Graphoids and separoids in model theory

Irene Córdoba-Sánchez ICORDOBA@FI.UPM.ES
Concha Bielza MCBIELZA@FI.UPM.ES
Pedro Larrañaga PEDRO.LARRANAGA@FI.UPM.ES
Departamento de Inteligencia Artificial
Universidad Politecnica de Madrid
Campus de Montegancedo, s/n
28660 Boadilla del Monte, Madrid, Spain

Abstract

We treat graphoid and separoid structures within the mathematical framework of model theory, specially suited for representing and analysing axiomatic systems with multiple semantics. We represent the graphoid axiom set in model theory, and translate algebraic separoid structures to another axiom set over the same symbols as graphoids. This brings both structures to a common, sound theoretical ground where they can be fairly compared. Our contribution further serves as a bridge between the most recent developments in formal logic research, and the well-known graphoid applications in probabilistic graphical modelling.

Keywords: Graphoid, separoid, conditional independence, mathematical logic

1. Introduction

Probabilistic graphical models were originated at the interface between statistics, where Markov random fields were predominant, and artificial intelligence, with a focus on Bayesian networks. Since then, a wide spectrum of models has grown (chain graphs, ancestral graphs, etc.), and all of them share the same foundational concept: the relationship between statistical independence in a probability distribution and graph separation. This correspondence was originally made explicit via the set of *graphoid* axioms (Pearl and Paz, 1987). Informally, they are rules that hold for both conditional independence (Dawid, 1979) and certain separation criteria in graphs (Pearl, 1988). Their satisfiability is not only limited to the context of probabilistic graphical models; in fact, they usually apply whenever a notion of 'irrelevance' is being mathematically treated (Dawid, 2001), although sometimes they are too restrictive (Cozman and Walley, 2005). Such apparent similarities led to the definition of *separoids* (Dawid, 2001), aimed at unifying those mathematical notions of irrelevance. A separoid is an algebraic structure consisting of a ternary relation I (representing irrelevance) operating on a set. The properties of I parallel, although are not equivalent to, the graphoid axioms (Dawid, 2001).

In this paper we frame graphoids and separoids in model theory (Marker, 2000), which describes and reasons about formal structures using the tools of mathematical logic. The distinguishing characteristic of model theory is an explicit separation between syntactical or symbolic objects, and context-specific semantics. Our contribution is thus twofold. On the one hand, we exploit the natural suitability of model theory for treating axiom systems (syntax) with varied applications (semantics), such as graphoids. On the other hand, we translate the algebraic separoid structure to an axiom set over the same symbols used for graphoids, thus bringing both structures to a common, sound theoretical ground (model theory) where they can be fairly compared. Furthermore, recently researchers in mathematical logic have defined a new logic for intuitively reasoning about independence (Grädel and Väänänen, 2013), which is being actively developed and studied (Hannula et al., 2016). This logic resembles the separoid structure, rather than the more well-known graphoid axioms. Thus, our contribution serves as a bridge between these recent formal developments and graphoid applications in probabilistic graphical modelling, which can be exploited for the potential benefit of both research areas.

The rest of the paper is structured as follows. Section 2 gives a preliminary introduction to model theory. Separoids and graphoids are both defined in Section 3. We formalize them within model theory in Sections 4 and 5, using two complementary approaches, which are unified in Section 6. In Section 7 we summarize the main conclusions and discuss prospective research directions. Appendix A contains auxiliary results used throughout the paper.

2. Preliminaries in model theory

The main purpose of model theory is to describe mathematical structures from the perspective of logic. Intuitively, a mathematical structure consists of a set, together with special functions, relations, and elements. These are represented with a chosen language.

Definition 1. A set of symbols L is called a *first-order language* when it contains:

- *Relation*, *constant* and *function* symbols. For relations and functions a strictly positive number n must be specified, called its *arity* (n-ary relation or function). Intuitively, this represents the number of elements operated or related by the represented function or relation, respectively. Sometimes, for technical convenience, constants are considered as functions with 0 arity.

- Logical symbols: *variables* ($x, y, z, v_1, v_2 \ldots$), *connectors* ($\neg$, $\rightarrow$) and the *universal quantifier* ($\forall$).

- A 2-ary relation symbol, $=$, called *equality*.

The set of function symbols (including constants) is denoted as $L_{\mathcal{F}}$, the set of relation symbols (including $=$) as $L_{\mathcal{R}}$. Logical symbols and the relation $=$ are always present in all first-order languages, and thus they are not explicitly specified when defining them.

Just as first-order languages represent the syntax in model theory, the semantics is provided by *interpreting* those symbols according to the mathematical structure we want to describe.

Definition 2. Let L be a first-order language. A L-*structure* $\mathcal{A}$ is composed of:

- a non-empty set A;

- a set $R^{\mathcal{A}} \subseteq A^m$ for each m-ary relation $R \in L_{\mathcal{R}}$;

- a function $f^{\mathcal{A}} : A^n \mapsto A$ for each n-ary function $f \in L_{\mathcal{F}}$,

and denoted as $\mathcal{A} = \langle A, \{f^{\mathcal{A}}\}_{f \in L_{\mathcal{F}}}, \{R^{\mathcal{A}}\}_{R \in L_{\mathcal{R}}} \rangle$. A is called the *universe* of $\mathcal{A}$, and for each $s \in L_{\mathcal{F}} \cup L_{\mathcal{R}}$, $s^{\mathcal{A}}$ is called the *interpretation* of s in $\mathcal{A}$. $=^{\mathcal{A}}$ is always interpreted as the equality among elements in A.

Example 1. Consider the first-order language $L = \{\leq\}$, where $\leq$ is a 2-ary relation symbol, aimed at describing ordered mathematical structures. The following are representations of standard mathematical structures as L-structures:

- $\mathcal{R} := \langle \mathbb{R}, \leq^{\mathcal{R}} \rangle$, where $\leq^{\mathcal{R}}$ is the usual less or equal than relation in $\mathbb{R}$.
- $\mathcal{S} := \langle 2^S, \leq^{\mathcal{S}} \rangle$, where S is a set, 2^S denotes its power set, and $\leq$ is interpreted as the set inclusion.
- $\mathcal{V} := \langle V, \leq^{\mathcal{V}} \rangle$ where V is the vertex set of an acyclic digraph and $(u, v) \in \leq^{\mathcal{V}}$ when v is reachable from u by a directed path.

2.1 Syntax: Formulas and theories

Our aim is to use finite sequences of the symbols in a first-order language L to construct logical statements about L-structures. Thus, we need to establish which of such finite sequences are going to be considered valid. These valid sequences are called formulas of L, which are in turn defined using the more basic notion of language term:

Definition 3. The *set of terms* of L, $\mathrm{Ter}(L)$, is the smallest set of finite sequences of symbols in L containing the variables and such that if $t_1, \ldots, t_n \in \mathrm{Ter}(L)$ and $f \in L_{\mathcal{F}}$ is n-ary, then $ft_1 \ldots t_n \in \mathrm{Ter}(L)$.

Definition 4. An *atomic formula* is $Rt_1 \ldots t_m$ where $R \in L_{\mathcal{R}}$ is m-ary and $t_i \in \mathrm{Ter}(L)$ for $1 \leq i \leq m$. The *set of formulas* of L, $\mathrm{For}(L)$, is the smallest set of finite sequences of symbols in L containing the atomic formulas and such that if $F, G \in \mathrm{For}(L)$, then $\neg F$, $F \to G$ and $\forall v F$ belong to $\mathrm{For}(L)$. The following abbreviations are commonly used: if $F, G \in \mathrm{For}(L)$, $F \vee G$ for $\neg F \to G$, $F \wedge G$ for $\neg(\neg F \vee \neg G)$, $F \leftrightarrow G$ for $(F \to G) \wedge (G \to F)$ and $\exists x$ for $\neg \forall x \neg$.

If $R \in L_{\mathcal{R}}$ is 2-ary, we will write the atomic formula $R t_1 t_2$ as $t_1 R t_2$; for m-arity such formulas will be sometimes parenthesized as $R(t_1, \ldots, t_m)$. Analogously, when $f \in L_{\mathcal{F}}$ is 2-ary, ft_1t_2 will be sometimes denoted as $t_1 f t_2$, whereas function symbols of any arity (except constants) will appear parenthesized most times.

In order to arrive at our main syntactic object of interest, we first need to formalize the notion of closed formulas. This allows us to define L-theories, which, intuitively, are sets of formulas stating properties about the functions and relations in our mathematical structure of interest.

Definition 5. Let $F \in \mathrm{For}(L)$. The *set of sub-formulas* of F, $\mathrm{SubFor}(F)$, is recursively defined as: $\{F\}$ if F is atomic, $\{F\} \cup \mathrm{SubFor}(G)$ if F is $\neg G$ or $\forall v G$, and $\{F\} \cup \mathrm{SubFor}(G) \cup \mathrm{SubFor}(H)$ if F is $G \to H$. An occurrence of a variable v in F is said to be *free* if it does not exist $G \in \mathrm{SubFor}(F)$ such that $\forall v G \in \mathrm{SubFor}(F)$. The variable v is said to be free in

F if it has a free occurrence in F. We indicate that the free variables of F are contained in $\{x_1, \ldots, x_n\}$ as $F(x_1, \ldots, x_n)$. F is said to be a *closed formula* if it contains no free variables, and the set of closed formulas of L is denoted as $\overline{\mathrm{For}}(L)$. If F is $F(x_1, \ldots, x_n)$, then $\forall x_1 \ldots \forall x_n\, F$ is called the *universal closure* of F.

Definition 6. A set $\mathrm{T} \subseteq \overline{\mathrm{For}}(L)$ is called a *L-theory*, or *theory in L*. An element $F \in \mathrm{T}$ is called an *axiom* of T.

Example 2. If $L = \{\leq\}$, $\mathrm{Ter}(L)$ consists on the set of variables, since there is no function symbol in L. Thus, an atomic formula, composed only by terms, is for example $x \leq y$, where the variables x and y are free. If instead $L = \{\sqcup, \sqcap\}$, where $\sqcup$ and $\sqcap$ are 2-ary function symbols, then $x \sqcup y$ is a non-variable term. A formula in such language would be $\forall x \forall y\, x \sqcup y = x \sqcap y$; in fact, this is a closed formula, and thus could be part of a L-theory.

2.2 Semantics: Satisfiability and models

We will now formalize how formulas are interpreted; that is, how to provide semantics for those sequences of symbols with respect to a L-structure. The building blocks of formulas, terms, are interpreted directly with respect to the L-structure, whereas formulas rely on those term interpretations and specified logical symbol semantics.

Definition 7. Let $t \in \mathrm{Ter}(L)$, where the variables appearing in t are $x_1, \ldots, x_n$. The interpretation of t is recursively defined as the function $t^{\mathcal{A}} : A^n \mapsto A$ such that for each $\bar{a} = (a_1, \ldots, a_n) \in A^n$, $t^{\mathcal{A}}(\bar{a}) = a_i$ if t is the variable x_i, and $t^{\mathcal{A}}(\bar{a}) = f^{\mathcal{A}}({t_1}^{\mathcal{A}}(\bar{a}), \ldots, {t_l}^{\mathcal{A}}(\bar{a}))$ if t is $ft_1 \ldots t_l$ with f a l-ary function symbol.

Definition 8. Let $\mathcal{A}$ be a L-structure, $F(x_1, \ldots, x_n) \in \mathrm{For}(L)$ and $\bar{a} = (a_1, \ldots, a_n) \in A^n$. *$F$ is satisfied in $\mathcal{A}$ for $\bar{a}$*, and denoted as $\mathcal{A} \models F(\bar{a})$, if:

- When F is $Rt_1 \ldots t_n$ with $R \in L_{\mathcal{R}}$, R n-ary, and $t_1, \ldots, t_n \in \mathrm{Ter}(L)$, that is, when F is atomic: $({t_1}^{\mathcal{A}}(\bar{a}), \ldots, {t_n}^{\mathcal{A}}(\bar{a})) \in R^{\mathcal{A}}$;
- When F is $\neg G$: $\mathcal{A} \not\models G(\bar{a})$ ($\not\models$ means not satisfied);
- When F is $G \to H$: $\mathcal{A} \models G(\bar{a})$ implies $\mathcal{A} \models H(\bar{a})$;
- When F is $\forall x G$: for all $d \in A$, $\mathcal{A} \models G(\bar{a}, d)$ (d substitutes a_i if x is x_i).

Note that abbreviations $(\wedge, \vee, \leftrightarrow, \exists)$ obtain their usual semantics (and, or, iff, exists).

Finally, we have arrived at our main semantic object of interest: models for L-theories. Intuitively, these are the mathematical structures that comply with the explicit properties that we have described closed formulas in a theory.

Definition 9. Let T be a L-theory. The *set of models* of T is $\mathrm{Mod}(\mathrm{T}) \coloneqq \{\mathcal{A}\ L\text{−structure} : \mathcal{A} \models F \text{ for all } F \in \mathrm{T}\}$. The notation $\mathcal{A} \in \mathrm{Mod}(\mathrm{T})$ is used interchangeably with $\mathcal{A} \models \mathrm{T}$.

A summary of the preliminaries introduced in this section is presented in Figure 1.

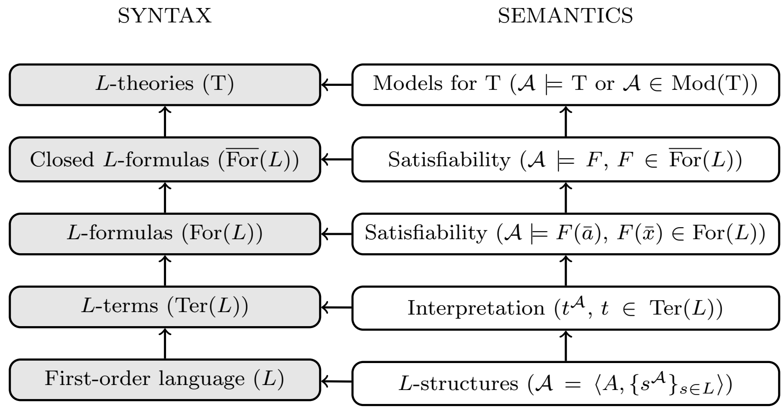

Figure 1: Key concepts in model theory, separated by syntax (gray) and semantics (white).

3. Lattices, separoids and graphoids

Separoid structures are special types of *lattices* (Grätzer, 2003). A lattice is a *partially ordered set* (poset) in which the join/supremum or meet/infimum exist between every two elements. When the former occurs, the structure is called *join semi-lattice*, whereas if the latter holds, it is called *meet semi-lattice*. When both situations occur, the set is said to be *lattice ordered*.

Definition 10. A *semi-separoid* is a set A with join semi-lattice structure, together with a collection I of triples of A, satisfying, for all $a, b, c, d \in A$,

- $(a, b, a) \in I$;
- if $(a, b, c) \in I$ then $(b, a, c) \in I$;
- if $d \leq b$ and $(a, b, c) \in I$ then $(a, c, d) \in I$ and $(a, b, \sup\{c, d\}) \in I$;
- if $(a, b, c) \in I$ and $(a, d, \sup\{b, c\}) \in I$ then $(a, \sup\{b, d\}, c) \in I$;

where $\sup\{\cdot\}$ denotes the supremum function. The stronger notion of *separoid* involves a lattice instead of a semi-lattice, and has the additional property that if $c \leq b$ and $d \leq b$, then $(a, b, c) \in I$ and $(a, b, d) \in I$ implies $(a, b, \inf\{c, d\}) \in I$, being $\inf\{\cdot\}$ the infimum function.

Definition 11. A *semi-graphoid* is a set I of triples of pairwise disjoint subsets of a finite set V, satisfying the following properties, for all $A, B, C, D \in 2^V$,

- $(A, \emptyset, A) \in I$;
- if $(A, B, C) \in I$ then $(B, A, C) \in I$;
- if $(A, B \cup D, C) \in I$ then $(A, B, C) \in I$, $(A, D, C) \in I$ and $(A, B, C \cup D) \in I$;
- if $(A, B, C \cup D) \in I$ and $(A, D, B \cup C) \in I$ then $(A, B \cup D, C) \in I$.

It is called a *graphoid* if, additionally, when $(A, B, C \cup D) \in I$ and $(A, D, B \cup C) \in I$, then $(A, B \cup D, C) \in I$.

Two important differences directly arise between separoid structures and graphoid axioms. The latter involve the notion of *pairwise disjointness*, while the former does not. By contrast, separoid structures are a explicit extension of semi-lattices, while graphoid axioms are indirectly related: the power set 2^V is a finite lattice where, given $A, B \in 2^V$, their join is $A \cup B$ and their meet is $A \cap B$. Thus, the graphoid axioms seem to define special cases of separoid structures (as claimed in Dawid (2001)). The first-order theories we will construct will make explicit these preliminary observations.

A final remark, regarding lattice theory, is relevant. Recall, from the beginning of this section, the definition of lattices as posets. An alternative, equivalent way of characterizing them is as algebraic structures with two operations satisfying certain dual properties (or one operation, if they are semi-lattices, see Grätzer, 2003). As such, this gives rise to alternative characterizations depending on how their underlying lattice structure is represented. Separoid structures are directly expressed as lattice ordered sets, while the graphoid axioms rely on the operations $\cup$ and $\cap$.

4. Separoids as lattice ordered sets

We will use $L_{Ord} \coloneqq \{\leq\}$ (see Example 1) for formalizing lattices as posets. Consider the following closed L_{Ord}-formulas:

$$\forall x\, x \leq x, \tag{\textbf{O1}}$$

$$\forall x \forall y\, x \leq y \wedge y \leq z \rightarrow x \leq z, \tag{\textbf{O2}}$$

$$\forall x \forall y\, x \leq y \wedge y \leq x \rightarrow x = y, \tag{\textbf{O3}}$$

$$\forall x \forall y \exists z\, x \leq z \wedge y \leq z \wedge (\forall w\, x \leq w \wedge y \leq w \rightarrow z \leq w), \tag{\textbf{O4}}$$

$$\forall x \forall y \exists z\, z \leq x \wedge z \leq y \wedge (\forall w\, w \leq x \wedge w \leq y \rightarrow w \leq z). \tag{\textbf{O5}}$$

O1, **O2** and **O3** syntactically express the *reflexive*, *transitive* and *antisymmetric* properties of posets, respectively, while **O4** and **O5** denote the existence of joins and meets, respectively. Thus, we define the *partial order* theory as PO $\coloneqq \{\mathbf{O1}, \ldots, \mathbf{O3}\}$, the *join semi-lattice* theory as JL $\coloneqq$ PO $\cup \{\mathbf{O4}\}$, the *meet semi-lattice* theory as ML $\coloneqq$ PO $\cup \{\mathbf{O5}\}$, and the *lattice-ordered set* theory OL $\coloneqq$ JL $\cup$ ML$=$ PO $\cup \{\mathbf{O4}, \mathbf{O5}\}$. Models for these theories represent the lattice structures as described in the beginning of Section 3.

Example 3. The L_{Ord}-structures in Example 1 are models for some of the previous theories. $\mathcal{R} \models$ OL and $\mathcal{S} \models$ OL. Indeed, joins (meets) are maximum (minimum) elements and union (intersection) sets, respectively. $\mathcal{V} \models$ PO, but it may happen that $\mathcal{V} \not\models \mathbf{O4}$ (Figure 2a) or $\mathcal{V} \not\models \mathbf{O5}$ (Figure 2b).

Consider the abbreviation, for expressing that we are referring to the join (meet) element, such that **O4** (**O5**) would be written $\forall x \forall y \exists z\, z = \sup\{x, y\}$ ($\forall x \forall y \exists z\, z = \inf\{x, y\}$). Using the expanded language $L_{OrdSep} \coloneqq L_{Ord} \cup \{I\} = \{\leq, I\}$, where I is a 3-ary relation, we can

(a) Violation of **O4** (b) Violation of **O5**

Figure 2: Non-transitive acyclic digraphs.

directly formulate the properties of I, in Definition 10, as:

$$\forall x \forall y\, I(x, y, x) \tag{\textbf{OS1}}$$
$$\forall x \forall y \forall z\, I(x, y, z) \rightarrow I(y, x, z) \tag{\textbf{OS2}}$$
$$\forall x \forall y \forall z \forall w\, w \leq y \wedge I(x, y, z) \rightarrow I(x, z, w) \tag{\textbf{OS3}}$$
$$\forall x \forall y \forall z \forall w \forall t\, w \leq y \wedge I(x, y, z) \wedge t = \sup\{z, w\} \rightarrow I(x, y, t) \tag{\textbf{OS4}}$$
$$\forall x \forall y \forall z \forall w \forall t \forall s\, I(x, z, w) \wedge I(x, y, t) \wedge t = \sup\{z, w\} \wedge s = \sup\{y, w\} \rightarrow I(x, s, z), \tag{\textbf{OS5}}$$
$$\forall x \forall y \forall z \forall w \forall t\, I(x, y, z) \wedge I(x, y, w) \wedge z \leq y \wedge w \leq y \wedge t = \inf\{z, w\} \rightarrow I(x, y, t). \tag{\textbf{OS6}}$$

Semi-separoids and separoids are thus represented by models for the L_{OrdSep}-theories OSS $:=$ JL $\cup$ $\{\mathbf{OS1}, \ldots, \mathbf{OS5}\}$ and OS $:=$ OL $\cup$ $\{\mathbf{OS1}, \ldots, \mathbf{OS6}\}$, respectively. The ‘O’ has been prepended to indicate that the underlying lattice is characterized as a poset.

5. Graphoids as algebraic lattices

As discussed in Section 3, graphoids are more related with the characterization of lattices as algebraic structures with two dual operations, instead of as posets. Thus, let $\sqcup$, $\sqcap$ denote 2-ary function symbols and consider the first-order language $L_{AlgLat} := \{\sqcup, \sqcap\}$ (see Example 2) where the following closed formulas are defined:

$$\forall x\, x \sqcup x = x, \quad (\mathbf{L1}) \qquad \forall x\, x \sqcap x = x, \quad (\mathbf{L2})$$
$$\forall x \forall y\, x \sqcup y = y \sqcup x, \quad (\mathbf{L3}) \qquad \forall x \forall y\, x \sqcap y = y \sqcap x, \quad (\mathbf{L4})$$
$$\forall x \forall y \forall z\, x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z, \quad (\mathbf{L5}) \qquad \forall x \forall y \forall z\, x \sqcap (y \sqcap z) = (x \sqcap y) \sqcap z, \quad (\mathbf{L6})$$
$$\forall x \forall y\, x \sqcup (x \sqcap y) = x, \quad (\mathbf{L7}) \qquad \forall x \forall y\, x \sqcap (x \sqcup y) = x. \quad (\mathbf{L8})$$

L1, **L3** and **L5** respectively express *idempotency*, *commutativity* and *associativity* for $\sqcup$, as dually do **L2**, **L4** and **L6** for $\sqcap$. These properties are satisfied by and characteristic of joins/meets in semi-lattices (Grätzer, 2003). **L7** and **L8** relate both symbols and are commonly called the *absorption* laws. We thus define the L_{AlgLat}theory of *(algebraic) semi-lattices* as ASL $:=$ $\{\mathbf{L1}, \mathbf{L3}, \mathbf{L5}\}$, and the theory of *(algebraic) lattices* as AL $=$ $\{\mathbf{L1}, \ldots, \mathbf{L8}\}$. An important notion that we need to syntactically reflect for graphoids is that of *pairwise disjointness*. Define in $L_{AlgLat} \cup \{0\}$, where 0 is a constant symbol, the formulas:

$$\forall x\, x \sqcap 0 = 0. \tag{L19}$$

$$\forall x \forall y \forall z\, x \sqcup (y \sqcap z) = (x \sqcup y) \sqcap (x \sqcup z), \tag{L10}$$

L19 enforces 0 to be a minimal element, while **L10** expresses the distributivity of $\sqcup$ and $\sqcap$. Finite distributive lattices have been shown to be in bijection with finite lattice ordered sets (see Section 4), a result that is known as Birkhoff's representation theorem (Grätzer, 2003); we will define for convenience the theory MDL $\coloneqq$ AL $\cup \{\mathbf{L19}, \mathbf{L10}\}$.

Example 4. Let V be finite and consider $\mathcal{P} \coloneqq \langle 2^V, \sqcup^{\mathcal{P}}, \sqcap^{\mathcal{P}}, 0^{\mathcal{P}} \rangle$, with $\sqcup^{\mathcal{P}}(A, B) = A \cup B$ and $\sqcap^{\mathcal{P}}(A, B) = A \cap B$. Then $\mathcal{P} \models$ AL. By the *syntactic* duality between $\sqcup$ and $\sqcap$, we could have reversed the *semantics*, and defined a distinct L_{AlgLat}-structure $\mathcal{P}'$, with the same universe 2^V, and such that $\sqcup^{\mathcal{P}'}(A, B) = A \cap B$ and $\sqcap^{\mathcal{P}'}(A, B) = A \cap B$, which is also a model for AL. In the former interpretation, if $0^{\mathcal{P}} = \emptyset$, then $\mathcal{P} \models$ MDL; while in the latter interpretation, this happens only if $0^{\mathcal{P}'} = V$.

Consider the extended language $L_{AlgGraph} \coloneqq L_{AlgLat} \cup \{I, 0\}$, where 0 is a constant symbol and I is a 3-ary relation, and abbreviate $\wedge_{i,j \in \{1,\ldots,n\}, i \neq j} x_i \sqcap x_j = 0$ (pairwise disjointness) in $L_{AlgGraph}$ as $\mathrm{dis}(x_1, \ldots, x_n)$. The graphoid axioms (Definition 11) are thus expressed as:

$$\forall x \forall z\ \mathrm{dis}(x, z) \to I(x, 0, z), \tag{G1}$$

$$\forall x \forall y \forall z\ \mathrm{dis}(x, y, z) \wedge I(x, y, z) \to I(y, x, z), \tag{G2}$$

$$\forall x \forall y \forall z \forall w\ \mathrm{dis}(x, y, z) \wedge \mathrm{dis}(x, z, w) \wedge I(x, y \sqcup w, z) \to I(x, z, w), \tag{G3}$$

$$\forall x \forall y \forall z \forall w\ \mathrm{dis}(x, y, z) \wedge \mathrm{dis}(x, z, w) \wedge I(x, y \sqcup w, z) \to I(x, y, z \sqcup w), \tag{G4}$$

$$\forall x \forall y \forall z \forall w\ \mathrm{dis}(x, y, z, w) \wedge I(x, z, w) \wedge I(x, y, z \sqcup w) \to I(x, y \sqcup w, z), \tag{G5}$$

$$\forall x \forall y \forall z \forall w\ \mathrm{dis}(x, y, z, w) \wedge I(x, y, z \sqcup w) \wedge I(x, w, y \sqcup z) \to I(x, y \sqcup w, z). \tag{G6}$$

Mathematical structures complying with graphoid axioms are represented by models for the L_{AlgLat}-theories ASG $\coloneqq$ MDL $\cup \{\mathbf{G1}, \ldots, \mathbf{G5}\}$ (semi-graphoids) and AG $\coloneqq$ ASG $\cup \{\mathbf{G6}\}$.

6. L_{Irr}: unifying first-order languages and theories

Let $L_{Irr} \coloneqq \{\sqcup, \sqcap, \leq, I\} = L_{Ord} \cup L_{AlgLat} \cup \{I\}$. This language combines the symbols for both the algebraic and order-theoretic characterization of underlying lattices in separoids and graphoids. In fact, a correspondence can be established between models for ASL and JL/ML, as well as between AL and OL, that gives raise to an equivalence of the respective expansions in the language $L_{Ord} \cup L_{AlgLat}$ (see Appendix A). This equivalence allows us to find expansions, into L_{Irr}, of models for OSS/OS that satisfy the analogous formulas

$$\forall x \forall y\, I(x, y, x), \tag{S1}$$

$$\forall x \forall y \forall z\, I(x, y, z) \to I(y, x, z), \tag{S2}$$

$$\forall x \forall y \forall z \forall w\, w \leq y \wedge I(x, y, z) \to I(x, z, w), \tag{S3}$$

$$\forall x \forall y \forall z \forall w\, w \leq y \wedge I(x, y, z) \to I(x, y, z \sqcup w), \tag{S4}$$

$$\forall x \forall y \forall z \forall w\, I(x, z, w) \wedge I(x, y, z \sqcup w) \to I(x, y \sqcup w, z), \tag{S5}$$

$$\forall x \forall y \forall z \forall w\, z \leq y \wedge w \leq y \wedge I(x, y, z) \wedge I(x, y, w) \to I(x, y, z \sqcap w), \tag{S6}$$

resorting to Corollary 1 in Appendix A. Furthermore, this process can be inverted to express separoid properties using symbols exclusively in $L_{AlgLat} \cup \{I\}$, as

$$\forall x \forall y\, I(x,y,x), \quad (\mathbf{AS1})$$
$$\forall x \forall y \forall z\, I(x,y,z) \to I(y,x,z), \quad (\mathbf{AS2})$$
$$\forall x \forall y \forall z \forall w\, y \sqcup w = y \wedge I(x,y,z) \to I(x,z,w), \quad (\mathbf{AS3})$$
$$\forall x \forall y \forall z \forall w\, y \sqcup w = y \wedge I(x,y,z) \to I(x,y,z \sqcup w), \quad (\mathbf{AS4})$$
$$\forall x \forall y \forall z \forall w\, I(x,z,w) \wedge I(x,y,z \sqcup w) \to I(x,y \sqcup w,z), \quad (\mathbf{AS5})$$
$$\forall x \forall y \forall z \forall w\, y \sqcup w = y \wedge y \sqcup z = y \wedge I(x,y,z) \wedge I(x,y,w) \to I(x,y,z \sqcap w). \quad (\mathbf{AS6})$$

Again, in this case if we let ASS $:=$ ASL $\cup \{\mathbf{AS1}, \ldots, \mathbf{AS5}\}$ and *algebraic separoids* AS $:=$ AL $\cup \{\mathbf{AS1}, \ldots, \mathbf{AS6}\}$, we can find a expansion for each model of ASS (AS) into L_{Irr} that satisfies $\mathbf{S1}, \ldots, \mathbf{S6}$. We will call the unifying L_{Irr}-theories SS $:=$ JL $\cup$ ASL $\cup \{\mathbf{S1}, \ldots, \mathbf{S5}\}$ and S $:=$ OL $\cup$ AL $\cup \{\mathbf{S1}, \ldots, \mathbf{S6}\}$, the I and *separoid* theories, respectively.

There is a huge resemblance between $\mathbf{AS2}, \ldots, \mathbf{AS5}$ and $\mathbf{G2}, \ldots,$ **G5**: **AS2** and **G2** are identical, and if we directly substitute the formula $x \sqcup y = y$, we obtain

$$I(x, y \sqcup w, z) \to I(x,z,w),$$
$$I(x, y \sqcup w, z) \to I(x,y,z \sqcup w),$$
$$I(x,z,w) \wedge I(x,y,z \sqcup w) \to I(x,y \sqcup w,z),$$

which are the same as **G3**, ..., **G5** omitting the disjointness sub-formulas and universal quantifiers. However, the formula $x \sqcup y = y$ cannot be omitted, since without it there is no correspondence with the statement $x \leq y$ when considering the transformation into the equivalent partial order (Corollary 3 in Appendix A).

A diagram relating this hierarchy of separoid formal characterizations is depicted in Figure 3.

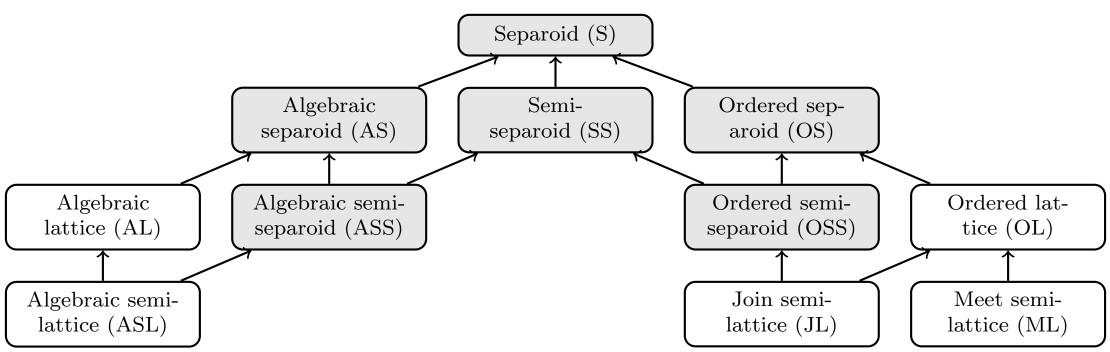

Figure 3: Hierarchy of theories related to separoids (highlighted in gray).

Similarly, formulas $\mathbf{G1}, \ldots, \mathbf{G6}$ can be reformulated in L_{Ord} and L_{Irr} by exploiting the outlined correspondences among lattices and an extension of the results in Appendix A to Birkhoff's representation theorem, obtaining the respective hierarchy of theories (Figure 4).

Comparing both Figures 3 and 4, we see directly that the stronger notion of separoid comes from strengthening the underlying lattice structure in a semi-separoid (from JL to

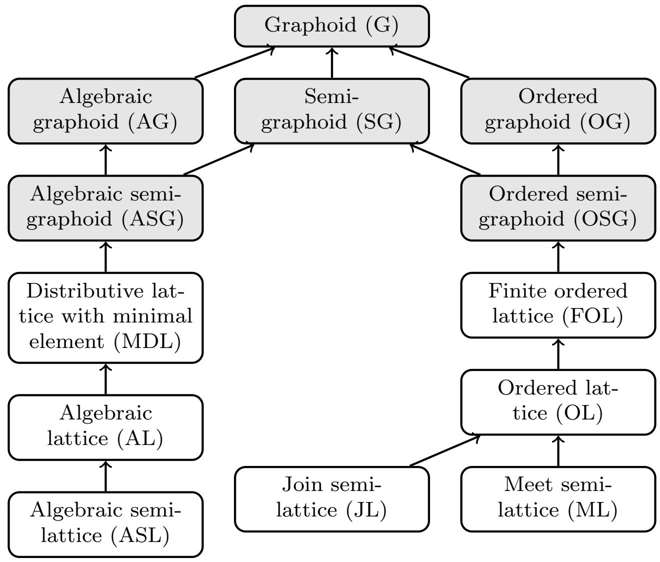

Figure 4: Hierarchy of theories related to graphoids (highlighted in gray).

OL); whereas in the case of graphoids, the underlying lattice structure is always MDL, and the strengthening comes axiomatically from **G6**.

7. Conclusions and future work

We have translated graphoids and separoids into the nomenclature of model theory. Our main objective has been to provide a common conceptual framework under which both structures can be studied and analysed. Separoids and graphoids both appear in contexts where some kind of 'irrelevance' is being treated, and as such we have shown how an underlying axiomatic set-up exists for such models. In addition, we have pointed out the main similarities and differences between them from a purely syntactic point of view.

When considering graphoids as algebraic structures (corresponding in our context to the L_{Irr}-theory SG), Dawid (2001) defined, under some assumptions, the following algebraic correspondence between separoids and graphoids: $(x, y, z) \in I$ if and only if $x \sqcap y \leq z$ and $(x \setminus z, y \setminus z, z) \in I$. This could be used to define a correspondence between models for graphoids and models for separoids in an analogous way to Propositions 2 and 3 in Appendix A. The *independence logic* that is currently being developed in model theory (Hannula et al., 2016) is closer to separoids than graphoids, and has been applied mainly to database modelling. Thus, another future direction of research is to further connect these new findings with graphoid applications, instead of separoids. This would provide a bridge between both research communities to advance towards a universal axiomatization of 'irrelevance', which has been and continues to be the core of probabilistic graphical modelling.

Acknowledgments

Research partially supported by the Spanish Ministry of Economy and Competitiveness through the Cajal Blue Brain (C080020-09; the Spanish partner of the Blue Brain initiative from EPFL) and TIN2013-41592-P projects, by the Regional Government of Madrid through the S2013/ICE-2845-CASI-CAM-CM project, and by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 604102 (Human Brain Project).

Appendix A. Correspondence between models for ASL and JL, ML

The following results are based on well known facts from lattice theory (Grätzer, 2003). We will need an additional concept from model theory: L'-expansions of L-structures, when $L \subseteq L'$.

Definition 12. Let L and L' be first-order languages with $L \subseteq L'$. If $\mathcal{A}$ is a L-structure and $\mathcal{A}'$ is a L'-structure, $\mathcal{A}'$ is said to be a *L'-expansion* of $\mathcal{A}$ when, for all $s \in L$, $s^{\mathcal{A}}$ and $s^{\mathcal{A}'}$ coincide.

Proposition 1. *Let $\mathcal{A} \models \mathrm{JL}$ with universe A and $a, b \in A$. The element $c \in A$ such that $\mathcal{A} \models \sup(a, b, c)$ is unique. Furthermore, for all $a, b, c \in A$, $\mathcal{A} \models \sup(a, b, c)$ if and only if $\mathcal{A} \models \sup(b, a, c)$. These results hold analogously for $\mathcal{A}' \models \mathrm{ML}$ and $\inf(x, y, z)$.*

Proof. The existence is guaranteed by the fact that $\mathcal{A} \models \mathrm{JL}$ and $\mathbf{O4} \in \mathrm{JL}$. This implies that $\mathcal{A} \models \mathbf{O4}$, which by definition of satisfiability leads to the existence, for all $a, b \in A$, of an element $c \in A$ such that $\mathcal{A} \models \sup(a, b, c)$. Assume that c is not unique and denote as $\tilde{c}$ another element in A such that $c \neq \tilde{c}$ and $\mathcal{A} \models \sup(a, b, \tilde{c})$. From $\mathcal{A} \models \sup(a, b, c)$ we get $x \leq^{\mathcal{A}} c$, $y \leq^{\mathcal{A}} c$ and for all $k \in A$, if $x \leq^{\mathcal{A}} k$ and $y \leq^{\mathcal{A}} k$, then $c \leq^{\mathcal{A}} k$; analogously, from $\mathcal{A} \models \sup(a, b, \tilde{c})$ we get $x \leq^{\mathcal{A}} \tilde{c}$, $y \leq^{\mathcal{A}} \tilde{c}$ and for all $k \in A$, if $x \leq^{\mathcal{A}} k$ and $y \leq^{\mathcal{A}} k$, then $\tilde{c} \leq^{\mathcal{A}} k$. Thus, taking k as $\tilde{c}$ and c, we have that $\tilde{c} \leq^{\mathcal{A}} c$ and $c \leq^{\mathcal{A}} \tilde{c}$, respectively. However, since $\mathbf{O3} \in \mathrm{JL}$, we get that $c = \tilde{c}$, which is a contradiction. Finally, the symmetry in $\mathbf{O4}$ directly gives $\mathcal{A} \models \sup(a, b, c)$ if and only if $\mathcal{A} \models \sup(b, a, c)$.

The proof regarding theory ML and $\inf(x, y, z)$ is analogous. □

Proposition 2. *Let $\mathcal{A} \models \mathrm{JL}$ with universe A. Define $\mathcal{B}$ as the L_{AlgLat}-structure with the same universe, A, and such that, for $a, b \in A$, $\sqcup^{\mathcal{B}}(a, b) = c$ where c is such that $\mathcal{A} \models \sup(a, b, c)$. Then, $\mathcal{B}$ is a model for ASL. If $\mathcal{A} \models \mathrm{ML}$ and the interpretation of $\sqcup$ is defined in terms of the element satisfying $\inf(x, y, z)$ instead of $\sup(x, y, z)$, then $\mathcal{B}$ is a model for ASL.*

Proof. Since $\mathcal{A}$ and $\mathcal{B}$ share the universe, $=^{\mathcal{A}}$ coincides with $=^{\mathcal{B}}$, thus we will omit the superscript throughout the proof. Note that Proposition 1 directly gives us that the function $\sqcup^{\mathcal{B}}$ is well defined. Recall that $\mathcal{B}$ is a model for ASL if $\mathcal{B} \models F$ for all $F \in \mathrm{ASL}$. $\mathcal{B} \models \mathbf{L1}$ if for all $a \in A$ we have $\sqcup^{\mathcal{B}}(a, a) = a$. From the definition of $\sqcup^{\mathcal{B}}$, we have that $\mathcal{A} \models \sup(a, a, c)$ for some unique $c \in A$, thus, it is enough to show that $\mathcal{A} \models \sup(a, a, a)$. Using the definition of satisfiability, we arrive at it by observing that $a \leq^{\mathcal{A}} a$, since $\mathcal{A} \models \mathbf{O1}$. Satisfiability of $\mathbf{L3}$ is directly obtained from Proposition 1, which gives that, for all $a, b \in A$, $\sqcup^{\mathcal{B}}(a, b) = \sqcup^{\mathcal{B}}(b, a)$. Finally, $\mathbf{L5}$ is interpreted as $\sqcup^{\mathcal{B}}(a, e) = \sqcup^{\mathcal{B}}(d, c)$, where $d = \sqcup^{\mathcal{B}}(a, b)$ and $e = \sqcup^{\mathcal{B}}(b, c)$. Let $f = \sqcup^{\mathcal{B}}(d, c)$ and $g = \sqcup^{\mathcal{B}}(a, e)$. It is enough to see that $\mathcal{A} \models \sup(a, e, f)$ and $\mathcal{A} \models \sup(d, b, g)$.

As an example, $a \leq^{\mathcal{A}} f$ follows from $a \leq^{\mathcal{A}} d$ ($\mathcal{A} \models \sup(a, b, d)$), $d \leq^{\mathcal{A}} f$ ($\mathcal{A} \models \sup(d, c, f)$) and **L3**, interpreted accordingly; $e \leq^{\mathcal{A}} f$, $d \leq^{\mathcal{A}} g$ and $b \leq^{\mathcal{A}} g$ follow analogously. Assume now a $k \in A$ such that $a \leq^{\mathcal{A}} k$ and $e \leq^{\mathcal{A}} k$; we want to arrive at $f \leq^{\mathcal{A}} k$, since we would have proved $\mathcal{A} \models \sup(a, e, f)$. Since $\mathcal{A} \models \sup(b, c, e)$ and $\mathcal{A} \models$ **O2**, we have $b \leq^{\mathcal{A}} k$ and $c \leq^{\mathcal{A}} k$. From the former, $a \leq^{\mathcal{A}} k$ and $\mathcal{A} \models \sup(a, b, d)$, we get $d \leq^{\mathcal{A}} k$; this, combined with the latter and the fact that $\mathcal{A} \models \sup(d, c, f)$, finally leads to $f \leq^{\mathcal{A}} k$. The proof for $\mathcal{A} \models \sup(d, b, g)$ proceeds in an analogous way.

When $\mathcal{A} \models$ ML and, for $a, b \in A$, $\sqcup^{\mathcal{B}}(a, b) = c$ with c such that $\mathcal{A} \models \inf(a, b, c)$, the proof is analogous. □

Proposition 3. *Let $\mathcal{B} \models$ ASL with universe A. Define $\mathcal{A}$ as the L_{Ord}-structure with the same universe, A, and such that the interpretation of $\leq^{\mathcal{A}}$ is the set $\{(a, b) \in A^2 \ : \ \mathcal{B} \models a \sqcup^{\mathcal{B}} b = b\}$. Then, $\mathcal{A}$ is a model for* JL. *If the interpretation of $\leq$ is instead characterized by $a \sqcup^{\mathcal{B}} b = a$, then $\mathcal{A}$ is a model for* ML.

Proof. Since $\mathcal{A}$ and $\mathcal{B}$ share the universe, $=^{\mathcal{A}}$ coincides with $=^{\mathcal{B}}$, and such we will omit the superscript throughout the proof. Recall that $\mathcal{A} \models$ JL if $\mathcal{A} \models F$ for all $F \in$ JL. $\mathcal{A} \models$ **O1** if $(a, a) \in \leq^{\mathcal{A}}$ for all $a \in A$. From the definition of $\leq^{\mathcal{A}}$, this is equivalent to proving that $\mathcal{B} \models a \sqcup^{\mathcal{B}} a = a$, which we have from the fact that $\mathcal{B} \models$ ASL and **L1** $\in$ ASL. For **O2** we need to show that if $(a, b) \in \leq^{\mathcal{A}}$ and $(b, c) \in \leq^{\mathcal{A}}$, then $(a, c) \in \leq^{\mathcal{A}}$, which is again equivalent to showing that if $\mathcal{B} \models a \sqcup^{\mathcal{B}} b = b$ and $\mathcal{B} \models b \sqcup^{\mathcal{B}} c = c$, then $\mathcal{B} \models a \sqcup^{\mathcal{B}} c = c$. This is obtained from $\mathcal{B} \models$ **L5** as $a \sqcup^{\mathcal{B}} c = a \sqcup^{\mathcal{B}} (b \sqcup^{\mathcal{B}} c) = (a \sqcup^{\mathcal{B}} b) \sqcup^{\mathcal{B}} c = b \sqcup^{\mathcal{B}} c = c$. Similarly, for **O3** we have to show that if $\mathcal{B} \models a \sqcup^{\mathcal{B}} b = b$ and $\mathcal{B} \models b \sqcup^{\mathcal{B}} a = a$, then $a = b$; this is directly obtained from $\mathcal{B} \models$ **L3**, since it gives $a = b \sqcup^{\mathcal{B}} a = a \sqcup^{\mathcal{B}} b = b$. Finally, for **O4**, we need to find, for all $a, b \in A$, an element $c \in A$ such that $\mathcal{B} \models a \sqcup^{\mathcal{B}} c = c$, $\mathcal{B} \models b \sqcup^{\mathcal{B}} c = c$, and if $\mathcal{B} \models a \sqcup^{\mathcal{B}} k = k$ and $\mathcal{B} \models b \sqcup^{\mathcal{B}} k = k$, then $\mathcal{B} \models c \sqcup^{\mathcal{B}} k$, for all $k \in A$. Let $c = a \sqcup^{\mathcal{B}} b$. Using $\mathcal{B} \models$ **L1**$, \ldots, \mathcal{B} \models$ **L5**, we can operate $b \sqcup^{\mathcal{B}} (a \sqcup^{\mathcal{B}} b = b \sqcup^{\mathcal{B}} (b \sqcup^{\mathcal{B}} a)) = (b \sqcup^{\mathcal{B}} b) \sqcup^{\mathcal{B}} a = b \sqcup^{\mathcal{B}} a = a \sqcup^{\mathcal{B}} b = c$, thus getting $\mathcal{B} \models b \sqcup^{\mathcal{B}} c = c$. We obtain $\mathcal{B} \models a \sqcup^{\mathcal{B}} c = c$ similarly. Now, assume that $\mathcal{B} \models a \sqcup^{\mathcal{B}} k = k$ and $\mathcal{B} \models b \sqcup^{\mathcal{B}} k = k$; then $c \sqcup^{\mathcal{B}} k = (a \sqcup^{\mathcal{B}} b) \sqcup^{\mathcal{B}} k = a \sqcup^{\mathcal{B}} (b \sqcup^{\mathcal{B}} k) = a \sqcup^{\mathcal{B}} k = k$, thus we have $\mathcal{B} \models c \sqcup^{\mathcal{B}} k = k$, as we wanted to prove.

When $\leq^{\mathcal{A}} = \{(a, b) \in A^2 \ : \ \mathcal{B} \models a \sqcup^{\mathcal{B}} b = a\}$, $\mathcal{A}$ is shown to be a model for ML in an analogous way. □

Corollary 1. *Let $\mathcal{A} \models$ JL with universe A, and let $\mathcal{B}$ be the L_{Irr}-expansion of $\mathcal{A}$ such that $\sqcup^{\mathcal{B}}$ is defined as in Proposition 2. Then, $\mathcal{B}$ is a model for the L_{Irr}-theory* JL $\cup$ ASL. *Conversely, if $\mathcal{B} \models$ ASL with universe A, and $\mathcal{A}$ is the L_{Irr}-expansion of $\mathcal{B}$ with $\leq^{\mathcal{A}}$ defined as in Proposition 3, then $\mathcal{A}$ is a model for the L_{Irr}-theory* JL $\cup$ ASL. *These results hold analogously for the theory* ML.

Corollary 2. *Let $\mathcal{A} \models$ JL $\cup$ ASL with universe A. Then for all $a, b \in A$, $\sqcup^{\mathcal{A}}(a, b) = c$ is the unique element in A satisfying $\mathcal{A} \models \sup(a, b, c)$. Conversely, if $\mathcal{A} \models$ ML $\cup$ ASL, then $\sqcap^{\mathcal{A}}(a, b) = c$ is the unique element satisfying $\mathcal{A} \models \inf(a, b, c)$.*

Corollary 3. *The following bijective equivalences hold, via the mappings from Corollaries 1 and 2:* Mod(JL)$\cup$Mod(ML) $\equiv$ Mod(ASL), Mod(JL$\cup$ASL) $\equiv$ Mod(JL), Mod(ML$\cup$ASL) $\equiv$ Mod(ML) *and* Mod(OL) $\equiv$ Mod(AL).

References

F. G. Cozman and P. Walley. Graphoid properties of epistemic irrelevance and independence. *Annals of Mathematics and Artificial Intelligence*, 45(1):173–195, 2005.

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

A. P. Dawid. Separoids: A mathematical framework for conditional independence and irrelevance. *Annals of Mathematics and Artificial Intelligence*, 32(1):335–372, 2001.

E. Grädel and J. Väänänen. Dependence and independence. *Studia Logica*, 101(2):399–410, 2013.

G. Grätzer. *General Lattice Theory*. Birkhäuser Basel, 2003.

M. Hannula, J. Kontinen, and S. Link. On the finite and general implication problems of independence atoms and keys. *Journal of Computer and System Sciences*, 82(5):856–877, 2016.

D. Marker. *Model Theory: An Introduction*. Springer, 2000.

J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, 1988.

J. Pearl and A. Paz. Graphoids: A graph-based logic for reasoning about relevance relations. In *Advances in Artificial Intelligence*, volume 2, pages 357–363. Elsevier, 1987.