theory, and Belief Functions.

Part I. Abstract structure

In this part we describe the mathematical structures of the *separoid* and *strong separoid,* and develop some purely algebraic aspects. The general definitions are introduced and some simple properties illustrated in section 1. In section 1.1 we further observe that a separoid can always be factored over the kernel of its underlying quasiorder. In section 1.2 two standard algebraic constructions are applied to separoids: the transfer of the separoid structure by means of a semilattice homomorphism, and the construction of the least separoid containing a given ternary relation. In section 2 we show the existence in any separoid of a natural congruence: a separoid is called *basic* if this congruence is trivial.

1. Separoids

Let $\mathcal{S}$, with elements denoted by $x, y, \ldots$, be a set equipped with a quasiorder $\leqslant$: that is, $\leqslant$ is *reflexive* ($x \leqslant x$, all $x \in \mathcal{S}$) and *transitive* ($x \leqslant y$ and $y \leqslant z \Rightarrow x \leqslant z$), but not necessarily anti-symmetric. If $\leqslant$ is also anti-symmetric, i.e., $x \leqslant y$ and $y \leqslant x \Rightarrow x = y$, it is a (*regular*) *order*. However, throughout this paper we do not assume this unless explicitly stated. In particular, this applies to our definitions of 'poset', 'lattice', 'semilattice', etc., which are thus slightly weaker than usual. In general, we write $x \doteq y$ if $x \leqslant y$ and $y \leqslant x$. This is an equivalence relation (the *kernel* of the quasiorder $\leqslant$), and we then call x and y *equivalent*.

Definition 1.1. Given a ternary relation $\cdot \perp\!\!\!\perp \cdot \mid \cdot$ on $\mathcal{S}$, we call $\perp\!\!\!\perp$ *a separoid* (*on* $(\mathcal{S}, \leqslant)$), or the triple $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ a *separoid,* if:

S1: $(\mathcal{S}, \leqslant)$ is a join semilattice,

and

$$
\begin{array}{lll}
\text{P1:} & x \perp\!\!\!\perp y \mid x, & \\
\text{P2:} & x \perp\!\!\!\perp y \mid z & \Rightarrow \quad y \perp\!\!\!\perp x \mid z, \\
\text{P3:} & x \perp\!\!\!\perp y \mid z \text{ and } w \leqslant y & \Rightarrow \quad x \perp\!\!\!\perp w \mid z, \\
\text{P4:} & x \perp\!\!\!\perp y \mid z \text{ and } w \leqslant y & \Rightarrow \quad x \perp\!\!\!\perp y \mid (z \vee w), \\
\text{P5:} & x \perp\!\!\!\perp y \mid z \text{ and } x \perp\!\!\!\perp w \mid (y \vee z) & \Rightarrow \quad x \perp\!\!\!\perp (y \vee w) \mid z.
\end{array}
$$

Condition S1 requires that any two elements $x, y \in \mathcal{S}$ should have a *least upper bound* (or *join*) in $\mathcal{S}$, denoted by $x \vee y$, with the property that $x \leqslant z$ and $y \leqslant z$ if and only if $x \vee y \leqslant z$. Then $x \vee y$ is unique up equivalence. P4 and P5 should be interpreted as applying to all possible equivalent choices for the join terms appearing in them.

It is readily seen that, if $\perp\!\!\!\perp$ is a separoid on $(\mathcal{S}, \leqslant)$ and $\mathcal{S}' \subseteq \mathcal{S}$ is a sub-join semilattice of $(\mathcal{S}, \leqslant)$ so that $x, y \in \mathcal{S}' \Rightarrow x \vee y \in \mathcal{S}'$, then the restriction $\perp\!\!\!\perp'$ of $\perp\!\!\!\perp$ to $\mathcal{S}'$ will be a separoid on $(\mathcal{S}', \leqslant)$. We then call $(\mathcal{S}', \leqslant, \perp\!\!\!\perp')$ a *sub-separoid* of $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$. This applies in particular if we take $\mathcal{S}'$ to be the set of all joins of collections of elements taken from some fixed finite subset $\{x_1, \dots, x_N\}$ of $\mathcal{S}$.

We also introduce the *strong separoid*:

Definition 1.2. The triple $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ is a *strong separoid* if, in definition 1.1, S1 is strengthened to:

S1′: $(\mathcal{S}, \leqslant)$ is a lattice;

and, in addition to P1–P5, we require

P6: if $z \leqslant y$ and $w \leqslant y$, then
$x \perp\!\!\!\perp y \mid z$ and $x \perp\!\!\!\perp y \mid w \quad \Rightarrow \quad x \perp\!\!\!\perp y \mid (z \wedge w)$.

If $\perp\!\!\!\perp$ is a strong separoid on $(\mathcal{S}, \leqslant)$, and $\mathcal{S}' \subseteq \mathcal{S}$ is a sublattice of $(\mathcal{S}, \leqslant)$, then the restriction $\perp\!\!\!\perp'$ of $\perp\!\!\!\perp$ to $\mathcal{S}'$ will be a strong separoid on $(\mathcal{S}', \leqslant)$ – a *sub-strong separoid* of $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$.

The principal motivation for introducing the above abstract structures arises from the specific model in which $\perp\!\!\!\perp$ expresses the property of probabilistic conditional independence between random variables [16,19,48] – see section 6 below. However, as we shall see, there is a wide variety of other interesting and useful models of the separoid and strong separoid axioms. Many, though by no means all, of these have an interpretation in terms of some kind of ‘irrelevance’ concept within some formal uncertainty calculus. Indeed, if we understand $X \perp\!\!\!\perp Y \mid Z$ as asserting “Once Z is known, Y becomes irrelevant to X”, properties P1 and P3–P5, at least, appear intuitively compelling, however we interpret ‘irrelevance’.

1.1. Formal manipulation

A major advantage of the abstract axiomatic approach is that it makes it possible to derive further properties of $\perp\!\!\!\perp$ by formal manipulation of the properties P1–P5 (or P1–P6), rather than, e.g., by calling on more specific properties of probability distributions. As a simple example, we have:

Lemma 1.1. When P1–P5 hold,

$$x \perp\!\!\!\perp y \mid z \quad \text{if and only if} \quad (x \vee z) \perp\!\!\!\perp (y \vee z) \mid z \tag{1}$$

(for any choice of the join terms).

Proof. From P1 and P2, we have $x \perp\!\!\!\perp (y \vee z) \mid (y \vee z)$, whence, by P4, $x \perp\!\!\!\perp z \mid (y \vee z)$. If now $x \perp\!\!\!\perp y \mid z$, these two properties combine, using P5, to give $x \perp\!\!\!\perp (y \vee z) \mid z$. Using P1, we can similarly extend x to $x \vee z$.

Conversely, if $(x \vee z) \perp\!\!\!\perp (y \vee z) \mid z$, then $(x \vee z) \perp\!\!\!\perp y \mid z$ follows from P3, and then, using P2, we can similarly contract $(x \vee z)$ to x. □

Corollary 1.2. Suppose $x \perp\!\!\!\perp y \mid z$, and that $x' \doteq x,\ y' \doteq y,\ z' \doteq z$. Then $x' \perp\!\!\!\perp y' \mid z'$.

Proof. Applying P4 and P2 to (1) we obtain $x' \perp\!\!\!\perp y' \mid z$. This in turn implies $x' \perp\!\!\!\perp (y' \vee z) \mid z$. Since $z' \leqslant (y' \vee z)$, and we can take $z \vee z' = z'$, applying P4 yields $x' \perp\!\!\!\perp (y' \vee z) \mid z'$, and then a further application gives $x' \perp\!\!\!\perp y' \mid z'$. □

In algebraic terminology [10, chapter VI, section 14], we have shown that the kernel $\doteq$ of the quasiorder $\leqslant$ is a *congruence* for the relation $\perp\!\!\!\perp$. The relation $x \perp\!\!\!\perp y \mid z$ can thus be regarded as determining, and being determined by, a separoid relation $\dot{\perp\!\!\!\perp}$ on the quotient join semilattice $\mathcal{S}/\doteq$ of equivalence classes in $\mathcal{S}$ under $\doteq$, with the induced order $\dot{\leqslant}$. We call $(\mathcal{S}/\doteq, \dot{\leqslant}, \dot{\perp\!\!\!\perp})$ the *regular reduction* of $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ – by corollary 1.2, this still carries all the useful information about $\perp\!\!\!\perp$, but its order is now regular.

Example 1.1 (Markov chain). There is a wide variety of more substantial results that can be derived by such formal manipulation. The following [16] is just one of many such applications.

Consider five random variables, arranged in sequence as

$$X_1 \to X_2 \to X_3 \to X_4 \to X_5,$$

with a joint probability distribution P such that each variable is conditionally independent of all earlier variables, given only its immediate predecessor. If we interpret $\perp\!\!\!\perp$ as expressing probabilistic conditional independence under P, and $\leqslant$ as expressing functional dependence, so that $X \vee Y \doteq (X, Y)$, then P1–P5 hold (see section 6 below).

We are given:

(a) $X_3 \perp\!\!\!\perp X_1 \mid X_2$,

(b) $X_4 \perp\!\!\!\perp (X_1, X_2) \mid X_3$,

(c) $X_5 \perp\!\!\!\perp (X_1, X_2, X_3) \mid X_4$.

Then we can deduce

(d) $X_3 \perp\!\!\!\perp (X_1, X_5) \mid (X_2, X_4)$.

For (b) with P4, P3 $\Rightarrow$ (e): $X_4 \perp\!\!\!\perp X_1 \mid (X_2, X_3)$. (a) and (e) with P2, P5 $\Rightarrow$ (f): $(X_3, X_4) \perp\!\!\!\perp X_1 \mid X_2$. (f) with P2, P3, P4 $\Rightarrow$ (g): $X_3 \perp\!\!\!\perp X_1 \mid (X_2, X_4)$. (c) with P4, P2, P3 $\Rightarrow$ (h): $X_3 \perp\!\!\!\perp X_5 \mid (X_1, X_2, X_4)$. Finally, (g) and (h) with P5 $\Rightarrow$ (d).

We have thus deduced the 'nearest neighbour' property of a Markov chain: each variable is independent of all the others, given only its immediate neighbours. But we can further deduce – and this is a major advantage of the abstract axiomatic approach – that the corresponding property would hold for any other model of the separoid axioms. □

1.2. Algebraic concepts

Another advantage of the abstract algebraic approach is the availability of general concepts and tools for relating different models and building new models out of old ones. In particular, we introduce the following definitions.

Definition 1.3. Let $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ and $(\mathcal{S}', \leqslant', \perp\!\!\!\perp')$ be two separoids. A map $f : \mathcal{S} \to \mathcal{S}'$ is a *separoid homomorphism* if:

(i) It is a join semilattice homomorphism, i.e., $f(x \vee y) \doteq' f(x) \vee' f(y)$ (and so, in particular, $x \leqslant y \Rightarrow f(x) \leqslant' f(y)$).

(ii) $x \perp\!\!\!\perp y \mid z \Rightarrow f(x) \perp\!\!\!\perp' f(y) \mid f(z)$.

When both $\perp\!\!\!\perp$ and $\perp\!\!\!\perp'$ are strong separoids, we call f a *strong separoid homomorphism* if, in addition to (i) and (ii), we have:

(iii) $f(x \wedge y) \doteq' f(x) \wedge' f(y)$

so that f is a full lattice homomorphism.

Definition 1.4. Let $(\mathcal{S}, \leqslant)$ be a join semilattice, $(\mathcal{S}', \leqslant', \perp\!\!\!\perp')$ a separoid, and $f : \mathcal{S} \to \mathcal{S}'$ a join semilattice homomorphism. Define a relation $\perp\!\!\!\perp$ on $\mathcal{S}$ by:

$$x \perp\!\!\!\perp y \mid z \Leftrightarrow f(x) \perp\!\!\!\perp' f(y) \mid f(z). \tag{2}$$

It is readily seen that $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ is a separoid, and that f is a separoid homomorphism. We call $\perp\!\!\!\perp$ the separoid relation *induced by* $\perp\!\!\!\perp$ *and* f.

Note that if, further, f is a lattice homomorphism and $\perp\!\!\!\perp'$ a strong separoid, then so is $\perp\!\!\!\perp$ defined by (2), and f is then a strong separoid homomorphism.

Definition 1.5. Let $\{\perp\!\!\!\perp_\alpha\colon \alpha \in \mathcal{A}\}$ be a collection of separoid relations, all defined on the same join semilattice $(\mathcal{S}, \leqslant)$. Then their *intersection*, denoted by $\bigcap\{\perp\!\!\!\perp_\alpha\colon \alpha \in \mathcal{A}\}$, is the relation $\perp\!\!\!\perp$ given by:

$$x \perp\!\!\!\perp y \mid z \quad \Leftrightarrow \quad x \perp\!\!\!\perp_\alpha y \mid z \text{ for all } \alpha \in \mathcal{A}.$$

This is just the set-theoretic intersection of the relations $\{\perp\!\!\!\perp_\alpha\}$, regarded as subsets of $\mathcal{S} \times \mathcal{S} \times \mathcal{S}$. Clearly, if each $\perp\!\!\!\perp_\alpha$ is a separoid, then so their intersection. A corresponding property holds for strong separoids.

Now let $(\mathcal{S}, \leqslant)$ be a join semilattice, and ρ a three-place relation on $\mathcal{S}$, regarded as a subset of $\mathcal{S} \times \mathcal{S} \times \mathcal{S}$.

Definition 1.6. The separoid $\perp\!\!\!\perp(\rho)$ *generated by* ρ (or the *separoid closure of* ρ) on $(\mathcal{S}, \leqslant)$ is the intersection of all separoid relations on $(\mathcal{S}, \leqslant)$ containing ρ.

Then $\perp\!\!\!\perp(\rho)$ is the smallest separoid relation on $(\mathcal{S}, \leqslant)$ containing ρ. It can be equivalently defined as the set of all triples $(x, y, z) \in \mathcal{S} \times \mathcal{S} \times \mathcal{S}$ for which it is possible to derive $x \perp\!\!\!\perp y \mid z$, starting from the 'input list' $x \perp\!\!\!\perp y \mid z$ (all $(x, y, z) \in \rho$) and applying, a finite number of times, the 'inference rules' embodied in the axioms P1–P5.

Given a separoid $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$, and a subset ρ of $\perp\!\!\!\perp$, we always have $\perp\!\!\!\perp(\rho) \subseteq \perp\!\!\!\perp$.

Definition 1.7. We call ρ a (*separoid*) *spanning relation for* $\perp\!\!\!\perp$, or say that ρ (*separoid*) *spans* $\perp\!\!\!\perp$, if $\perp\!\!\!\perp(\rho) = \perp\!\!\!\perp$.

When $(\mathcal{S}, \leqslant)$ is a lattice, we can similarly define the *strong separoid generated by* ρ, and a *(strong separoid) spanning relation* for a strong separoid $\perp\!\!\!\perp$.

1.2.1. Span deduction

Clearly, if ρ spans $\perp\!\!\!\perp$, and we can show that, for some separoid of interest $\perp\!\!\!\perp'$ on $(\mathcal{S}, \leqslant)$, $x \perp\!\!\!\perp' y \mid z$ for all $(x, y, z) \in \rho$, we can then deduce that $x \perp\!\!\!\perp' y \mid z$ for all $(x, y, z) \in \perp\!\!\!\perp$. This is a simple yet powerful method of proving such properties: we term it *span deduction*. In particular, if f is a join semilattice homomorphism from $(\mathcal{S}, \leqslant)$ to $(\mathcal{S}', \leqslant', \perp\!\!\!\perp')$, and we can show that $f(x) \perp\!\!\!\perp' f(y) \mid f(z)$ for all $(x, y, z) \in \rho$, we can deduce $f(x) \perp\!\!\!\perp' f(y) \mid f(z)$ whenever $x \perp\!\!\!\perp y \mid z$.

2. Basic separoids

In any separoid, we have, by P1 and P3, $x \leqslant y \Rightarrow x \perp\!\!\!\perp x \mid y$. In this section we investigate the possibility of reversing this implication.

Definition 2.1. We call a separoid $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ *basic* if

$$x \perp\!\!\!\perp x \mid y \Rightarrow x \leqslant y. \tag{3}$$

Lemma 2.1. Suppose $(\mathcal{S}, \leqslant)$ is a lattice. Then $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ is basic if and only if

$$x \perp\!\!\!\perp y \mid z \Rightarrow (x \vee z) \wedge (y \vee z) \doteq z. \tag{4}$$

Note that we always have $(x \vee z) \wedge (y \vee z) \geqslant z$. Thus it is equivalent to replace $\doteq$ by $\leqslant$ in (4).

Proof. Suppose that $x \perp\!\!\!\perp y \mid z$. Then $(x \vee z) \perp\!\!\!\perp (y \vee z) \mid z$ by lemma 1.1. Defining $w := (x \vee z) \wedge (y \vee z)$, we have both $w \leqslant (x \vee z)$ and $w \leqslant (y \vee z)$, and so, by P3 and P2, $w \perp\!\!\!\perp w \mid z$. If now $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ is basic, we deduce $w \leqslant z$.

Conversely, suppose $x \perp\!\!\!\perp x \mid y$. If (4) holds, we deduce $x \vee y \doteq y$, i.e., $x \leqslant y$. □

Corollary 2.2. *Suppose* $\perp\!\!\!\perp$ is a basic separoid on a lattice $(\mathcal{S}, \leqslant)$. Then

$$x \perp\!\!\!\perp y \mid z \Rightarrow x \wedge y \leqslant z.$$

Lemma 2.3. Suppose $\perp\!\!\!\perp$ is a basic separoid on a lattice $(\mathcal{S}, \leqslant)$. Then

$$x \perp\!\!\!\perp y \mid x \wedge y \text{ and } z \leqslant y \Rightarrow (x \vee z) \wedge y \doteq (x \wedge y) \vee z.$$

Proof. That $z \leqslant y \Rightarrow (x \vee z) \wedge y \geqslant (x \wedge y) \vee z$ follows readily from the lattice properties. Also, by P4, $x \perp\!\!\!\perp y \mid x \wedge y$ and $z \leqslant y \Rightarrow x \perp\!\!\!\perp y \mid (x \wedge y) \vee z$. From this and lemma 1.1, we readily deduce $(x \vee z) \perp\!\!\!\perp y \mid (x \wedge y) \vee z$. The result now follows from corollary 2.2. □

In a basic separoid, we can fully recover the order from knowledge of the separoid relation $\perp\!\!\!\perp$. It follows that the properties S1 and P1–P5 (and P6 where relevant) could then all be expressed in terms of $\perp\!\!\!\perp$ alone; however, they do not look particularly elegant when this is done.

2.1. Basic and minimal representations

Here we show that any separoid can be regarded as basic, so long as we use an appropriate quasiorder.

Let $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ be a separoid. Define a relation $\preceq$ on $\mathcal{S}$ by:

$$x \preceq y \Leftrightarrow x \perp\!\!\!\perp x \mid y. \tag{5}$$

Lemma 2.4. $x \preceq y \Leftrightarrow x \perp\!\!\!\perp z \mid y$ for all $z \in \mathcal{S}$.

Proof. We only need show $\Rightarrow$. By P1, $x \perp\!\!\!\perp (y \vee z) \mid x$. So by P4 and P3, $x \perp\!\!\!\perp z \mid (x \vee y)$. Combining this with $x \perp\!\!\!\perp x \mid y$ using P5 gives $x \perp\!\!\!\perp (x \vee z) \mid y$, and hence, by P3, $x \perp\!\!\!\perp z \mid y$. □

Theorem 2.5. $\preceq$ is a quasiorder on $\mathcal{S}$.

Proof. P1 implies $x \perp\!\!\!\perp x \mid x$, i.e., $x \leqslant x$, so that $\preceq$ is reflexive. For transitivity, suppose $x \preceq y$, $y \preceq z$. By lemma 2.4, $x \perp\!\!\!\perp (x \vee z) \mid y$, whence $x \perp\!\!\!\perp x \mid (y \vee z)$; and (using P2) $y \perp\!\!\!\perp x \mid z$. Together these yield, using P5, $x \perp\!\!\!\perp (x \vee y) \mid z$, so that $x \perp\!\!\!\perp x \mid z$, i.e., $x \preceq z$. □

We call $\preceq$ the *natural quasiorder* on $\mathcal{S}$, and its kernel $\simeq$ (where $x \simeq y$ if $x \preceq y$ and $y \preceq x$) the *natural equivalence*. Clearly, $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ is basic if and only if $\preceq$ is identical with the original quasiorder $\leqslant$. In general, even if $\leqslant$ is a regular order, $\preceq$ need not be.

Lemma 2.6. If $x \leqslant y$ then $x \preceq y$.

Proof. From P1 and P2, $x \perp\!\!\!\perp y \mid y$, and then, if $x \leqslant y$, P3 yields $x \perp\!\!\!\perp x \mid y$. □

Corollary 2.7. If $x \doteq y$ then $x \simeq y$.

Theorem 2.8. $(\mathcal{S}, \preceq)$ is a join semilattice. If $x \smile y$ is a least upper bound of $\{x, y\}$ under $\preceq$, then $x \smile y \simeq x \vee y$.

Proof. It follows from lemma 2.6 that $z = x \vee y$ is an upper bound of $\{x, y\}$ under $\preceq$. Let w be any upper bound. Since $x \preceq w$, $x \perp\!\!\!\perp z \mid w$ by lemma 2.4, whence, since $y \leqslant z$, $x \perp\!\!\!\perp z \mid (y \vee w)$ by P4, again by lemma 2.6. Also, $y \preceq w \Rightarrow y \perp\!\!\!\perp z \mid w$. Combining these using P5 (and P2) gives $(x \vee y) \perp\!\!\!\perp z \mid w$, i.e., $z \perp\!\!\!\perp z \mid w$, so that $z \preceq w$. □

Corollary 2.9. $x \preceq y$ if and only if there exists z such that $x \leqslant z$ and $y \simeq z$.

Proof. "If" follows from lemma 2.6. Suppose $x \preceq y$, and take $z = x \vee y$. Clearly $x \leqslant z$. Also, $y \simeq x \smile y, \simeq z$ by theorem 2.8. □

In general we may not be able to perform a similar substitution for x.

Lemma 2.10.

$$x \perp\!\!\!\perp y \mid z \text{ and } w \preceq y \quad \Rightarrow \quad x \perp\!\!\!\perp w \mid z, \tag{6}$$

$$x \perp\!\!\!\perp y \mid z \text{ and } w \preceq y \quad \Rightarrow \quad x \perp\!\!\!\perp y \mid (w \vee z). \tag{7}$$

Proof. By lemma 2.4, $w \preceq y \Rightarrow w \perp\!\!\!\perp (x \vee z) \mid y$, so that $x \perp\!\!\!\perp w \mid (y \vee z)$ by P4 and P2. Combining this with $x \perp\!\!\!\perp y \mid z$ using P5 yields $x \perp\!\!\!\perp (y \vee w) \mid z$, whence (6) and (7) follow from P3 and P4. □

Theorem 2.11. Suppose $x \perp\!\!\!\perp y \mid z$, and let $x' \simeq x,\ y' \simeq y,\ z' \simeq z$. Then $x' \perp\!\!\!\perp y' \mid z'$.

Proof. To show that we can replace z by z', we note that, since $z' \preceq z$, we have $z' \perp\!\!\!\perp (x \vee y) \mid z$ by lemma 2.4, so that $y \perp\!\!\!\perp z' \mid (x \vee z)$ by P4 and P2. Combining with $y \perp\!\!\!\perp x \mid z$ using P5 gives $y \perp\!\!\!\perp (x \vee z') \mid z$, and thus, again by P4 and P2, $x \perp\!\!\!\perp y \mid (z \vee z')$. Also, since $z \preceq z'$, $x \preceq z \mid z'$, by lemma 2.4 and P2. We thus obtain, from P5, $x \perp\!\!\!\perp (y \vee z) \mid z'$, and hence, by P3, $x \perp\!\!\!\perp y \mid z'$.

Now $x \perp\!\!\!\perp y' \mid z'$ follows from (6), and then, using P2, we similarly obtain $x' \perp\!\!\!\perp y' \mid z'$. □

Theorem 2.12. $(\mathcal{S}, \preceq, \perp\!\!\!\perp)$ is a basic separoid.

Proof. For P1 and P2 there is nothing to show. P3 is equation (6) of lemma 2.10. P4 follows from equation (7) of lemma 2.10, together with theorems 2.8 and 2.11. Similarly P5 follows from the same property of $\leqslant$, on again using theorems 2.8 and 2.11. Finally, (3) holds by definition. □

We call $(\mathcal{S}, \preceq, \perp\!\!\!\perp)$ the *basic representation* of $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$.

We have shown that the natural equivalence $\simeq$ is in fact a congruence relation – the *natural congruence* – for $\perp\!\!\!\perp$. The quotient of $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ by $\simeq$ is identical with the regular reduction of $(\mathcal{S}, \preceq, \perp\!\!\!\perp)$. This still carries all the information necessary to apply $\perp\!\!\!\perp$, and provides the maximum possible reduction without loss of information about $\perp\!\!\!\perp$: we call it the *minimal representation* of $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$. We can render this representation concrete by forming a *transversal* $\mathcal{S}'$ of $\mathcal{S}$ under $\simeq$, containing exactly one member of each $\simeq$-equivalence class. Then $(\mathcal{S}', \preceq, \perp\!\!\!\perp)$ will be a sub-separoid of $(\mathcal{S}, \preceq, \perp\!\!\!\perp)$, containing all the information about $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$. (It need not, however, be the case that $(\mathcal{S}', \preceq, \perp\!\!\!\perp)$ is a sub-separoid of $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$, since $\mathcal{S}'$ need not be closed under $\vee$).

We can extend the above analysis to the case of a strong separoid:

Theorem 2.13. Suppose $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ is a strong separoid. Then a greatest lower bound $X \frown Y$ of $\{X, Y\}$ under $\preceq$ exists, and $X \frown Y \simeq X \wedge Y$.

Proof. Since $X \wedge Y \leqslant X$, $X \wedge Y \preceq X$, by lemma 2.6. So $X \wedge Y$ is a lower bound for $\{X, Y\}$ under $\preceq$. Now let W be any lower bound for $\{X, Y\}$ under $\preceq$. Since $W \preceq X$, $W \perp\!\!\!\perp Y \mid X$; and similarly $W \perp\!\!\!\perp X \mid Y$. From P6, $W \perp\!\!\!\perp (X \vee Y) \mid (X \wedge Y)$, whence, since $W \preceq X \vee Y$, $W \perp\!\!\!\perp W \mid (X \wedge Y)$ by (6), i.e., $W \preceq X \wedge Y$. The result follows. $\square$

Corollary 2.14. If $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ is a strong separoid then so is $(\mathcal{S}, \preceq, \perp\!\!\!\perp)$.

Again, the natural equivalence $\simeq$ is a congruence for $\perp\!\!\!\perp$, and we obtain a *minimal representation* of $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ from its quotient by $\simeq$, which will yield a basic strong separoid.

3. Alternative axioms systems

Here we introduce some other general mathematical structures, *orthogonoids* (section 3.1) and *semi-graphoids* (section 3.2), and show their close connections with separoids.

3.1. Orthogonoids

The following axioms generalise properties of perpendicularity (relative orthogonality) between possibly intersecting subspaces of an inner product space (see section 4.3 below).

Definition 3.1. Let $\perp$ be a two-place relation on a lattice $(\mathcal{S}, \leqslant)$. We call $\perp$ an *orthogonoid* (on $(\mathcal{S}, \leqslant)$), or the triple $(\mathcal{S}, \leqslant, \perp)$) an *orthogonoid*, if the following properties are satisfied:

G1: $x \leqslant y \Rightarrow x \perp y$,

G2: $x \perp y \Rightarrow y \perp x$,

G3: If $x \wedge y \leqslant z \leqslant y$, then $x \perp y \Rightarrow x \perp z$,

G4: If $x \wedge y \leqslant z \leqslant y$, then

$$x \perp y \Rightarrow \begin{cases} (x \vee z) \wedge y \doteq z \text{ and} & \text{(G4a)} \\ (x \vee z) \perp y, & \text{(G4b)} \end{cases}$$

G5: If $(x \vee z) \wedge y \doteq z$, then

$$x \perp z \text{ and } (x \vee z) \perp y \quad \Rightarrow \quad x \perp y.$$

Note that the condition for G5 implies that for G4.

We call $\perp$ a *strong orthogonoid* if, in addition,

G6: $x \perp z$ and $y \perp z \Rightarrow (x \wedge y) \perp z$.

Recall that a lattice $(\mathcal{S}, \leqslant)$ is *modular* if

$$z \leqslant y \Rightarrow (x \wedge y) \vee z \doteq (x \vee z) \wedge y.$$

In that case, the conditions on z in G4 and G5 coincide, and so the requirement (G4a) is automatically satisfied. Then we can combine G3–G5 into the single condition:

G7: If $x \wedge y \leqslant z \leqslant y$, then

$$x \perp z \text{ and } (x \vee z) \perp y \quad \Leftrightarrow \quad x \perp y.$$

There is a close connection between separoids and orthogonoids:

Theorem 3.1. Let $\perp\!\!\!\perp$ be a separoid relation on a lattice $(\mathcal{S}, \leqslant)$. Suppose that *either* the separoid is basic, *or* the lattice is modular. Define a relation $\perp$ on $\mathcal{S}$ by:

$$x \perp y \Leftrightarrow x \perp\!\!\!\perp y \mid (x \wedge y). \tag{8}$$

Then $(\mathcal{S}, \leqslant, \perp)$ is an orthogonoid. If $\perp\!\!\!\perp$ is a strong separoid, then $(\mathcal{S}, \leqslant, \perp)$ is a strong orthogonoid.

Proof. For G1, if $x \leqslant y$ then $x \perp\!\!\!\perp y \mid x$ by P1, and so, since $x \wedge y \doteq x$, $x \perp y$.

G2 follows immediately from P2.

For G3 and G4, we assume that $x \perp y$ and $x \wedge y \leqslant z \leqslant y$, noting that then

$$x \wedge y \doteq x \wedge z. \tag{9}$$

From $x \perp\!\!\!\perp y \mid (x \wedge y)$ we have $x \perp\!\!\!\perp z \mid (x \wedge y)$ by P3, whence $x \perp z$ follows from (9).

Also, by P4 we have $x \perp\!\!\!\perp y \mid \{(x \wedge y) \vee z\}$, i.e., $x \perp\!\!\!\perp y \mid z$, whence $(x \vee z) \perp\!\!\!\perp y \mid z$. If the lattice is modular, then

$$(x \vee z) \wedge y \doteq (x \wedge y) \vee z \doteq z. \tag{10}$$

Alternatively, if $\perp\!\!\!\perp$ is basic, (10) follows from lemma 2.3. Thus G4 holds in either case.

Suppose now the conditions on the left-hand side of G5 hold. From $(x \vee z) \perp\!\!\!\perp y \mid (x \vee z) \wedge y$ and $(x \vee z) \wedge y \doteq z$ we obtain $x \perp\!\!\!\perp y \mid z$. Also, $x \perp\!\!\!\perp z \mid (x \wedge z)$.

Combining these using P5 yields $x \perp\!\!\!\perp (y \vee z) \mid (x \wedge z)$, whence $x \perp\!\!\!\perp y \mid (x \wedge z)$ by P3, and thus, since $x \wedge y \doteq x \wedge z$, $x \perp y$.

Finally, suppose $x \perp z$ and $y \perp z$. Then $x \perp\!\!\!\perp z \mid (x \wedge z)$, whence $(x \wedge y) \perp\!\!\!\perp z \mid (x \wedge z)$; and similarly $(x \wedge y) \perp\!\!\!\perp z \mid (y \wedge z)$. If now P6 holds, we deduce $(x \wedge y) \perp\!\!\!\perp z \mid (x \wedge y \wedge z)$, i.e., G6. □

The following theorem goes in the converse direction.

Theorem 3.2. Let $\perp$ be an orthogonoid on the lattice $(\mathcal{S}, \leqslant)$. Define $\perp\!\!\!\perp$ by:

$$x \perp\!\!\!\perp y \mid z \Leftrightarrow \begin{cases} (x \vee z) \perp (y \vee z) \text{ and} \\ (x \vee z) \wedge (y \vee z) \doteq z. \end{cases} \tag{11}$$

Then $\perp\!\!\!\perp$ is a basic separoid on $(\mathcal{S}, \leqslant)$. If, moreover, $\perp$ is a strong orthogonoid, then $\perp\!\!\!\perp$ is a strong separoid.

Proof. For P1: $x \perp\!\!\!\perp y \mid x$ requires $x \perp (x \vee y)$, which holds by G1; and $x \wedge (x \vee y) \doteq x$, which is so.

P2 is immediate.

Now suppose $X \perp\!\!\!\perp Y \mid Z$, and $W \leqslant Y$. Define $x = X \vee Z$, $y = Y \vee Z$, $z = W \vee Z$. Then $x \perp y$, and $x \wedge y = Z \leqslant z \leqslant y$. From G3 we deduce $x \perp z$, and using (9) this is equivalent to $X \perp\!\!\!\perp W \mid Z$, i.e., P3.

Also, G4 gives $(x \vee z) \perp y$ and $(x \vee z) \wedge y \doteq z$. That is, $X \perp\!\!\!\perp Y \mid (W \vee Z)$, i.e., P4.

Suppose now $X \perp\!\!\!\perp W \mid (Y \vee Z)$ and $X \perp\!\!\!\perp Y \mid Z$. Define $x = X \vee Z$, $y = W \vee Y \vee Z$, $z = Y \vee Z$. Then $(x \vee z) \perp y$ and $x \perp z$. Also, $(x \vee z) \wedge y \doteq z \leqslant y$, and $x \wedge z \doteq Z$ – from which we readily deduce $x \wedge y \doteq x \wedge z \doteq Z$. From G5 we obtain $x \perp y$, i.e., $X \perp\!\!\!\perp (W \vee Y) \mid Z$, so we have P5.

We have thus shown that $\perp\!\!\!\perp$ is a separoid. Note further that $x \perp\!\!\!\perp x \mid y \Rightarrow (x \vee y) \wedge (x \vee y) = y \Rightarrow x \vee y = y \Rightarrow x \leqslant y$, so that $\perp\!\!\!\perp$ is basic.

Finally, suppose $Z \leqslant Y$, $W \leqslant Y$, $X \perp\!\!\!\perp Y \mid Z$ and $X \perp\!\!\!\perp Y \mid W$. Defining $x = X \vee Z$, $y = X \vee W$, $z = Y$, we have both $x \perp z$ and $y \perp z$, while $x \wedge z = Z$, $y \wedge z = W$. If now G6 holds, we deduce $(x \wedge y) \perp z$, i.e., $(x \wedge y) \perp\!\!\!\perp z \mid (x \wedge y \wedge z)$. Noting that $x \wedge y \wedge z = (x \wedge z) \wedge (y \wedge z) = Z \wedge W$, that $z = Y$, and that $X \leqslant x \wedge y$, and using P3 and P2 (already shown), we deduce $X \perp\!\!\!\perp Y \mid (Z \wedge W)$, i.e., P6 holds, and we have a strong separoid. □

3.2. *Semi-graphoids*

Suppose that $(\mathcal{S}, \leqslant)$ is a distributive lattice having a minimal element 0, and possessing relative complements: i.e., if $x \leqslant y$ there exists $x' \in \mathcal{S}$ such that $x \wedge x' \doteq 0$, and $x \vee x' = y$. Then x' is unique up to equivalence, and we denote it by $y \backslash x$. For general x, y, we define $y \backslash x$ to be $(x \vee y) \backslash x$, or equivalently $y \backslash (x \wedge y)$. If $\mathcal{S}$ possesses a maximal element, it is a Boolean algebra. Otherwise, it becomes one when restricted to $\{x\colon x \leqslant M\}$, for any $M \in \mathcal{S}$, where complementation is relative to M.

In the most usual application, $(\mathcal{S}, \leqslant)$ is $(2^{*U}, \subseteq)$, the lattice of all finite subsets of some space U, ordered by inclusion. We shall term this the *finitary case*. When U is finite (the *finite case*), this becomes the Boolean algebra $(2^U, \subseteq)$.

Definition 3.2. A three-place relation $\amalg$ on $\mathcal{S}$ is called a *semi-graphoid* (*on* $(\mathcal{S}, \leqslant)$) if it contains only pairwise disjoint triples, and:

Q1: $x \wedge y \doteq 0 \Rightarrow x \amalg 0 \mid y,$

Q2: $x \amalg y \mid z \Rightarrow y \amalg x \mid z,$

Q3: $x \amalg y \mid z$ and $w \leqslant y \Rightarrow x \amalg w \mid z,$

Q4: $x \amalg y \mid z$ and $w \leqslant y \Rightarrow x \amalg (y \setminus w) \mid (z \vee w),$

Q5: $x \amalg y \mid z$ and $x \amalg w \mid (y \vee z) \Rightarrow x \amalg (y \vee w) \mid z.$

Definition 3.3. A semi-graphoid is a *graphoid* if, in addition to Q1–Q5, it satisfies:

Q6: $x \amalg y \mid (z \vee w)$ and $x \amalg w \mid (z \vee y) \Rightarrow x \amalg (y \vee w) \mid z.$

For the finite case, the (semi-)graphoid axioms of definitions 3.2 and 3.3 were first given by Pearl and Paz [48].

Given $\amalg$, define a ternary relation $\perp\!\!\!\perp$ on $\mathcal{S}$ by:

$$x \perp\!\!\!\perp y \mid z \quad \Leftrightarrow \quad x \wedge y \leqslant z \text{ and } (x \setminus z) \amalg (y \setminus z) \mid z. \tag{12}$$

It is readily verified that, if $\amalg$ is a semi-graphoid, then $\perp\!\!\!\perp$ is a basic separoid; and if $\amalg$ is graphoid then $\perp\!\!\!\perp$ is a basic strong separoid. Indeed, $\perp\!\!\!\perp$ is just the separoid (or strong separoid) closure of the relation $\amalg$.

Conversely, given an arbitrary separoid relation $\perp\!\!\!\perp$ on $(\mathcal{S}, \leqslant)$, its restriction $\amalg$ to pairwise disjoint terms will be a semi-graphoid (a graphoid if $\perp\!\!\!\perp$ is a strong separoid). By first so restricting $\perp\!\!\!\perp$ to pairwise disjoint terms, and then constructing a new relation using (12), we obtain a basic separoid $\perp\!\!\!\perp_*$, a subrelation of $\perp\!\!\!\perp$. We shall have $\perp\!\!\!\perp_* = \perp\!\!\!\perp$ if and only if $\perp\!\!\!\perp$ was already basic.

If f is a join semilattice homomorphism from $(\mathcal{S}, \leqslant)$ to $(\mathcal{S}', \leqslant', \perp\!\!\!\perp')$, and $\perp\!\!\!\perp$ is the induced separoid relation on $(\mathcal{S}, \leqslant)$, its restriction $\amalg$ will be a semi-graphoid on $(\mathcal{S}, \leqslant)$. Given such a homomorphism, we can apply the span deduction method of section 1.2 to deduce $f(x) \perp\!\!\!\perp' f(y) \mid f(z)$ for all $(x, y, z) \in \perp\!\!\!\perp$ whenever we can demonstrate this property for all $(x, y, z) \in \amalg$.

3.2.1. The finitary case

In this case, any join semilattice homomorphism f is fully specified (up to equivalence) by its values $f(i) := f(\{i\})$ on singleton sets (which values can be chosen arbitrarily), since for any finite $A \subseteq U$

$$f(A) \doteq \bigvee \{f(i): i \in A\}. \tag{13}$$

We call $f : (2^{*U}, \subseteq) \to (\mathcal{S}', \leqslant')$, constructed using (13) from a given *point map* $f : U \to \mathcal{S}$, the *join extension* of that function.

As we vary the range separoid $(\mathcal{S}', \leqslant', \perp\!\!\!\perp')$ and the point map f, we obtain a variety of induced separoid, and thus semi-graphoid, relations on $(2^{*U}, \subseteq)$ (note that in general, and contrary to an assumption implicit in most investigations to date, there is no requirement that $f(i)$ and $f(j)$ should differ for $i \neq j$). In particular, we can perform this construction using the various special separoids to be introduced in part III below. These yield: *probabilistic semi-graphoids* when f associates a random variable $X_i := f(i)$ with each $i \in U$, and $\perp\!\!\!\perp'$ is $\perp\!\!\!\perp_p[P]$ for some P, as described in section 6; *variation semi-graphoids* using $\perp\!\!\!\perp_v[\Omega]$ as in section 7.1; *etc*. From section 5 we also obtain, directly, definitions of *undirected graph* and *directed graph* semi-graphoids *etc*. A good deal of attention has been paid to problems of characterising and relating such special semi-graphoids, albeit with generally negative findings [61,62].

There are various other axiomatic systems and constructions which yield special finite semi-graphoids. We mention in particular the *valuation-based systems* of Shenoy [52]; the *conditional products* of Dawid and Studený [27]; and the *structural semi-graphoids* of Studený [63] or *semi-matroids* of Matúš [42], these last being alternative representations of essentially the same structures.

3.2.2. Semi-graphoid representation

Here we show that any separoid $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ can be represented as a homomorphic image of a separoid corresponding to some finitary semi-graphoid.

Take $U = \mathcal{S} \times \{1, 2, 3\}$, and let the point map $f : U \to \mathcal{S}$ be projection onto $\mathcal{S}$. As in section 3.2.1, the join extension of f supplies a separoid homomorphism from $(2^{*U}, \subseteq, \perp\!\!\!\perp_{\mathcal{S}})$ (where $\perp\!\!\!\perp_{\mathcal{S}}$ is the induced separoid relation) onto $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$. If now $\amalg_{\mathcal{S}}$ denotes the semi-graphoid restriction of $\perp\!\!\!\perp_{\mathcal{S}}$, and $\perp\!\!\!\perp_{\mathcal{S}*}$ the corresponding basic separoid relation on $(2^{*U}, \subseteq)$ (so that $A \perp\!\!\!\perp_{\mathcal{S}*} B \mid C$ if $A \perp\!\!\!\perp_{\mathcal{S}} B \mid C$ and $A \cap B \subseteq C$), we have $x \perp\!\!\!\perp y \mid z$ if and only if $\{(x, 1)\} \perp\!\!\!\perp_{\mathcal{S}*} \{(y, 2)\} \mid \{(z, 3)\}$ (or, equivalently, $\{(x, 1)\} \amalg_{\mathcal{S}*} \{(y, 2)\} \mid \{(z, 3)\}$) – thus allowing us to "read off" the former property from the latter.

The above construction associates a (special) finitary semi-graphoid with any separoid; while (12) did the reverse. Nevertheless, it should be noted that we have not demonstrated any kind of structure-preserving morphism between these two families of objects, or subfamilies of them. This might be an interesting algebraic direction to pursue.

Part II. Mathematical models

In this part we display some purely mathematical examples and applications of separoids.

Section 4 considers some special separoid relations on a lattice, directly related to the lattice properties. In particular, in section 4.1 it is shown that the property of *distributivity* of a lattice can be re-expressed as the requirement that a certain ternary relation should define a separoid. A similar characterisation of *modularity* is demonstrated in section 4.2. In section 4.5 a *submodular function* defined on a lattice is shown to generate a separoid relation.

In section 5 we consider some graphical separation properties that have been introduced for different kinds of graph, and their relationships to general separoid structures.

4. Lattice properties

Let $(\mathcal{L}, \leqslant)$ be a lattice. It is *distributive* if, for all $x, y, z \in \mathcal{L}$, $x \wedge (y \vee z) \doteq (x \wedge y) \vee (x \wedge z)$. The dual property, $x \vee (y \wedge z) \doteq (x \vee y) \wedge (x \vee z)$, is then automatically satisfied. An important example is the *Boolean algebra* of all subsets of a given set, ordered by inclusion.

The lattice $(\mathcal{L}, \leqslant)$ is *modular* if $z \leqslant x \Rightarrow (x \wedge y) \vee z \doteq x \wedge (y \vee z)$. In fact we can replace $\doteq$ by $\geqslant$, since the opposite inequality always holds. Important examples are the lattice of normal subgroups of a group, and the lattice of subspaces of a vector space, in each case ordered by inclusion.

4.1. Distributivity

We introduce the following ternary relation $\perp\!\!\!\perp_d$ on a lattice $(\mathcal{L}, \leqslant)$:

Definition 4.1. $x \perp\!\!\!\perp_d y \mid z \Leftrightarrow x \wedge y \leqslant z$.

Theorem 4.1. $(\mathcal{L}, \leqslant, \perp\!\!\!\perp_d)$ is a separoid if and only if the lattice $(\mathcal{L}, \leqslant)$ is distributive. In that case $(\mathcal{L}, \leqslant, \perp\!\!\!\perp_d)$ is a strong separoid.

Proof. 1. Suppose that $(\mathcal{L}, \leqslant, \perp\!\!\!\perp_d)$ is a separoid. Take $x, y, z \in \mathcal{L}$, and define $w = (x \wedge y) \vee (x \wedge z)$. We then have $x \wedge y \leqslant w$ and $x \wedge z \leqslant w \leqslant y \vee w$. That is, $x \perp\!\!\!\perp_d y \mid w$ and $x \perp\!\!\!\perp_d z \mid (y \vee w)$. Using P5 we deduce $x \perp\!\!\!\perp_d (y \vee z) \mid w$, i.e., $x \wedge (y \vee z) \leqslant (x \wedge y) \vee (x \wedge z)$. Since the opposite inequality always holds, we have shown that the lattice $(\mathcal{L}, \leqslant)$ is distributive.

2. Conversely, suppose that $(\mathcal{L}, \leqslant)$ is distributive. The separoid properties P1 and P2 are trivially seen to hold for $\perp\!\!\!\perp_d$. For P3 and P4, suppose $x \perp\!\!\!\perp_d y \mid z$, that is $x \wedge y \leqslant z$, and let $w \leqslant y$. Then $x \wedge w \leqslant x \wedge y \leqslant z$, i.e., P3 holds; and $x \wedge y \leqslant w \vee z$, i.e., P4 holds. For P5, if $x \perp\!\!\!\perp_d y \mid z$ and $x \perp\!\!\!\perp_d w \mid (y \vee z)$, we have $x \wedge y \leqslant z$ and $x \wedge w \leqslant y \vee z$. Thus $x \wedge (y \vee w) \doteq (x \wedge y) \vee (x \wedge w) \leqslant z \wedge (y \vee z) \doteq z$, i.e., $x \perp\!\!\!\perp_d (y \vee w) \mid z$. Finally, suppose $x \perp\!\!\!\perp_d y \mid z$ and $x \perp\!\!\!\perp_d z \mid y$, i.e., $x \wedge y \leqslant z$ and $x \wedge z \leqslant y$. We then have $x \wedge y \doteq y \wedge (x \wedge y) \leqslant y \wedge z$; and similarly $x \wedge z \leqslant y \wedge z$. Thus $x \wedge (y \vee z) \doteq (x \wedge y) \vee (x \wedge z) \leqslant y \wedge z$. That is, $z \perp\!\!\!\perp_d (y \vee z) \mid (y \wedge z)$, and P6 is proved. □

It is easily seen that, when $(\mathcal{L}, \leqslant)$ is distributive, the property $x \perp\!\!\!\perp_d y \mid z$ will hold if and only if $x \perp\!\!\!\perp_M y \mid z$, holds, as defined by:

$$x \perp\!\!\!\perp_M y \mid z \Leftrightarrow (x \vee z) \wedge (y \vee z) \doteq z. \tag{14}$$

More generally, from (4) we shall have $x \perp\!\!\!\perp_M y \mid z$ whenever $x \perp\!\!\!\perp y \mid z$ for some basic separoid relation $\perp\!\!\!\perp$ on $(\mathcal{L}, \leqslant)$.

If $\perp\!\!\!\perp$ is any separoid on the distributive lattice $(\mathcal{L}, \leqslant)$, then the intersection $\perp\!\!\!\perp \cap \perp\!\!\!\perp_d$ is a basic separoid (and a strong separoid if $\perp\!\!\!\perp$ is). When $\mathcal{S}$, as well as being distributive, has a minimal element and relative complements, $\perp\!\!\!\perp \cap \perp\!\!\!\perp_d$ coincides with the basic separoid relation $\perp\!\!\!\perp_*$ constructed in section 3.2.

4.1.1. *Possibility functions*

Let $(\mathcal{L}, \leqslant)$ be a distributive lattice, and $(\mathcal{S}, \leqslant')$ an arbitrary join semilattice. Let $\phi : \mathcal{S} \to \mathcal{L}$ be a join semilattice homomorphism. We then obtain a separoid relation $\perp\!\!\!\perp_d[\phi]$ on $(\mathcal{S}, \leqslant')$, induced by ϕ and $\perp\!\!\!\perp_d$, and given by:

$$x \perp\!\!\!\perp_d y \mid z\,[\phi] \Leftrightarrow \phi(x) \wedge \phi(y) \leqslant \phi(z). \tag{15}$$

In particular, consider the case $(\mathcal{L}, \leqslant) = (\Re, \leqslant)$, the real line under its usual order. Then ϕ will be a join semilattice homomorphism if

$$\phi(x \vee y) = \max\bigl\{\phi(x), \phi(y)\bigr\}. \tag{16}$$

In this case we call ϕ a *possibility function*. The relation (15) now becomes:

$$x \perp\!\!\!\perp_d y \mid z\,[\phi] \Leftrightarrow \min\bigl\{\phi(x), \phi(y)\bigr\} \leqslant \phi(z). \tag{17}$$

We deduce that this will define a separoid whenever ϕ is a possibility function.

4.2. *Modularity*

Consider the following ternary relation $\perp\!\!\!\perp_m$ on a lattice $(\mathcal{L}, \leqslant)$:

Definition 4.2. $x \perp\!\!\!\perp_m y \mid z \Leftrightarrow x \wedge (y \vee z) \leqslant z.$

Theorem 4.2. $(\mathcal{L}, \leqslant, \perp\!\!\!\perp_m)$ is a separoid if and only if the lattice $(\mathcal{L}, \leqslant)$ is modular. In that case it is a strong separoid, and

$$x \perp\!\!\!\perp_m y \mid z \Leftrightarrow x \perp\!\!\!\perp_M y \mid z. \tag{18}$$

Proof. 1. Suppose that $(\mathcal{L}, \leqslant, \perp\!\!\!\perp_m)$ is a separoid. Take $z \leqslant x$. Then $y \wedge \{x \vee (x \wedge y) \vee z\} \doteq y \wedge x \leqslant (y \wedge x) \vee z$, i.e., $y \perp\!\!\!\perp_m x \mid (x \wedge y) \vee z$, and thus, by P2, $x \perp\!\!\!\perp_m y \mid (x \wedge y) \vee z$. Since $z \leqslant (x \wedge y) \vee z$, we then easily deduce from the separoid properties that $x \perp\!\!\!\perp_m (y \vee z) \mid (x \wedge y) \vee z$, i.e., $x \wedge (y \vee z) \leqslant (x \wedge y) \vee z$, thus demonstrating modularity.

2. Conversely, suppose $(\mathcal{L}, \leqslant)$ is modular. Property P1 requires $x \wedge (x \vee y) \leqslant x$, which holds with equality. Next note that $x \perp\!\!\!\perp_m y \mid z$ if and only if $x \wedge (y \vee z) \leqslant z$, i.e., $\{(y \vee z) \wedge x\} \vee z \doteq z$, and under modularity (using $z \leqslant y \vee z$) this is equivalent to $(y \vee z) \wedge (x \vee z) \doteq z$. This symmetrical restatement of $\perp\!\!\!\perp_m$ demonstrates P2 and (18). For

P3 and P4, suppose $x \perp\!\!\!\perp_m y \mid z$, i.e., $x \wedge (y \vee z) \leqslant z$, and let $w \leqslant y$. Then $x \wedge (w \vee z) \leqslant x \wedge (y \vee z) \leqslant z$, so that $x \perp\!\!\!\perp_m w \mid z$, i.e., P3 holds; while $x \wedge (y \vee z \vee w) \doteq x \wedge (y \vee z) \leqslant z \leqslant z \vee w$, so that $x \perp\!\!\!\perp_m y \mid (z \vee w)$, and P4 holds. For P5, supposing $x \perp\!\!\!\perp_m y \mid z$ and $x \perp\!\!\!\perp_m w \mid (y \vee z)$, we have $x \wedge (y \vee z) \leqslant z$ and $x \wedge (w \vee y \vee z) \leqslant y \vee z$. Since $x \wedge (w \vee y \vee z) \leqslant x$, we deduce $x \wedge (w \vee y \vee z) \leqslant x \wedge (y \vee z) \leqslant z$, i.e., $x \perp\!\!\!\perp_m (w \vee y) \mid z$, and P5 follows. Finally, for P6, suppose $x \perp\!\!\!\perp_m y \mid z$ and $x \perp\!\!\!\perp_m z \mid y$, so that $x \wedge (y \vee z) \leqslant z$ and also $x \wedge (y \vee z) \leqslant y$. Then $x \wedge (y \vee z) \leqslant y \wedge z$, i.e., $x \perp\!\!\!\perp_m (y \vee z) \mid (y \wedge z)$. □

Corollary 4.3. When $(\mathcal{L}, \leqslant)$ is modular, the strong separoid $(\mathcal{L}, \leqslant, \perp\!\!\!\perp_m)$ is basic.

Theorem 4.4. $(\mathcal{L}, \leqslant, \perp\!\!\!\perp_M)$ is a separoid if and only if the lattice $(\mathcal{L}, \leqslant)$ is modular.

Proof. By theorem 4.2, if $(\mathcal{L}, \leqslant)$ is modular, then $\perp\!\!\!\perp_M$ is identical to $\perp\!\!\!\perp_m$, and thus defines a separoid.

Conversely, suppose $\perp\!\!\!\perp_M$ defines a separoid, and let $z \leqslant x$. From the definition, for any y we have $x \perp\!\!\!\perp_M y \mid (x \wedge y)$, so that, by P4, $x \perp\!\!\!\perp_M y \mid (x \wedge y) \vee z$. Again from the definition, this is equivalent to $x \wedge (y \vee z) \doteq (x \wedge y) \vee z$, showing modularity. □

Combining theorems 4.2 and 4.4, we see that $(\mathcal{L}, \leqslant)$ is modular if and only if either $\perp\!\!\!\perp_m$ or $\perp\!\!\!\perp_M$ defines a separoid, in which case they are equivalent, and define a strong separoid. When $(\mathcal{L}, \leqslant)$ is distributive it is also modular, and all three relations $\perp\!\!\!\perp_d$, $\perp\!\!\!\perp_m$ and $\perp\!\!\!\perp_M$ coincide and define a strong separoid.

Note that the orthogonoid relation (8) corresponding to $\perp\!\!\!\perp_m$ on a modular lattice is trivial: $x \perp y$ for all x, y.

If we start with a non-basic separoid $\perp\!\!\!\perp$ on a modular lattice, we can construct a basic separoid $\perp\!\!\!\perp_*$ by first constructing $\perp$ using (8), and then $\perp\!\!\!\perp_*$ using (11). It is easy to see that $\perp\!\!\!\perp_* = \perp\!\!\!\perp \cap \perp\!\!\!\perp_m$. This generalises the construction in section 4.1.

4.2.1. Vector space

Let $\mathcal{L}$ be the lattice of all subspaces of a vector space $\mathcal{V}$, ordered by inclusion. This lattice is modular, and therefore, by theorem 4.2, the relation $\perp\!\!\!\perp_M$ given by (14) is a strong separoid. This relation is sometimes called *linear independence*.

4.3. Hilbert space

Now let $\mathcal{V}$, equipped with inner product I, be a Hilbert space, and $\mathcal{L}$ its lattice of closed subspaces, ordered by inclusion.

Definition 4.3. We say X *is orthogonal independent of* Y *given* Z (*with respect to* I), and write $X \perp\!\!\!\perp_O Y \mid Z\,[I]$, if, for each $x \in X$, the projection $\Pi_{Y \vee Z} x$ of x onto $Y \vee Z$ lies in Z.

More generally, I might only be a semi-inner product, so that we can have $I(x, x) = 0$ for $x \neq 0$. Then projection is not uniquely defined. In this case, we define $X \perp\!\!\!\perp_O Y \mid Z\,[J]$ if there exists some projection of x onto $Y \vee Z$ that lies in Z.

An equivalent symmetrical requirement is: for all $x \in X$, $y \in Y$, the 'residual vectors' $x - \Pi_Z x$ and $y - \Pi_Z y$ are I-orthogonal. (If I is only a semi-inner product, $\Pi_Z x$ can here represent any choice for the projection of x onto Z.)

It is easy to confirm that, under this definition, $\perp\!\!\!\perp_O$ defines a separoid (a basic strong separoid if I is a full inner product). Note that the lattice $\mathcal{L}$ is modular, and the corresponding orthogonoid relation, given by (8), is just geometric perpendicularity (relative orthogonality) of subspaces.

4.4. Join orthogonality

Let $(\mathcal{L}, \leqslant)$, $(\mathcal{L}', \leqslant')$ be modular lattices, and $f : \mathcal{L} \to \mathcal{L}'$ a *meet semilattice homomorphism*, i.e., $f(x \wedge y) \doteq' f(x) \wedge' f(y)$. It readily follows that $x \leqslant y \Rightarrow f(x) \leqslant' f(y)$, and $f(x \vee y) \geqslant' f(x) \vee' f(y)$.

Definition 4.4. For $x, y \in \mathcal{L}$, we say that x *is join orthogonal to* y *under* f, and write $x \perp_J y\,[f]$, if

$$f(x \vee y) \doteq' f(x) \vee' f(y). \tag{19}$$

Evidently, it is equivalent to require $f(x \vee y) \leqslant' f(x) \vee' f(y)$.

Theorem 4.5. The relation $\perp_J\ [f]$ defines an orthogonoid on $\mathcal{L}$.

Proof. G1 and G2 are immediate.

For G3 and G4, suppose $x \wedge y \leqslant z \leqslant y$, or equivalently $z \doteq (x \vee z) \wedge y$; and $x \perp_J y\,[f]$, so that (19) holds. We then have $f(x) \vee' f(z) \doteq' f(x) \vee' \{f(x \vee z) \wedge' f(y)\}$, and by modularity in $\mathcal{L}'$ this is $f(x \vee z) \wedge' \{f(x) \vee' f(y)\} \doteq' f(x \vee z) \wedge' f(x \vee y) \doteq' f\{(x \vee z) \wedge (x \vee y)\} \doteq' f(x \vee z)$. So we have shown G3.

We then deduce $f(x \vee z) \vee' f(y) \doteq' f(x) \vee' f(y) \vee' f(z) \doteq' f(x \vee y) \vee' f(z) \doteq f(x \vee y)$ (since $z \leqslant x \vee y$). That is, $(x \vee z) \perp_J y\,[f]$.

Finally, if $x \perp_J z\,[f]$ and $(x \vee z) \perp_J y\,[f]$, we have $f(x \vee z) \doteq' f(x) \vee' f(z)$ and $f(x \vee y) \doteq' f(x \vee z) \vee' f(y)$. We deduce $f(x \vee y) \doteq' f(x) \vee' f(z) \vee' f(y) \doteq' f(x) \vee' f(y)$, i.e., $x \perp_J y\,[f]$. Hence G5 is shown. □

Corollary 4.6. If we define $\perp\!\!\!\perp_J\,[f]$ by:

$$x \perp\!\!\!\perp_J y \mid z\,[f] \Leftrightarrow \begin{cases} f(x \vee y \vee z) \doteq' f(x \vee z) \vee' f(y \vee z) \text{ and} \\ (x \vee z) \wedge (y \vee z) \doteq z \end{cases}$$

then $\perp\!\!\!\perp_J\,[f]$ defines a basic separoid on $(\mathcal{L}, \leqslant)$.

4.5. Submodularity

Let $(\mathcal{L}, \leqslant)$ be a lattice. A function $\phi : \mathcal{L} \to \Re$ is called *submodular* if, for all $x, y \in \mathcal{L}$,

$$\phi(x \vee y) + \phi(x \wedge y) \leqslant \phi(x) + \phi(y). \tag{20}$$

If $(\mathcal{L}, \leqslant)$ is a *semi-modular lattice*, its height function is a strictly increasing integer-valued submodular function [34, p. 173].

Now let $\phi : \mathcal{L} \to \Re$ be submodular and increasing (i.e., $x \leqslant y \Rightarrow \phi(x) \leqslant \phi(y)$).

Lemma 4.7. In this case,

$$\phi(x \vee z) + \phi(y \vee z) \geqslant \phi(x \vee y \vee z) + \phi(z). \tag{21}$$

Proof. By submodularity, $\phi(x \vee z) + \phi(y \vee z) \geqslant \phi(x \vee y \vee z) + \phi\{(x \vee z) \wedge (y \vee z)\}$. Also $(x \vee z) \wedge (y \vee z) \geqslant z$, so, since ϕ is increasing, $\phi\{(x \vee z) \wedge (y \vee z)\} \geqslant \phi(z)$. □

We introduce the following relation on $\mathcal{L}$:

Definition 4.5. We write $x \perp\!\!\!\perp_s y \mid z\, [\phi]$, and say that x is (*submodular*) *independent of y given z under ϕ*, if

$$\phi(x \vee z) + \phi(y \vee z) = \phi(x \vee y \vee z) + \phi(z). \tag{22}$$

By lemma 4.7, it is enough to require the same property with $\leqslant$ replacing $=$ in (22). Note that if ϕ is a possibility function it is submodular and increasing, and the relation $\perp\!\!\!\perp_s [\phi]$ is then identical with $\perp\!\!\!\perp_d [\phi]$ given by (17).

The above definition generalises that of a *semimatroid* relation on the finite lattice $(2^U, \subseteq)$ – see [42], where this is shown to define a semi-graphoid.

Theorem 4.8. $\perp\!\!\!\perp_s [\phi]$ is a separoid on $(\mathcal{L}, \leqslant)$.

Proof. P1 and P2 are trivial. For P3 and P4, suppose that $x \perp\!\!\!\perp_s y \mid z\, [\phi]$, so that (22) holds, and let $w \leqslant y$. By submodularity, we have $\phi(x \vee w \vee z) + \phi(y \vee z) \geqslant \phi(x \vee y \vee z) + \phi\{(x \vee w \vee z) \wedge (y \vee z)\}$, and since ϕ is increasing we deduce

$$\phi(x \vee w \vee z) + \phi(y \vee z) \geqslant \phi(x \vee y \vee z) + \phi(w \vee z). \tag{23}$$

Subtracting (22) from (23) and rearranging then yields $\phi(x \vee w \vee z) + \phi(z) \geqslant \phi(x \vee z) + \phi(w \vee z)$, and P3 is verified.

Again, by the submodularity and increasing properties of ϕ,

$$\begin{aligned}\phi(x \vee z) + \phi(w \vee z) &\geqslant \phi(x \vee w \vee z) + \phi\big\{(x \vee z) \wedge (w \vee z)\big\} \\ &\geqslant \phi(x \vee w \vee z) + \phi(z),\end{aligned}$$

and now subtracting (22) and rearranging yields $\phi(x \vee y \vee z) + \phi(w \vee z) \geqslant \phi(x \vee w \vee z) + \phi(y \vee z)$, which demonstrates P4.

Finally, suppose both $x \perp\!\!\!\perp_s y \mid z\,[\phi]$, so that (22) holds, and also $x \perp\!\!\!\perp_s w \mid (y \vee z)\,[\phi]$, so that

$$\phi(x \vee y \vee w \vee z) + \phi(y \vee z) = \phi(x \vee y \vee z) + \phi(y \vee w \vee z). \tag{24}$$

Subtracting (22) from (24) now yields $\phi(x \vee z) + \phi(y \vee w \vee z) = \phi(x \vee y \vee z \vee w) + \phi(z)$, and P5 follows. □

The natural quasiorder $\preceq_\phi$, induced from $\perp\!\!\!\perp_s [\phi]$ as in (5), is given by: $x \preceq_\phi y$ if $\phi(x \vee y) = \phi(y)$. In particular, we deduce that this relation must be transitive, which is not immediately obvious. The separoid relation $\perp\!\!\!\perp_\phi$ is thus basic if and only if ϕ is strictly increasing. In that case the corresponding orthogonoid relation $\perp\!\!\!\perp_s [\phi]$ is given by:

$$x \perp\!\!\!\perp_s y\,[\phi] \text{ if } \phi(x \vee y) + \phi(x \wedge y) = \phi(x) + \phi(y). \tag{25}$$

This is analogous to the definition of *modular pair* in matroid theory [43, section 6.9]. By submodularity, it is sufficient to require $\geqslant$ in place of $=$ in (25).

It is instructive to note that, even when ϕ is strictly increasing, so that $\perp\!\!\!\perp_s [\phi]$ is a basic separoid relation, it need not yield a strong separoid.

Example 4.1. Let $\mathcal{L}$ be the lattice of subsets of $\{1, 2, 3\}$, ordered by inclusion, and for $x \in \mathcal{L}$ define $\phi(x)$ to be: 0 if the size $|x|$ of x is 0; 2 if $|x| = 1$; 3 if $|x| = 2$; and 4 if $|x| = 3$. This is a strictly increasing submodular function, and the corresponding separoid relation $\perp\!\!\!\perp_s [\phi]$ is that generated by the properties: $\{1\} \perp\!\!\!\perp_s \{2\} \mid \{3\}\,[\phi]$, $\{2\} \perp\!\!\!\perp_s \{3\} \mid \{1\}\,[\phi]$, and $\{3\} \perp\!\!\!\perp_s \{1\} \mid \{2\}\,[\phi]$. This yields a basic, but not a strong separoid. □

Even when ϕ is not required to be increasing, we can still obtain a separoid so long as the lattice $(\mathcal{L}, \leqslant)$ is modular. Again consider the relation $\perp_s\ [\phi]$ given by (25). We can show directly that this defines an orthogonoid.

Theorem 4.9. Let ϕ be a submodular function on a modular lattice $(\mathcal{L}, \leqslant)$. Then $\perp\ [\phi]$ is an orthogonoid relation on $(\mathcal{L}, \leqslant)$.

Proof. G1 and G2 are trivial. For G3 and G4, suppose $x \perp_s y\,[\phi]$ and $x \wedge y \leqslant z \leqslant y$, whence $x \wedge y \doteq x \wedge z$. We have $(x \vee y) \vee z \doteq x \vee y$. By modularity,

$$(x \vee z) \wedge y \doteq (x \wedge y) \vee z \doteq z. \tag{26}$$

Then (20) yields

$$\phi(x \vee z) + \phi(y) \geqslant \phi(x \vee y) + \phi(z). \tag{27}$$

On combining this with (25), and using $x \wedge y \doteq x \wedge z$, we obtain

$$\phi(x \vee z) + \phi(x \wedge z) \geqslant \phi(x) + \phi(z), \tag{28}$$

from which we deduce $x \perp_s z\,[\phi]$. Hence G3 holds.

We must now have equality in (28), i.e.,

$$\phi(x \vee z) + \phi(x \wedge z) = \phi(x) + \phi(z). \tag{29}$$

Working the preceding steps backwards starting from (29), we see that we also have equality in (27). But this is just the condition that $(x \vee z) \perp_s y\,[\phi]$. Hence G4.

Finally, suppose that $(x \vee z) \wedge y \doteq z \leqslant y$, that $x \perp_s z\,[\phi]$ and that $(x \vee z) \perp_s y\,[\phi]$. We obtain

$$\begin{aligned} \phi(x \vee z) + \phi(x \wedge z) &= \phi(x) + \phi(z), \\ \phi(x \vee y) + \phi(z) &= \phi(x \vee z) + \phi(y). \end{aligned}$$

We readily deduce

$$\phi(x \vee y) + \phi(x \vee z) = \phi(x) + \phi(y),$$

and thus, using $x \wedge z \doteq x \wedge y$, $x \perp_s y\,[\phi]$. Hence G5 holds. □

Corollary 4.10. Given a submodular function $\phi : \mathcal{L} \to \Re$ on a modular lattice $(\mathcal{L}, \leqslant)$, define a relation $\perp\!\!\!\perp_{s'}[\phi]$ on $\mathcal{L}$ by: $x \perp\!\!\!\perp_{s'} y \mid z\,[\phi]$ if (22) holds and $(x \vee z) \wedge (y \vee z) = z$. Then $\perp\!\!\!\perp_{s'}[\phi]$ defines a basic separoid on $(\mathcal{L}, \leqslant)$.

If, in addition, ϕ is strictly increasing, then $\perp\!\!\!\perp_{s'}[\phi]$ will be identical with $\perp\!\!\!\perp_s[\phi]$. Alternatively, suppose $\mathcal{L}$ is finite. Then its height function h is modular (i.e., $x \perp_s y\,[h]$ for all $x, y \in \mathcal{L}$) and strictly increasing; and so by choosing K large enough we can ensure that the function $\phi_K := \phi + Kh$ is submodular and strictly increasing. Moreover, the relations $\perp_s[\phi]$ and $\perp_s[\phi_K]$ are identical. It follows that $\perp\!\!\!\perp_{s'}[\phi]$ is identical to $\perp\!\!\!\perp_s[\phi_K]$.

We note that obvious dual analogues of all the results in this section will apply to *supermodular* functions, which are just the negatives of submodular functions. In particular, if ϕ is both supermodular and decreasing, $\perp\!\!\!\perp_s[\phi]$, as given in definition 4.5, will be a separoid.

All the above theory could be generalized by allowing the function ϕ to take values in a partially ordered linear space, rather than $\Re$.

5. Graphs

The connection between graphical definitions of separation [45,46] and notions (especially probabilistic) of conditional independence has fuelled much of the recent research on separoids – as well as being responsible for such terminology as '(semi)-graphoid' and 'separoid'. In this section we use the graph-theory terminology of Cowell et al. [13, chapter 3]. See chapter 4 of that book for further details of the concepts introduced here. We deal only with finite graphs, although much of the analysis applies equally to infinite graphs.

5.1. *Undirected graphs*

Let $\mathcal{G} = (U, E)$ be an undirected graph. We can define a relation $\perp\!\!\!\perp_{ug}[\mathcal{G}]$ on 2^U by: $A \perp\!\!\!\perp_{ug} B \mid C\,[\mathcal{G}]$ if every path in $\mathcal{G}$ joining a point in A to a point in B intersects C. This is readily seen to be a strong separoid on $(2^U, \subseteq)$. (It could therefore be, and typically is, characterised by its semi-graphoid restriction, where we only require this separation condition for A, B and C pairwise disjoint)

Let $\pi_{ug}[\mathcal{G}]$ be the subrelation of $\perp\!\!\!\perp_{ug}[\mathcal{G}]$ where we restrict to triples of the form $A = \{i\},\ B = \{j\},\ C = U \setminus \{i, j\}\,(i, j \in U;\ i \neq j)$. It is easily seen that this relation holds if and only if there is no edge between i and j in $\mathcal{G}$. It can then be shown, by an argument essentially identical to that of Pearl and Paz [48] (see also [47, appendix 3-A]) that $\pi_{ug}[\mathcal{G}]$ is a strong separoid spanning relation for $\perp\!\!\!\perp_{ug}[\mathcal{G}]$.

Another such strong separoid spanning relation is given by $\lambda_{ug}[\mathcal{G}]$, in which we restrict $\perp\!\!\!\perp$ to triples of the form $A = \{i\},\ B = U \setminus \{i\},\ C = \mathrm{bd}(i)$, for $i \in U$, where $\mathrm{bd}(i)$ denotes the 'boundary', or set of 'neighbours', of i in $\mathcal{G}$.

Suppose now that $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ is a strong separoid, and that we have a join semi-lattice homomorphism $f : U \to \mathcal{S}$ (perhaps constructed from a point map, as in (13)).

Definition 5.1. We call $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ (*globally*) *Markov* (*with respect to* $\mathcal{G}$, *under* f) if:

(i) f is a lattice homomorphism, i.e.,

$$f(A) \wedge f(B) \doteq f(A \cap B),$$

(ii) f is a separoid (and thus a strong separoid) homomorphism, i.e.,

$$A \perp\!\!\!\perp_{ug} B \mid C\,[\mathcal{G}] \Rightarrow f(A) \perp\!\!\!\perp f(B) \mid f(C).$$

We call $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ *pairwise Markov* if (ii) is only required for triples $(A, B, C) \in \pi_{ug}[\mathcal{G}]$, and *local Markov* if (ii) is only required for triples $(A, B, C) \in \lambda_{ug}[\mathcal{G}]$.

Using the method of span deduction, it now follows that the global, pairwise and local Markov properties are all equivalent. In particular, we can deduce the global Markov property whenever we have $f(i) \perp\!\!\!\perp f(j) \mid f(U \setminus \{i, j\})$ for all distinct $i, j \in U$ that are not joined by an edge of $\mathcal{G}$.

It is important to note, for the above, the requirements that $\perp\!\!\!\perp$ be a strong separoid on a lattice $(\mathcal{S}, \leqslant)$, and that f be a lattice homomorphism. These are needed because $\pi_{ug}[\mathcal{G}]$ is not in general a (weak) separoid spanning relation for $\perp\!\!\!\perp_{ug}[\mathcal{G}]$.

5.2. *Directed graphs*

Now let $\mathcal{D} = (U, E)$ be a directed acyclic graph. Its *moral graph* $\mathcal{D}^m$ is the undirected graph obtained by first adding an undirected edge between i and j in U whenever they have a common child (and are not already joined); and then ignoring the directions of edges. For $A \subseteq U$, the *induced graph* $\mathcal{D}_A$ is the directed subgraph obtained from $\mathcal{D}$ by simply ignoring all vertices outside A and all edges involving them. For any subgraph $\mathcal{D}'$ of $\mathcal{D}$, its *ancestral graph*, $\mathrm{An}(\mathcal{D}')$, is the subgraph induced by the set of nodes in $\mathcal{D}'$ together with all their ancestors in $\mathcal{D}$.

We introduce the following relation on 2^U:

Definition 5.2.

$$A \perp\!\!\!\perp_{dg} B \mid C\ [\mathcal{D}] \Leftrightarrow A \perp\!\!\!\perp_{ug} B \mid C\ \left[\left\{\mathrm{An}(\mathcal{D}_{A\cup B\cup C})\right\}^m\right].$$

This property has an alternative expression by means of the 'd-separation' criterion of Pearl [47] (section 3.3.1). It can be shown that $\perp\!\!\!\perp_{dg}[\mathcal{D}]$ is a strong separoid relation on $(2^U, \subseteq)$.

Let now $\lambda_{dg}[\mathcal{D}]$ denote the set of triples (A, B, C) of the form: $A = \{i\}$; $B = \mathrm{nd}(i)$; $C = \mathrm{pa}(i)$ $(i \in U)$, where $\mathrm{pa}(i)$ denotes the 'parents' of i, and $\mathrm{nd}(i)$ the 'non-descendants' of i. This is easily seen to be a subrelation of $\perp\!\!\!\perp_{dg}[\mathcal{D}]$, and is in fact a separoid spanning relation for it. This may be shown by an argument essentially the same as that in section 6 of Lauritzen et al. [40].

We can now proceed in parallel to section 5.1 – although things are here simpler, since we can restrict attention to separoids, rather than strong separoids. Thus suppose we have a join-semilattice homomorphism $f : (2^U, \subseteq) \to (\mathcal{S}, \leqslant)$. We call $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$ (*globally*) *directed Markov* (*with respect to* $\mathcal{D}$*, under* f) if f is a separoid homomorphism from $f : (2^U, \subseteq, \perp\!\!\!\perp_{dg}[\mathcal{D}])$ to $(\mathcal{S}, \leqslant, \perp\!\!\!\perp)$. We call it *locally directed Markov* if $f(i) \perp\!\!\!\perp f\{\mathrm{nd}(i)\} \mid \{\mathrm{pa}(i)\}$ for all $i \in U$. Then these two properties are equivalent. (There is a slightly stronger result available, using the concept of 'local well-numbering Markov property'; see Lauritzen et al. [40].)

5.3. *Other graphical relations*

There have been numerous other definitions of graph separation, in a variety of graphical frameworks, usually yielding separoid or strong separoid relations on $(2^U, \subseteq)$, sometimes with useful attendant spanning relations. We give no further details here, but mention in particular extensions to 'chain graphs' [3,30,66], to 'reciprocal graphs' [38], to 'annotated graphs' [44], and to 'MC graphs' [39]. See [65] for a brief overview of some of this work.

Part III. Uncertainty and irrelevance

In this part we survey a number of examples of separoids, intended to explicate a variety of concepts of 'irrelevance' arising in different calculi of uncertainty. We particularly draw attention to the fact that most previous definitions and investigations of these concepts have confined themselves to problems involving a finite number of variables – or, equivalently, to a domain space Ω that is a subset of a (generally Euclidean) product space, and to functions given by coordinate projections – thus allowing the structures introduced to be regarded as semi-graphoids. In contrast, with the exception of sections 6.2.1 and 6.2.2 below, we impose no such restriction here.

Section 6 considers *Probabilistic Independence*, which historically has been of crucial importance to the development of Probability Theory, and formed the principal motivation for more abstract development of separoid properties.

Section 7 treats another classic case, embodying logical, rather than probabilistic, relations. Three seemingly different separoid structures, *Variation Independence* of functions, *Equivalence Independence* of equivalence relations, and *Qualitative Independence* of τ-fields, are shown to be isomorphic.

In section 8 we consider various extensions of probabilistic independence when there is a family of possible distributions: *Statistical Independence, Meta Independence* and *Hyper Independence*. These incorporate structural, as well as probabilistic, aspects of the family of distributions, and have important applications to problems of statistical inference.

Section 9 studies another non-probabilistic form of independence, *Natural Independence,* while section 10 considers *Covariance Independence,* a weaker form of probabilistic independence based on second-moment properties only. In section 11 we introduce various notions of *Belief Independence* for Dempster-Shafer belief functions, and show that these are not equivalent.

Finally, section 12 points to some limitations of the separoid as an all-purpose representation of irrelevance, and some possible generalisations.

6. Probabilistic independence

Although the specific application to probabilistic conditional independence was the original motivation for considering separoids, it is somewhat more complex than many other models of the axioms. Here we relate it to the general abstract structure, and in particular identify a concrete minimal representation (see section 2.1). We also consider connections with *multi-information.*

6.1. σ-fields

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and $\mathcal{S}$ the lattice of sub-σ-fields of $\mathcal{F}$, ordered by inclusion. For $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathcal{S}$, we write $\mathcal{A} \perp\!\!\!\perp_p \mathcal{B} \mid \mathcal{C}\,[P]$ to denote that $\mathcal{A}$ and $\mathcal{B}$ are (*conditionally*) *independent, given* $\mathcal{C}$ (*under* P). When $\mathcal{C}$ is the trivial σ-field $\{\emptyset, \Omega\}$, we write just $\mathcal{A} \perp\!\!\!\perp_p \mathcal{B}\,[P]$, and say that $\mathcal{A}$ is (*marginally*) *independent of* $\mathcal{B}$ (*under* P).

The property of probabilistic conditional independence can be expressed mathematically in numerous logically equivalent ways, including the following [16]:

(i) For each $A \in \mathcal{A}$,

$$P(A \mid \mathcal{B} \vee \mathcal{C}) = P(A \mid \mathcal{C}), \text{ a.s. } [P];$$

(ii) For each $A \in \mathcal{A}$, there exists some $\mathcal{C}$-measurable function a such that

$$P(A \mid \mathcal{B} \vee \mathcal{C}) = a, \text{ a.s. } [P];$$

(iii) For each $A \in \mathcal{A}$, $B \in \mathcal{B}$,

$$P(A \cap B \mid \mathcal{C}) = P(A \mid \mathcal{C}) \times P(B \mid \mathcal{C}), \text{ a.s. } [P].$$

In the above, "a.s. $[P]$" denotes "almost surely under P", this qualification being needed because a conditional probability such as $P(A \mid \mathcal{C})$ is a non-uniquely defined $\mathcal{C}$-measurable function, any two possible versions of which may differ on a P-null set in $\mathcal{C}$. It can be shown [16,19] that $(\mathcal{S}, \subseteq, \perp\!\!\!\perp_p [P])$ is a separoid, but not, in general, a strong separoid.

Let $\mathcal{I}(P)$ denote the σ-ideal of all P-null sets in $\mathcal{F}$. For any $\mathcal{A} \in \mathcal{S}$, let $\widetilde{\mathcal{A}}$ be the smallest σ-field containing $\mathcal{A}$ and $\mathcal{I}(P)$. If $\mathcal{I}(P) \subseteq \mathcal{A}$, so that $\widetilde{\mathcal{A}} = \mathcal{A}$, we call $\mathcal{A}$ a *completed σ-field,* and we denote the family of all completed σ-fields by $\widetilde{\mathcal{S}}$. It is not difficult to see that the natural quasiorder $\preceq$ is given by: $\mathcal{A} \preceq \mathcal{B}$ if $\widetilde{\mathcal{A}} \subseteq \widetilde{\mathcal{B}}$. It readily follows that $(\widetilde{\mathcal{S}}, \subseteq, \perp\!\!\!\perp_p [P])$ is a concrete minimal representation of $(\mathcal{S}, \subseteq, \perp\!\!\!\perp_p [P])$, which is in fact a sub-separoid of $(\mathcal{S}, \subseteq, \perp\!\!\!\perp_p [P])$, since $\mathcal{A}, \mathcal{B} \in \widetilde{\mathcal{S}} \Rightarrow \mathcal{A} \vee \mathcal{B} \in \widetilde{\mathcal{S}}$. Moreover, $(\widetilde{\mathcal{S}}, \subseteq)$ is a sublattice of $(\mathcal{S}, \subseteq)$, and $(\widetilde{\mathcal{S}}, \subseteq, \perp\!\!\!\perp_p [P])$ is a strong separoid [19] – even though $(\mathcal{S}, \subseteq, \perp\!\!\!\perp_p [P])$ may not be.

In general, the completion map $\widetilde{\ } : \mathcal{S} \to \widetilde{\mathcal{S}}$ does not preserve intersection. However, if P is strictly positive, i.e., $\mathcal{I}(P) = \{\emptyset\}$, which can only happen if Ω is at most countable, $\widetilde{\ }$ is just the identity, so that in this case $(\mathcal{S}, \subseteq, \perp\!\!\!\perp_p)$ is a strong separoid. More generally, we may be concerned with a sub-separoid $\mathcal{S}'$ of $\mathcal{S}$. If we can show that, for all $\mathcal{A}, \mathcal{B} \in \mathcal{S}'$, $\widetilde{\mathcal{A}} \cap \widetilde{\mathcal{B}} = \widetilde{\mathcal{A} \cap \mathcal{B}}$, so that $\widetilde{\ } : \mathcal{S}' \to \widetilde{\mathcal{S}'}$ is a lattice homomorphism, then $(\mathcal{S}', \subseteq, \perp\!\!\!\perp_p [P])$ will be a strong separoid, induced by this map.

6.2. *Random variables*

By a *random variable*, say X, we understand simply a function from Ω into an arbitrary set, R_X say. The set $\mathbf{F}$ of all such functions, quasiordered by contraction ($X \leqslant Y$ if $Y(\omega) = Y(\eta)) \Rightarrow X(\omega) = X(\eta)$) forms a lattice, with join $X \vee Y \doteq (X, Y)$.

The associated *induced sub-σ-field* $\mathcal{F}_X$ of $\mathcal{F}$ is defined by: $\mathcal{F}_X = \mathcal{F} \cap \{X^{-1}(S)\colon S \subseteq R_X\}$. Then $X \leqslant Y$ if and only if $\mathcal{F}_X \subseteq \mathcal{F}_Y$, and $\mathcal{F}_{(X,Y)} = \mathcal{F}_X \vee \mathcal{F}_Y$.

The map $X \mapsto \mathcal{F}_X$ is a join semilattice homomorphism from $(\mathbf{F}, \leqslant)$ into $(\mathcal{S}, \subseteq)$. Consequently, given any separoid relation on $\mathcal{S}$, we can construct the induced separoid relation on $(\mathbf{F}, \leqslant)$. In particular, we obtain the following separoid relation on $(\mathbf{F}, \leqslant)$:

$$X \perp\!\!\!\perp_p Y \mid Z\, [P] \Leftrightarrow \mathcal{F}_X \perp\!\!\!\perp_p \mathcal{F}_Y \mid \mathcal{F}_Z\, [P].$$

Again, $(\mathbf{F}, \leqslant, \perp\!\!\!\perp_p [P])$ will not usually be a strong separoid.

The relation $X \preceq Y\, [P]$, equivalent to $X \perp\!\!\!\perp_p X \mid Y\, [P]$, now holds if and only if $\widehat{X} \subseteq \widehat{Y}$, where the map $\widehat{\ } : \mathbf{F} \to \widetilde{\mathcal{S}}$ given by: $\widehat{X} = \widetilde{\mathcal{F}_X}$. Thus $X \simeq Y\, [P] \Leftrightarrow \widehat{X} = \widehat{Y}$. We could equivalently construct the induced separoid using the map $\widehat{\ } : \mathbf{F} \to \widetilde{\mathcal{S}}$. Note that this depends (through the σ-ideal $\mathcal{I}(P)$) on the specific probability distribution P under consideration.

6.2.1. Markov properties

Suppose we have a point map f, associating a random variable $X_i \in \mathbf{F}$ with each i in a finite set U, and join-extended as in (13). Let $\mathcal{D}$ be a directed graph $\mathcal{D}$ on U. We can then apply the theory of section 5.2 to show that the global and local directed Markov properties with respect to $\mathcal{D}$ are equivalent. That is, if we know that $X_i \perp\!\!\!\perp_p X_{\mathrm{nd}(i)} \mid X_{\mathrm{pa}}(i)\,[P]$, we can deduce $X_A \perp\!\!\!\perp_p X_B \mid X_C\,[P]$ whenever $A \perp\!\!\!\perp_{dg} B \mid C\,[\mathcal{D}]$.

For suitable P it may further be possible to demonstrate that that the map f is a lattice homomorphism from $(2^U, \subseteq)$ into $(\mathbf{F}, \preceq)$: equivalently, that $X_A \wedge X_B \simeq X_{A\cap B}\,[P]$ for $A, B \subseteq U$. This holds, in particular, when $(X_1, \ldots, X_n)$ has an everywhere positive joint density under P. In that case the map $\widehat{f} : A \mapsto \widehat{X}_A$ is a lattice homomorphism from $(2^U, \subseteq)$ into $(\widetilde{\mathcal{S}}, \subseteq)$. Since $(2^U, \subseteq)$ is a strong separoid, and $\perp\!\!\!\perp_p\,[P]$ is a strong separoid on $(\widetilde{\mathcal{S}}, \subseteq)$, we can then use the theory of section 5.1 to show that the (undirected) global, pairwise and local Markov properties (with respect to an arbitrary undirected graph $\mathcal{G}$ on U) are equivalent. In this case, if we know $X_i \perp\!\!\!\perp_p X_j \mid X_{U\setminus\{i,j\}}\,[P]$, whenever distinct i and j are not neighbours in $\mathcal{G}$, we can deduce $X_A \perp\!\!\!\perp_p X_B \mid X_C\,[P]$ whenever $A \perp\!\!\!\perp_{ug} B \mid C\,[\mathcal{G}]$.

Such ‘theorem-proving’ techniques have formed a major and fruitful application of the theory of graphical separoids to date. Of course, they apply equally (and generally more straightforwardly) to other ‘range separoids’, not just probabilistic independence.

The connections between graphical representations and probabilistic conditional independence are also fundamental in the theory of *Probabilistic Expert Systems* [13]. This too can be axiomatized at a more general level, and so applied in other separoid contexts [37,41,53].

6.2.2. Multi-information

For $A \subseteq U$, let P_A denote the distribution for X_A induced by P, and $P_i := P_{\{i\}}$ ($i \in U$). Also, let $Q_A := \otimes\{P_i\colon i \in A\}$ denote the distribution for X_A having the same marginals P_i as does P, but under which the (X_i) are independent. We suppose that P_A is absolutely continuous with respect to Q_A for $A = U$, and thus for all $A \subseteq U$.

For $A \subseteq U$ the *multi-information* $I(A)$ (with respect to P) is defined as the Kullback–Leibler divergence between P_A and Q_A. It can be shown [60] that I is a supermodular function on the (Boolean, hence modular) lattice $(2^U, \subseteq)$. Consequently, by a dual version of corollary 4.10 it defines a basic (but not necessarily strong) separoid relation $\perp\!\!\!\perp_s\,[I]$ on $(2^U, \subseteq)$. In its equivalent semi-graphoid form, this is determined by: $x \amalg_s y \mid z[I]$ if $I(x \cup z) + I(y \cup z) = I(x \cup y \cup z) + I(z)$ (x, y, z pairwise disjoint). This can then be shown to be identical with the probabilistic semi-graphoid generated from $\perp\!\!\!\perp_p\,[P]$ using the construction in section 3.2.1 above.

As an example, suppose $U = \{1, 2, 3\}$ and, under P, $X_1 = X_2 = X_3$ with probability 1, each taking value 0 or 1 with probability one-half each. We then obtain the separoid given in example 4.1 – indeed, in terms of the increasing submodular function ϕ of that example and the (modular) height function $h(x) \equiv |x|$, $I(x) = \log 2 \times \{2h(x) - \phi(x)\}$, so that $\perp\!\!\!\perp_s\,[I]$ is the same as $\perp\!\!\!\perp_s\,[\phi]$. Note that, although

$X_1 \perp\!\!\!\perp_p X_1 \mid X_2\,[P]$, we do *not* have $\{1\} \perp\!\!\!\perp_s \{1\} \mid \{2\}\,[I]$ – such 'functional' properties of $\perp\!\!\!\perp_p\,[P]$ disappear in the transition to $\perp\!\!\!\perp_s\,[I]$.

The definition of multi-information used above depends heavily on the choice of a specific collection of random variables $(X_i\colon\ i \in U)$, and then applies only to random variables of the form X_A. While it would be pleasant to be able to extend this construction to define a supermodular function on the full lattice $(\mathbf{F}, \leqslant)$ of all random variables on Ω, it seems doubtful that this could be done; and even if it could, it is not clear that it could be used to define a separoid relation on $(\mathbf{F}, \leqslant)$, considering that neither is this lattice modular (as would be required to apply corollary 4.10), nor is I decreasing (as would be required for the dual version of theorem 4.8).

7. Some logical irrelevance relations

Here we introduce some irrelevance relations that depend only on the structure of the basic space, without requiring any additional ingredients such as Probability.

7.1. Variation independence

Variation independence is a fundamental logical irrelevance property, arising naturally in a number of applications. Its theory has been most specifically studied in the context of relational databases, where it is very closely related to the concept of *embedded multivalued dependency* [35,50,71].

We again deal with the lattice $(\mathbf{F}, \leqslant)$ of functions $X, Y, \dots$ defined on an arbitrary non-empty domain space Ω. However, we do not now need any underlying σ-field or probability measure. We define the *conditional image of X, given $Y = y$* by: $R(X \mid Y = y) := \{X(\omega)\colon\ \omega \in \Omega, Y(\omega) = y\}$. This is nonempty for any $y \in R(Y) := Y(\Omega)$, the (*unconditional*) *image* of Y. When the conditioning variable Y is clear, we may write simply $R(X \mid y)$. The conditional image represents the residual *logical* uncertainty about X after learning $Y = y$, and is analogous to the conditional distribution of X given $Y = y$ in a probabilistic model.

Again closely analogous to probabilistic conditional independence (as expressed, for example, in (i)–(iii) of section 6), we write $X \perp\!\!\!\perp_v Y \mid Z\,[\Omega]$ (or, if Ω is understood, just $X \perp\!\!\!\perp_v Y \mid Z$), and say that X *is variation* (*conditionally*) *independent of Y given Z* (*on* Ω), if any of the following properties (easily shown to be equivalent) hold:

(i) For any $(y, z) \in R(Y, Z)$,

$$R(X \mid y, z) = R(X \mid z).$$

(ii) There exists a function $a(z)$ such that, for any $(y, z) \in R(Y, Z)$,

$$R(X \mid y, z) = a(z).$$

(iii) For any $z \in R(Z)$,

$$R(X, Y \mid z) = R(X \mid z) \times R(y \mid z).$$

It is straightforward to verify that, with the above definitions, $(\mathbf{F}, \leqslant, \perp\!\!\!\perp_v)$ is a basic strong separoid. The corresponding orthogonoid relation $\perp_v$ is then given by:

$$X \perp_v Y\,[\Omega] \Leftrightarrow X \perp\!\!\!\perp_v Y \mid (X \wedge Y)\,[\Omega].$$

7.1.1. Weak conditional independence

Let P be a discrete probability measure, and let Ω be the support of P, so that $P(\{\omega\}) > 0$ for all $\omega \in \Omega$. It is then easy to see that $X \perp\!\!\!\perp_p Y \mid Z\,[P]$ can hold only if $X \perp\!\!\!\perp_v Y \mid Z\,[\Omega]$. In this context, Wong and Butz [72] have introduced a seemingly new property they term 'weak conditional independence' between X and Y given Z, which does not require $X \perp\!\!\!\perp_v Y \mid Z\,[\Omega]$ (they work in terms of the isomorphic separoid relation $\perp\!\!\!\perp_e$ introduced in section 7.2 below, rather than directly with $\perp\!\!\!\perp_v$). However, their definition may be seen to be equivalent to $X \perp\!\!\!\perp_p Y \mid (X \vee Z) \wedge (Y \vee Z)\,[P]$ (so requiring $(X \vee Z) \perp_v (Y \vee Z)\,[\Omega]$). That is, on re-specifying the conditioning variable appropriately, weak conditional independence reverts to ordinary conditional independence.

7.2. Equivalence relations

The set $\mathbf{E}$ of equivalence relations on a set Ω forms a lattice under *refinement:* $\sim$ is *less refined than* $\sim'$, written $\sim \leqslant \sim'$, if $\omega \sim' \eta \Rightarrow \omega \sim \eta$.

The *composition*, $\sim \circ \sim'$, of $\sim$ and $\sim'$ is defined by:

$$\omega(\sim \circ \sim')\eta \quad \Leftrightarrow \quad \text{there exists } \zeta \in \Omega \text{ with } w \sim \zeta \text{ and } \zeta \sim' \eta.$$

Typically, this relation is not an equivalence relation. It will be so if and only if $\sim$ and $\sim'$ *commute, i.e.,* $\sim \circ \sim' = \sim' \circ \sim$ – a property we denote by $\sim \perp_e \sim'$. In that case $\sim \circ \sim' = \sim' \circ \sim = \sim \wedge \sim'$.

Definition 7.1. We write $\sim \perp\!\!\!\perp_e \sim' \mid \sim''$, and call $\sim$ *equivalence* (*conditionally*) *independent of* $\sim'$ *given* $\sim''$, if:

$$(\sim \vee \sim'') \circ (\sim' \vee \sim'') = \sim'' .$$

Equivalently, we require:

$$(\sim \vee \sim'') \perp_e (\sim' \vee \sim'') \quad \text{and} \tag{30}$$

$$(\sim \vee \sim'') \wedge (\sim' \vee \sim'') = \sim'' . \tag{31}$$

Note that, with this definition, $\sim \perp_e \sim'$ is equivalent to $\sim \perp\!\!\!\perp_e \sim' \mid (\sim \wedge \sim')$.

There is a natural isomorphism ι between the quotient lattice $(\mathbf{F}/\doteq, \leqslant)$ of functions on Ω (identifying two functions if each is a contraction of the other), and the lattice $(\mathbf{E}, \leqslant)$ of equivalence relations on Ω: $X \in \mathbf{F}$ corresponds to $\sim \in \mathbf{E}$ under ι if $X(\omega) = X(\omega') \Leftrightarrow \omega \sim \omega'$. In [22] it is shown that ι also supplies an isomorphism between the relations $\perp\!\!\!\perp_v$ and $\perp\!\!\!\perp_e$ defined on these lattices. We immediately deduce that $\perp\!\!\!\perp_e$ is a basic strong separoid relation (equivalently, $\perp_e$ is an orthogonoid relation) on $(\mathbf{E}, \leqslant)$.

7.2.1. Contextual independence

We sometimes wish to consider a *contextual variation independence* property, with the interpretation that X and Y are variation independent given $Z = z$ for certain values of z (say for $z \in \mathcal{Z}_0$), but not necessarily otherwise. This can be expressed in our framework as follows. Define $\Omega_0 := \{\omega \in \Omega\colon\ Z(\omega) \in \mathcal{Z}_0\}$, and $\Omega_1 := \Omega \backslash \Omega_0$. Construct a relation $\sim^*$ on Ω by letting $\omega \sim^* \omega'$ if: *either* $\omega, \omega' \in \Omega_0$ and $Z(\omega) = Z(\omega')$; or $\omega, \omega' \in \Omega_1$ and $X(\omega) = X(\omega'),\ Y(\omega) = Y(\omega'),\ Z(\omega) = Z(\omega')$. This is readily seen to be an equivalence relation. Let the function $Z^* \in \mathbf{F}$ correspond, under the isomorphism ι, to the equivalence relation $\sim^* \in \mathbf{E}$. The desired contextual conditional independence can then be represented by the ordinary variation conditional independence property $X \perp\!\!\!\perp_v Y \mid Z^* \, [\Omega]$. In the discrete sample space setting of section 7.1.1, we can similarly define *contextual probabilistic independence* between X and Y given $Z = z$, for $z \in \mathcal{Z}_0$ (the basic idea can be extended to more general probability spaces – we refrain from going into detail). By combining the analysis of section 7.1.1 with that given here, we can go on to re-express the property of 'contextual weak independence', as defined by Wong and Butz [72], as an ordinary conditional independence property.

7.3. τ-fields

A class $\mathcal{T}$ of subsets of a set Ω forms a *τ-field* (or *complete algebra*) if it is closed under complementation and arbitrary (not just countable) unions – and thus, also, under arbitrary intersections. The set $\mathbf{T}$ of all τ-fields on Ω forms a lattice when ordered by inclusion.

Definition 7.2. We write $\mathcal{T} \perp\!\!\!\perp_q \mathcal{T}' \mid \mathcal{T}''$, and say $\mathcal{T}$ is *qualitatively independent of $\mathcal{T}'$ given $\mathcal{T}''$* if:

$$\left.\begin{array}{l} \alpha \in \mathcal{T} \vee \mathcal{T}'' \\ \beta \in \mathcal{T}' \vee \mathcal{T}'' \\ \alpha \cap \beta = \emptyset \end{array}\right\} \quad \Rightarrow \quad \text{there exists } \gamma \in \mathcal{T}'' \text{ such that } \alpha \subseteq \gamma,\ \beta \subseteq \overline{\gamma}.$$

When $\mathcal{T}''$ is trivial (in which case we may write $\mathcal{T} \perp\!\!\!\perp_q \mathcal{T}'$) the above condition becomes: $\alpha \in \mathcal{T}, \beta \in \mathcal{T}', \alpha \cap \beta = \emptyset \Rightarrow$ either $\alpha = \emptyset$ or $\beta = \emptyset$. When restricted to σ-fields, this is essentially the relation of qualitative independence due to Rényi [49]. The above extension is based on [4].

It is shown in [22], again by exhibiting an isomorphism with variation independence, that $(\mathbf{T}, \leqslant, \perp\!\!\!\perp_q)$ is a basic strong separoid.

8. Statistical independence

Let $\mathcal{P}$ be a family of distributions on $(\Omega, \mathcal{F})$. We can then define a new separoid relation $\perp\!\!\!\perp_S [\mathcal{P}]$ on $(\mathcal{S}, \subseteq)$ as the intersection of all the relations $\perp\!\!\!\perp_p [P]$ for $P \in \mathcal{P}$. That is, $\mathcal{A} \perp\!\!\!\perp_S \mathcal{B} \mid \mathcal{C} \, [\mathcal{P}]$ if $\mathcal{A} \perp\!\!\!\perp_p \mathcal{B} \mid \mathcal{C} \, [P]$, all $P \in \mathcal{P}$. In this case we call $\mathcal{A}$ *statistically* *independent of $\mathcal{B}$ given $\mathcal{C}$, with respect to $\mathcal{P}$.* We now have $\mathcal{A} \preceq \mathcal{B}\,[\mathcal{P}]$ if and only if $\mathcal{A} \preceq \mathcal{B}\,[P]$, all $P \in \mathcal{P}$.

We can again extend the above definition to random variables, which we henceforth consider. For $X \in \mathbf{F}$, $P \in \mathcal{P}$, let P_X denote the marginal distribution of X under P. For $X \in \mathbf{F}$, let π_X denote the function on $\mathcal{P}$ that maps P to P_X. In particular, $\pi_X \in \mathbf{G}$, the set of all functions on $\mathcal{P}$, which we endow with its ordering by functional contraction. Finally, let π denote the function on $\mathbf{F}$ that maps X to π_X. We note that π is increasing, in the sense:

$$X \preceq Y\,[\mathcal{P}] \Rightarrow \pi_X \leqslant \pi_Y, \tag{32}$$

$$\text{so that} \quad X \simeq Y\,[\mathcal{P}] \Rightarrow \pi_X \doteq \pi_Y. \tag{33}$$

We further have:

Lemma 8.1. Suppose $X \perp\!\!\!\perp_S Y \mid X \wedge Y\,[\mathcal{P}]$ and $X \wedge Y \leqslant Z \leqslant Y$. Then

$$\pi_{X \vee Y} \doteq \pi_X \vee \pi_Y \quad \text{and} \tag{34}$$

$$\pi_{(X \vee Z) \wedge Y} \doteq \pi_Z. \tag{35}$$

Proof. We have $X \perp\!\!\!\perp_p Y \mid (X \wedge Y)\,[P]$ for any $P \in \mathcal{P}$. But, under such conditional independence, the joint distribution of (X, Y) not only determines, but is entirely determined by, its margins for X and for Y. Thus (34) follows.

We further have, for all $P \in \mathcal{P}$, $(X \vee Z) \perp\!\!\!\perp_p Y \mid Z\,[P]$, whence $(X \vee Z) \wedge Y \preceq Z\,[P]$. Thus $(X \vee Z) \wedge Y \preceq Z\,[\mathcal{P}]$, and (35) now follows from (32) and the fact that $Z \leqslant (X \vee Z) \wedge Y$, so that

$$(X \vee Z) \wedge Y \simeq Z\,[\mathcal{P}]. \tag{36}$$

□

8.1. Meta independence

The above definition of statistical independence is of limited interest without further elaboration to account for the structure of the family $\mathcal{P}$.

Definition 8.1. We say X is *meta orthogonal to Y with respect to* $\mathcal{P}$, and write $X \perp_\mu Y\,[\mathcal{P}]$, if:

$$X \perp\!\!\!\perp_S Y \mid X \wedge Y\,[\mathcal{P}], \quad \text{and} \tag{37}$$

$$\pi_X \perp\!\!\!\perp_v \pi_Y \mid \pi_{X \wedge Y}\,[\mathcal{P}]. \tag{38}$$

Note that (38) expresses *variation independence*, considering π_X etc. as functions on the domain $\mathcal{P}$. For the motivation behind this definition, and some statistical applications, see [25].

Theorem 8.2. The relation $\perp_\mu\ [\mathcal{P}]$ is an orthogonoid on $(\mathbf{F}, \leqslant)$.

Proof. Conditions G1 and G2 for an orthogonoid are readily seen to hold.

For G3 and G4, suppose $X \perp_\mu Y\,[\mathcal{P}]$, and $X \wedge Y \leqslant Z \leqslant Y$ (so that, in particular, $X \wedge Z \doteq X \wedge Y$). From (37) we obtain $X \perp\!\!\!\perp_S Z \mid (X \wedge Z)\,[\mathcal{P}]$. In particular, from (34) we obtain

$$x_{X \vee Z} \doteq \pi_X \vee \pi_Z. \tag{39}$$

Also, from (38), since $\pi_Z \leqslant \pi_Y$, $\pi_X \perp\!\!\!\perp_v \pi_Z \mid \pi_{X \wedge Y}\,[\mathcal{P}]$, i.e., $\pi_X \perp\!\!\!\perp_v \pi_Z \mid \pi_{X \wedge Z}\,[\mathcal{P}]$. Thus G3 holds.

Also, since $\pi_{X \wedge Y} \leqslant \pi_Z \leqslant \pi_Y$, (38) gives $\pi_X \perp\!\!\!\perp_v \pi_Y \mid \pi_Z\,[\mathcal{P}]$, whence $(\pi_X \vee \pi_Z) \perp\!\!\!\perp_v \pi_Y \mid \pi_Z\,[\mathcal{P}]$, so that, using (39) and (35), $\pi_{X \vee Z} \perp\!\!\!\perp_v \pi_Y \mid \pi_{(X \vee Z) \wedge Y}\,[\mathcal{P}]$. Since we have $(X \vee Z) \perp\!\!\!\perp_S Y \mid Z\,[\mathcal{P}]$, G4 now follows from (36).

Finally, suppose the conditions for G5 hold. We readily find $X \perp\!\!\!\perp_S Y \mid (X \wedge Y)\,[\mathcal{P}]$. Also, we have $\pi_X \perp\!\!\!\perp_v \pi_Z \mid \pi_{X \wedge Z}\,[\mathcal{P}]$, while, from $\pi_{X \vee Z} \perp\!\!\!\perp_v \pi_Y \mid \pi_Z\,[\mathcal{P}]$, we derive $\pi_X \perp\!\!\!\perp_v \pi_Y \mid \pi_Z\,[\mathcal{P}]$. We deduce $\pi_X \perp\!\!\!\perp_v \pi_Y \mid \pi_{X \wedge Z}\,[\mathcal{P}]$, whence G5 follows since $X \wedge Z \doteq X \wedge Y$. □

Corollary 8.3 (Meta conditional independence). Define a relation $\perp\!\!\!\perp_\mu\,[\mathcal{P}]$ on $\mathcal{S}$ by: $X \perp\!\!\!\perp_\mu Y \mid Z\,[\mathcal{P}]$ if:

$$X \perp\!\!\!\perp_S Y \mid Z\,[\mathcal{P}], \quad \text{and} \tag{40}$$

$$\pi_{X \vee Z} \perp\!\!\!\perp_v \pi_{Y \vee Z} \mid \pi_Z\,[\mathcal{P}]. \tag{41}$$

Then $\perp\!\!\!\perp \mu$ is a separoid on $(\mathbf{F}, \preceq\,[\mathcal{P}])$.

Proof. From theorem 3.2, on noting that (40) implies $(X \vee Z) \wedge (Y \vee Z) \simeq Z\,[\mathcal{P}]$, whence also $\pi_{(X \vee Z) \wedge (Y \vee Z)} \doteq \pi_Z$. □

Remark. If we add the further condition $(X \vee Z) \wedge (Y \vee Z) \doteq Z$, we obtain a separoid on $(\mathbf{F}, \leqslant)$.

8.2. Hyper independence

The theory of section 8.1 applies just as well if we use any other separoid relation on $\mathbf{G}$ in (38), in place of variation independence. In particular, suppose we have a distribution Π defined over (a suitable σ-field in) $\mathcal{P}$ – we call such a distribution for a distribution a *law* over $\mathcal{P}$. We can then replace (38) by:

$$\pi_X \perp\!\!\!\perp_p \pi_Y \mid \pi_{X \wedge Y}\,[\Pi],$$

and, correspondingly, (41) by:

$$\pi_{X \vee Z} \perp\!\!\!\perp_p \pi_{Y \vee Z} \mid \pi_Z\,[\Pi].$$

The resulting relations $\perp_H$ [Π], $\perp\!\!\!\perp_H$ [Π] are termed *hyper* orthogonality and conditional independence. These have important applications in Bayesian statistical inference [25].

9. Natural independence

There are a number of uncertainty formalisms related to 'possibility theory' supporting various concepts of irrelevance, many of which lead to separoids [2,9].

Specifically, in Spohn's theory of 'deterministic epistemology' [59], uncertainty over a space Ω is represented by a *natural conditional function* $\kappa : \Omega \to \{0, 1, \ldots\}$. A point ω with $\kappa(\omega) = n$ is "implausible to degree n" (informally, it has probability of order ε^n). The implausibility of $A \subseteq \Omega$ is $\kappa(A) := \inf\{\kappa(x)\colon x \in A\}$. The implausibility of A given B is $\kappa(A \mid B) := \kappa(A \cap B) - \kappa(B)$. Note that $-\kappa$ is a possibility function on 2^Ω: the following analysis will apply essentially unchanged whenever this holds.

Given such a natural conditional function κ on Ω, we introduce the following relation $\perp\!\!\!\perp_N$ [κ] on $(\mathbf{F}, \leqslant)$:

Definition 9.1. We say that X *is naturally* (*conditionally*) *independent of* Y *given* Z (*with respect to* κ), and write $X \perp\!\!\!\perp_N Y \mid Z\,[\kappa]$, if:

(i) $X \perp\!\!\!\perp_v Y \mid Z\,[\Omega]$, and

(ii) $\kappa(X = x \mid Y = y,\ Z = z) = \kappa(X = x \mid Z = z)$, all $(x, y, z) \in R(X, Y, Z)$.

We can alternatively express (ii), in an obvious notation, as:

(ii)′ $\kappa(x, y, z) + \kappa(z) \equiv \kappa(x, z) + \kappa(y, z)$.

Lemma 9.1. When (i) holds, (ii) is equivalent to the existence of some function $a(\cdot)$ such that

(ii)* $\kappa(X = x \mid Y = y,\ Z = z) = a(x, z)$, all $(x, y, z) \in R(X, Y, Z)$.

Proof. Clearly (ii) $\Rightarrow$ (ii)*. So suppose (ii)* holds. We obtain $\kappa(x, y, z) \equiv \kappa(y, z) + a(x, z)$. Fix x, z, and take infima over y, noting that the range of such y may depend on z, but, by (i), does not further depend on x. We obtain $\kappa(x, z) \equiv \kappa(z) + a(x, z)$, whence $a(x, z) \equiv \kappa(x, z) - \kappa(z) \equiv \kappa(x \mid z)$. $\square$

The following theorem generalises related results in a semi-graphoid setting [36, 64].

Theorem 9.2. The relation $\perp\!\!\!\perp_N$ [κ] defines a strong separoid on $(\mathbf{F}, \leqslant)$.

Proof. P1 and P2 easily follow from (i) and (ii)′.

Now suppose that $X \perp\!\!\!\perp_N Y \mid Z\,[\kappa]$, and $W \leqslant Y$. By P2 we have $Y \perp\!\!\!\perp_N X \mid Z\,[\kappa]$. Then $\kappa(w \mid x, z) = \min\{\kappa(y \mid x, z)\colon Y = y \Rightarrow W = w\}$, and by (ii) this is

$\min\{\kappa(y \mid z)\colon Y = y \Rightarrow W = w\} = \kappa(w, z)$. Also, from (i) and the separoid properties of variation independence, $W \perp\!\!\!\perp_v Y \mid Z\,[\Omega]$. Thus $W \perp\!\!\!\perp_N X \mid Z[\kappa]$, and so, using P2 again, $X \perp\!\!\!\perp_N W \mid Z[\kappa]$.

Furthermore, for $(x, y, z, w) \in R(X, Y, Z, W)$ (which requires $Y = y \Rightarrow W = w$), $\kappa(x \mid y, z, w) = \kappa(x \mid y, z)$, and this is $\kappa(x \mid z)$ by (ii). Since $X \perp\!\!\!\perp_v Y \mid (Z \vee W)\,[\Omega]$, and $\kappa(x \mid z)$ can be regarded as a function of (x, w, z), lemma 9.1 now implies $\kappa(x \mid y, z, w) = \kappa(x \mid w, z)$, and thus P4 holds.

Now suppose $X \perp\!\!\!\perp_N Y \mid Z[\kappa]$ and $X \perp\!\!\!\perp_N W \mid (Y \vee Z)\,[\kappa]$. From (i) we obtain $X \perp\!\!\!\perp_v (Y \vee W) \mid Z\,[\Omega]$. Also, for $(x, y, w, z) \in R(X, Y, W, Z)$, we have successively $\kappa(x \mid y, w, z) = \kappa(x \mid y, z) = \kappa(x \mid z)$, and P5 follows.

Finally, suppose $X \perp\!\!\!\perp_N Y \mid Z\,[\kappa]$ and $X \perp\!\!\!\perp_N Y \mid W\,[\kappa]$, with both $Z \leqslant Y$ and $W \leqslant Y$. From (i) we have $X \perp\!\!\!\perp_N Y \mid (Z \wedge W)\,[\kappa]$. Also, for $(x, y) \in R(X, Y), \kappa(x \mid y) = \kappa(x \mid z) = \kappa(x \mid w)$, where $w,\ z$ are the values of $W,\ Z$ when $Y = y$. Fixing x, we see that, as members of **F**, $\kappa(x \mid Y) \leqslant Z$, and $\kappa(x \mid Y) \leqslant W$. Hence $\kappa(x \mid Y) \leqslant Z \wedge W$. Writing $U := Z \wedge W$, with value u when $Y = y$, we deduce $\kappa(x \mid y, u)$ has the form $a(x, u)$, whence P6 follows from lemma 9.1. $\square$

10. Covariance independence

An important application of definition 4.3 is to the case that $\mathcal{V}$ is the space of square-integrable random variables on a probability space $(\Omega, \mathcal{F}, P)$, endowed with the implied 'covariance (semi-)inner product' C: $C(X, Y) = E_P(XY) - E_P(X)E_P(Y)$ [55]. In this case, for linear spaces $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ of such random variables, $\mathcal{X} \perp\!\!\!\perp_O \mathcal{Y} \mid \mathcal{Z}\,[C]$ if, for any $X \in \mathcal{X}$, a best linear predictor based on $\mathcal{Y}$ and $\mathcal{Z}$ can be taken to be based on $\mathcal{Z}$ alone; or, equivalently, if the partial correlation between any $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, after allowing for $\mathcal{Z}$, is 0. The use of graphical representations (see section 5) to manipulate properties of $\perp\!\!\!\perp_O$ in this context has been considered by Smith [55,56], Goldstein [32] and Wilkinson [70].

In the case $P(\{\omega\}) > 0$ for all $\omega \in \Omega$, so that C is a proper inner product (*modulo* additive constants), we have $X \perp\!\!\!\perp_p Y \mid Z\,[P]$ if and only if $\mathcal{X} \perp\!\!\!\perp_O \mathcal{Y} \mid \mathcal{Z}\,[C]$, where $\mathcal{X}$ is now the collection of all square-integrable functions of X, etc.

11. Belief independence

In the Dempster–Shafer theory of belief functions [28,51], uncertainty over a space Ω is expressed by means of a probability distribution (the *basic belief distribution*) M for a random subset $\mathbf{S}$ of Ω. In the 'discrete case', when M puts all its mass on a finite or countable number of sets, this is typically described by its *basic belief assignment* m, the set function given by $m(S) = M(\mathbf{S} = S)$ for $S \subseteq \Omega$. We then call S a *focal set* of Bel if $m(S) > 0$. Usually it is supposed that $M(\mathbf{S} = \emptyset) = 0$, although in the 'transferable belief model' (TBM) [54] this requirement is removed. The *belief* $\mathrm{Bel}(A)$ in an event $A \subseteq \Omega$ is $\mathrm{Bel}(A) := M(\mathbf{S} \subseteq A)$, or $M(\mathbf{S} \subseteq A,\ \mathbf{S} \neq \emptyset)$ in TBM. For $X \in \mathbf{F}$, a belief function on $R(X)$ is equivalent, in an obvious way, to its 'vacuous extension' to Ω, a belief function on Ω supported on sets in the τ-field $\mathcal{T}_X$ generated by X (so that, in the discrete case, all its focal sets belong to $\mathcal{T}_X$), and we shall not distinguish between these.

For $X \in \mathbf{F}$ and $S \subseteq \Omega$, let $S_X := \{X(\omega)\colon \omega \in S\}$ denote the image of S under X. The *marginal belief function* Bel_X of Bel on $R(X)$ is defined as that having basic belief distribution M_X, the distribution of $\mathbf{S}_X$ when $\mathbf{S}$ has distribution M.

According to 'Dempster's rule', the *combination*, $\mathrm{Bel}^* = \mathrm{Bel}_1 \oplus \mathrm{Bel}_2$, of two belief functions Bel_1 and Bel_2 over Ω is given, in terms of their corresponding basic belief distributions M^*, M_1 and M_2, by first generating $\mathbf{S}_1$ and $\mathbf{S}_2$, independently, from M_1 and M_2, respectively; and then taking for M^* the distribution of $\mathbf{S}_1 \cap \mathbf{S}_2$, conditional on $\mathbf{S}_1 \cap \mathbf{S}_2 \neq \emptyset$. In TBM this conditioning is omitted, leading to the "conjunctive combination" $\mathrm{Bel}_1 \circledcap \mathrm{Bel}_2$. If Bel is a belief function on $R(X)$, then, for any Bel' on Ω, $(\mathrm{Bel} \oplus \mathrm{Bel}')_X = \mathrm{Bel} \oplus (\mathrm{Bel}')_X$, and similarly for $\circledcap$.

Shafer [51] introduced a concept of 'evidential independence' between two variables X and Y defined on Ω, with respect to a belief function Bel. Ben Yaghlane et al. [5,6] show that this is equivalent to:

(i) with probability one for $\mathbf{S}$ generated from the basic belief distribution M, $X \perp\!\!\!\perp_v Y\ [\mathbf{S}]$; together with:

(ii) $\mathbf{S}_X \perp\!\!\!\perp_p \mathbf{S}_Y\ [M]$.

Condition (i) requires that (with probability 1 under M – in the discrete case, for every focal set) we will obtain a set S which is a 'rectangle', in the sense that X and Y are variation independent on S; while (ii) requires that the 'sides' of this random rectangle vary independently under the probability distribution M. Under TBM, the independence condition (ii) has to be replaced by conditional independence, given $\mathbf{S} \neq \emptyset$.

An alternative expression of the same property is *non-interactivity* [5], defined by the property:

$$\mathrm{Bel}_{X\vee Y} = \mathrm{Bel}_X \oplus \mathrm{Bel}_Y \tag{42}$$

together with the implicit condition $X \perp\!\!\!\perp_v Y\ [\Omega]$. For TBM this becomes

$$\mathrm{Bel}_{X\vee Y} \circledcap \mathrm{Bel}_0 = \mathrm{Bel}_X \circledcap \mathrm{Bel}_Y, \tag{43}$$

where Bel_0 denotes the margin under Bel for the constant random variable 0, whose basic belief function M_0 gives positive probability to the empty set if M does. When $M(\emptyset) = 0$ this term can be omitted, (43) then becoming equivalent to (42).

Ben Yaghlane et al. [5] also introduce a property of *doxastic independence*, related to conditioning, and show that this too is equivalent to non-interactivity.

Although the above marginal independence definition appears generally appropriate, its extension to a suitable definition of conditional independence is more problematic, and there are several possible choices. Ben Yaghlane et al. [7,8] introduce and discuss some of these, in particular *conditional non-interactivity*, expressed by the following generalisation of (42):

$$\mathrm{Bel}_{X\vee Y\vee Z} \oplus \mathrm{Bel}_Z = \mathrm{Bel}_{X\vee Z} \oplus \mathrm{Bel}_{Y\vee Z} \tag{44}$$

together with the implicit assumption $X \perp\!\!\!\perp_v Y \mid Z\,[\Omega]$. The equivalent property for TBM is obtained by replacing $\oplus$ by $\textcircled{\cap}$ in (44). Following Shenoy [52], they demonstrate that (when specialised to coordinate projection functions on a product space) this definition satisfies the graphoid axioms.

However, the property (44) does not appear fully satisfactory, since the combination operation on each side involves independent, rather than identical, items of evidence about Z. An alternative definition of belief conditional independence, formally analogous to meta or hyper independence, can be developed by generalising the description of evidential independence as follows:

Definition 11.1. Let Bel be a belief function on Ω, and M its basic belief distribution. We say that X is *belief* (*conditionally*) *independent of* Y *given* Z (*with respect to Bel*), and write $X \perp\!\!\!\perp_b Y \mid Z\,[\mathrm{Bel}]$, if:

(i) With probability one for $\mathbf{S}$ generated from the distribution M (in the discrete case, for every focal set),

$$X \perp\!\!\!\perp_v Y \mid Z\,[\mathbf{S}].$$

(ii) $\mathbf{S}_{X\vee Z} \perp\!\!\!\perp_p \mathbf{S}_{Y\vee Z} \mid \mathbf{S}_Z\,[M]$.

It is easy to see that, when $X \perp\!\!\!\perp_b Y \mid Z\,[\mathrm{Bel}]$, $\mathrm{Bel}_{X\vee Y\vee Z}$ is fully determined by $\mathrm{Bel}_{X\vee Z}$ and $\mathrm{Bel}_{Y\vee Z}$. Using arguments similar to those of section 8.1, it can be shown that $(\mathbf{F}, \leqslant, \perp\!\!\!\perp_b)$ forms a separoid.

The belief conditional independence property of definition 11.1 is quite distinct from conditional non-interactivity as given by (44). For example, let $\Omega = \{0, 1\}^3$, with typical point $\omega = xyz$; and let $X,\ Y,\ Z$ be the coordinate projection functions: $X(\omega) = x$, etc. Suppose that M puts equal probability one-third on each of the sets $\{001\},\ \{110\}$, and $\{000,\ 111\}$. It is readily checked that $X \perp\!\!\!\perp_b Y \mid Z\,[\mathrm{Bel}]$. However, the set $\{000\}$ can be seen to be a focal set for $\mathrm{Bel} \oplus \mathrm{Bel}_Z$, but not for $\mathrm{Bel}_{X\vee Z} \oplus \mathrm{Bel}_{Y\vee Z}$ showing that these belief functions differ.

12. Non-symmetric irrelevance

Although many ‘irrelevance’ concepts are well described by the mathematical structure of the separoid, this is not universally so. In particular, the symmetry axiom P2 is sometimes not appropriate. Moreover, in this the case we have to split each of the other axioms into at least two versions, depending on how we order the first two terms in the relation $\perp\!\!\!\perp$.

The necessity for a non-symmetric relation was already apparent for the application of conditional independence to parameters in statistical models, as treated by Dawid [16,19]. Similar considerations apply in influence diagrams, and in many other instances: for example, martingale theory, and the theory of imprecise probabilities [12,14,69]. Two cases of asymmetric irrelevance have attracted particular attention in recent years. The first is causal irrelevance [31]: intuitively – whatever definition of causality is taken – if one variable is causally irrelevant for another, the reverse need not hold. The second example involves prediction in temporal structures, which underlies Granger (non-)causality [33]. The basic irrelevance concept here is that the past of a process $Y(t)$ is uninformative for prediction of another process $X(t)$ – which can hold even when the reverse does not. Local independence [1] extends this concept to event history data. Didelez [29] has developed a set of axioms satisfied by this particular non-symmetric irrelevance notion, and a related graphical representation, involving a non-symmetric notion of separation. A similar formal and graphical investigation has been undertaken by Vantaggi [67,68] for a non-symmetric concept of stochastic independence involving null events [11]. These are promising first steps along the road towards establishing a suitable framework for non-symmetric irrelevance more generally.

13. Conclusion

There has been widespread study of properties and applications of specific structures related to separoids, especially semi-graphoids and graphical models. It is hoped that, by developing a more general mathematical framework to support this enterprise, this paper may lead to cross-fertilisation across different fields, and so promote the further development of the subject area.

New concepts of 'irrelevance', in a wide variety of settings, have been proliferating rapidly, and this can be expected to continue. It should be regarded as a matter of course, when any such new definition is proposed, that a check be made to see whether it satisfies the separoid axioms P1–P5. If not, departures from the axioms should be carefully identified and justified.

Acknowledgements

I am grateful to Ross Shachter for suggesting the usefulness of the construction in equation (5); to Philippe Smets for helpful comments on section 11; to František Matúš and Milan Studený for many valuable discussions over several years, and in particular to František for suggesting the semi-graphoid representation of section 3.2.2; and to all the participants at the Workshop on Conditional Independence Structures and Graphical Models, held at the Fields Institute for Research in Mathematical Sciences, Toronto, in September 1999, for the stimulus they provided to develop the general ideas presented here.

References

[1] O.O. Aalen, Dynamic modeling and causality, Scandinavian Actuarial Journal (1987) 177–190.

[2] N.B. Amor, S. Benferhat, D. Dubois, H. Geffner and H. Prade, Independence for qualitative uncertain reasoning (abstract), in: *Abstracts of Workshop on Conditional Independence Structures and Graphical Models*, eds. F. Matúš and M. Studený, Toronto, Canada (The Fields Institute for Research in Mathematical Sciences, 1999) pp. 2–3.

[3] S.A. Andersson, D. Madigan and M.D. Perlman, An alternative Markov property for chain graphs, in: *Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence*, eds. E. Horvitz and F.V. Jensen (Morgan Kaufmann, San Francisco, CA, 1996) pp. 40–48.

[4] P. Bartfai and T. Rudas, Conditionally independent extension of measures, Preprint 57/1988, Mathematical Institute, Hungarian Academy of Sciences (1988).
[5] B. Ben Yaghlane, P. Smets and K. Mellouli, Belief function independence: I. The marginal case, Technical Report 2000-13, Institut de Recherches Interdisciplinaires et de Développements en Intelligence Artificielle, Université Libre de Bruxelles (2000).
[6] B. Ben Yaghlane, P. Smets and K. Mellouli, Independence concepts for belief functions, in: *Proceedings of 8th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2000)*, Madrid (2000) pp. 357–364.
[7] B. Ben Yaghlane, P. Smets and K. Mellouli, On conditional belief function independence, in: *Workshop on Partial Knowledge and Uncertainty: Independence, Conditioning, Inference*, eds. R. Scozzafava and B. Vantaggi, Rome (2000).
[8] B. Ben Yaghlane, P. Smets and K. Mellouli, Belief function independence: II. The conditional case, Technical Report, Institut de Recherches Interdisciplinaires et de Développements en Intelligence Artificielle, Université Libre de Bruxelles (2001).
[9] S. Benferhat, D. Dubois and H. Prade, Possibilistic independence and plausible reasoning, in: *Foundations and Applications of Possibility Theory*, Advances in Fuzzy Systems – Applications and Theory, Vol. 8, eds. G. de Cooman, D. Ruan and E.E. Kerre (World Scientific, Singapore, 1995) pp. 47–63.
[10] G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, Revised edn. (MacMillan, New York, 1953).
[11] G. Coletti and R. Scozzafava, Zero probabilities in stochastic independence, in: *Information, Uncertainty and Fusion*, eds. B. Bouchon-Meunier, R.R. Yager and L.A. Zadeh (Kluwer, Dordrecht, 2000) pp. 185–196.
[12] I. Couso, S. Moral and P. Walley, Examples of independence for imprecise probabilities, in: *Proceedings of the First International Symposium on Imprecise Probabilities and their Applications*, eds. G. de Cooman, F. Cozman, S. Moral and P. Walley, Ghent, Belgium (1999) pp. 121–130.
[13] R.G. Cowell, A.P. Dawid, S. Lauritzen and D.J. Spiegelhalter, *Probabilistic Networks and Expert Systems* (Springer, New York, 1999).
[14] F.G. Cozman, Irrelevance and independence axioms in quasi-Bayesian theory, in: *Symbolic and Quantitative Approaches to Reasoning with Uncertainty – European Conference, ECSQARU'99*, eds. A. Hunter and S. Parsons, Lecture Notes in Artificial Intelligence, Vol. 1638 (Springer, Berlin, 1999) pp. 128–136.
[15] A.P. Dawid, Properties of diagnostic data distributions, Biometrics 32 (1976) 647–658.
[16] A.P. Dawid, Conditional independence in statistical theory (with Discussion), Journal of the Royal Statistical Society, Series B 41 (1979) 1–31.
[17] A.P. Dawid, Some misleading arguments involving conditional independence, Journal of the Royal Statistical Society, Series B 41 (1979) 249–252.
[18] A.P. Dawid, A Bayesian look at nuisance parameters, in: *Bayesian Statistics*, eds. J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith (Valencia University Press, Valencia, 1980) pp. 167–184.
[19] A.P. Dawid, Conditional independence for statistical operations, Annals of Statistics 8 (1980) 598–617.
[20] A.P. Dawid, Invariance and independence in multivariate distribution theory, Journal of Multivariate Analysis 17 (1985) 304–315.
[21] A.P. Dawid, Conditional independence, in: *Encyclopedia of Statistical Science (Update Volume 2)*, eds. S. Kotz, C.B. Read and D.L. Banks, (Wiley Interscience, New York, 1998) pp. 146–155.
[22] A.P. Dawid, Some variations on variation independence, in: *Artificial Intelligence and Statistics 2001*, eds. T. Jaakkola and T. Richardson (Morgan Kaufmann, San Francisco, CA, 2001) pp. 187–191.
[23] A.P. Dawid and J.M. Dickey, Likelihood and Bayesian inference from selectively reported data, Journal of the American Statistical Association 72 (1977) 845–850.
[24] A.P. Dawid and I.W. Evett, Using a graphical method to assist the evaluation of complicated patterns of evidence, Journal of Forensic Science 42 (1997) 226–231.

[25] A.P. Dawid and S.L. Lauritzen, Hyper Markov laws in the statistical analysis of decomposable graphical models, Annals of Statistics 21 (1993) 1272–1317.
[26] A.P. Dawid and J. Mortera, Coherent analysis of forensic identification evidence, Journal of the Royal Statistical Society, Series B 58 (1996) 425–443.
[27] A.P. Dawid and M. Studený, Conditional products: An alternative approach to conditional independence, in: *Artificial Intelligence and Statistics 99*, eds. D. Heckerman and J. Whittaker (Morgan Kaufmann, San Francisco, CA, 1999) pp. 32–40.
[28] A.P. Dempster, A generalisation of Bayesian inference (with Discussion), Journal of the Royal Statistical Society, Series B 30 (1968) 205–247.
[29] V. Didelez, Local independence graphs for composable Markov processes, Discussion Paper 158, SFB 386, University of Munich (1999).
[30] M. Frydenberg, The chain graph Markov property, Scandinavian Journal of Statistics 17 (1990) 333–353.
[31] D. Galles and J. Pearl, Axioms of causal relevance, Artificial Intelligence 97 (1997) 9–43.
[32] M. Goldstein, Influence and belief adjustment, in: *Influence Diagrams, Belief Nets and Decision Analysis*, eds. R.M. Oliver and J.Q. Smith (Wiley, Chichester, 1990) 143–174.
[33] C. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37 (1969) 424–438.
[34] G. Grätzer, *General Lattice Theory* (Academic Press, 1978).
[35] J.R. Hill, Comment on Speigelhalter et al. (1993), Statistical Science 8 (1993) 258–261.
[36] D. Hunter, Graphoids and natural conditional functions, International Journal of Approximate Reasoning 5 (1991) 489–504.
[37] F.V. Jensen, Inference in directed acyclic graphs, Research Report R-97-5011, Institute for Electronic Systems, Aalborg University (1997).
[38] J.T.A. Koster, Markov properties of non-recursive causal models, Annals of Statistics 24 (1996) 2148–2177.
[39] J.T.A. Koster, Marginalizing and conditioning in graphical models. Technical Report EUR/FSW-Soc/2000.02, Erasmus University, Rotterdam (2000).
[40] S.L. Lauritzen, A.P. Dawid, B.N. Larsen and H.G. Leimer, Independence properties of directed Markov fields, Networks 20 (1990) 491–505.
[41] S.L. Lauritzen and F.V. Jensen, Local computation with valuations from a commutative semigroup, Annals of Mathematics and Artificial Intelligence 21 (1997) 51–69.
[42] F. Matúš, Probabilistic conditional independence structures and matroid theory: Background, International Journal of General Systems 22 (1994) 185–196.
[43] J.G. Oxley, *Matroid Theory* (Oxford University Press, Oxford, 1992).
[44] A. Paz, R.Y. Geva and M. Studený, Representation of irrelevance relations by annotated graphs, Fundamenta Informaticae 42 (2000) 149–199.
[45] A. Paz and J. Pearl, Axiomatic characterization of directed graphs. Technical Report R-234, Cognitive Systems Laboratory, Computer Science Department, University of California, Los Angeles (1994).
[46] A. Paz, J. Pearl and S. Ur, A new characterization of graphs based on interception relations, Journal of Graph Theory 22 (1996) 125–136.
[47] J. Pearl, *Probabilistic Reasoning in Intelligent Systems* (Morgan Kaufmann, San Mateo, CA, 1988).
[48] J. Pearl and A. Paz, Graphoids: A graph-based logic for reasoning about relevance relations, in: *Advances in Artificial Intelligence*, eds. D. Hogg and L. Steels (North-Holland, Amsterdam, 1987) pp. 357–363.
[49] A. Rényi, *Foundations of Probability* (Holden-Day, San Francisco, CA, 1970).
[50] Y. Sagiv and S.F. Walecka, Subset dependencies and a completeness result for a subclass of embedded multivalued dependencies, Journal of the Association for Computing Machinery 29 (1982) 103–117.
[51] G. Shafer, *A Mathematical Theory of Evidence* (Princeton University Press, 1976).
[52] P.P. Shenoy, Conditional independence in valuation-based systems, International Journal of Approximate Reasoning 10 (1994) 203–234.

[53] P.P. Shenoy and G. Shafer, Axioms for probability and belief function propagation, in: *Uncertainty in Artificial Intelligence*, eds. R.D. Shachter, T.S. Levitt, L.N. Kanal and J.F. Lemmer (North-Holland, Amsterdam, 1990) pp. 169–198.
[54] P. Smets and R. Kennes, The transferable belief model, Artificial Intelligence 66 (1994) 191–234.
[55] J.Q. Smith, Influence diagrams for statistical modelling, Annals of Statistics 17 (1989) 654–672.
[56] J.Q. Smith, Statistical principles on graphs, in: *Influence Diagrams, Belief Nets and Decision Analysis*, eds. R.M. Oliver and J.Q. Smith (Wiley, Chichester, 1990) pp. 89–120.
[57] D.J. Speigelhalter, A.P. Dawid, S.L. Lauritzen and R.G. Cowell, Bayesian analysis in expert systems (with Discussion), Statistical Science 8 (1993) 219–283.
[58] W. Spohn, Stochastic independence, causal independence, and shieldability, Journal of Philosophical Logic 9 (1980) 73–99.
[59] W. Spohn, Ordinal conditional functions: a dynamic theory of epistemic states, in: *Causation in Decision, Belief Change, and Statistics*, eds. W.L. Harper and B. Skyrms (Kluwer Academic, Dordrecht 1988) pp. 105–134.
[60] M. Studený, Multiinformation and the problem of characterization of conditional independence relations, Problems of Control and Information Theory 18 (1989) 3–16.
[61] M. Studený, Conditional independence relations have no finite complete characterization, in: *Transactions of the 11th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes* (Academia, Prague, 1992) pp. 377–396.
[62] M. Studený, Formal properties of conditional independence in different calculi of AI, in: *Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, eds. M. Clarke, R. Kruse and S. Moral (Springer, Berlin, 1993) pp. 341–348.
[63] M. Studený, Structural semigraphoids, International Journal of General Systems 22 (1994) 207–217.
[64] M. Studený, Conditional independence and natural conditional functions, International Journal of Approximate Reasoning 12 (1995) 934–945.
[65] M. Studený, Comparison of graphical approaches to description of conditional independence structures, in: *Proceedings of WUPES'97*, January 22–25, 1997, Prague, Czech Republic (1997) pp. 156–172.
[66] M. Studený and R.R. Bouckaert, On chain graph models for description of conditional independence structures, Annals of Statistics 26 (1998) 1434–1495.
[67] B. Vantaggi, Conditional independence in a coherent finite setting, Annals of Mathematics and Artificial Intelligence 32 (2001) 287–313.
[68] B. Vantaggi, The l-separation criterion for description of cs-independence models, Technical Report, Dipartimento di Metodi e Modelli Matematici, University of Rome "La Sapienza", submitted to International Journal of Approximate Reasoning (2001).
[69] P. Walley, *Statistical Reasoning with Imprecise Probabilities* (Chapman and Hall, London, 1991).
[70] D.J. Wilkinson, An object-oriented approach to local computation in Bayes linear belief networks, in: *Proceedings in Computational Statistics*, eds. P.J. Green and R.W. Payne (Physica Verlag, Heidelberg, 1998) pp. 491–496.
[71] S.K.M. Wong, An extended relational data model for probabilistic reasoning, Journal of Intelligent Information Systems 9 (1997) 181–202.
[72] S.K.M. Wong and C.J. Butz, Contextual weak independence in Bayesian networks, in: *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, eds. K.B. Laskey and H. Prade (Morgan Kaufmann, San Francisco, CA, 1999) pp. 670–679.