Dependency Structures and Locales

Category Theory Octoberfest, JHU, 26 October 2019

Gershom Bazerman

Awake Security

jww Raymond Puzio, Albert Einstein Institute

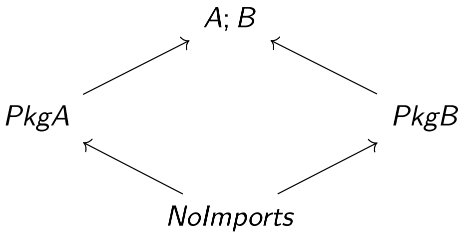
Outline

1. Dependency Structures
2. Distributive Lattices
3. Locales
4. Versioning
5. Free Distributive Lattices
6. Dependency Problems
7. Conclusion

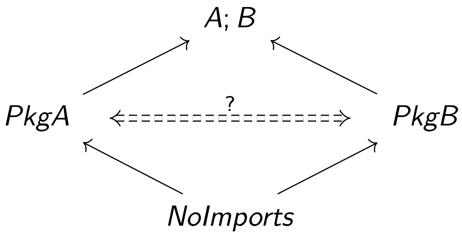
Disclaimer

Everything in sight is assumed to be finite (for now).

Motivation



Motivation



Dependency data is everywhere

- Package repositories
- Concurrent semantics (Petri nets, CCS, CSP, π-calculus, etc.)
- Knowledge representation (proof dependencies, course dependencies, chapter dependencies)

Two key questions

Two key questions: *compositionality* (external, gros), and *reachability* (internal, petit).

The plan of attack

Internal languages the "data first" way.

1. Raw dependency data (in the wild)
2. Familiar mathematical structure
3. Nice class of categories
4. "Read off" the internal logic
5. Apply combinatorics
6. Apply algebraic topology

One end goal: a (non-dependent) type theory *with* homological data. (*Not* a "homotopy type theory".)

Existing Order-Theoretic Models

General Event Structures (Neilsen, Plotkin, Winskel)

Model dependency, conflict, choice. Hard to reason about!

Event Structures (Winskel)

Model dependency, conflict, **not** choice. Good properties, form a domain! Correspond to safe Petri nets, CCS.

pomsets (partially ordered multisets) (Pratt)

Model dependency, **not** conflict, **not** choice. Compose beautifully. Relate to Kleene algebras.

Our Approach

Dependency Structures with Choice (B., Puzio)

Model dependency, **not** conflict, choice. Nice properties. Relate to locales and constructive logic. Haven't studied composition.

Dependency Structures with Choice and Conflict (B., Puzio)

Model dependency, conflict, choice. Future work! Should allow us to relate GES to Directed Topology.

Pre-DSCs

Definition

A **Pre-Dependency Structure with Choice** is a pair $(E, D : E \to \mathcal{P}(\mathcal{P}(E)))$

- E is thought of as a finite set of events.
- D is thought of as mapping each event to a set of alternative dependency requirements – i.e. to a predicate in disjunctive normal form.

DSCs

Definition

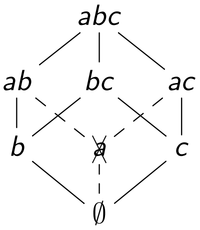
A **Dependency Structure with Choice** (DSC) is a pre-DSC with D satisfying appropriate conditions of transitive closure and cycle-freeness.

X is a **possible dependency set** of e if $X \in D(e)$. An event set X is a **complete event sent** if for every element e there is a possible dependency set Y of e such that $Y \subseteq X$. A pre-DSC is a DSC if every possible dependency set of every element is complete and cycle-free.

reachable dependency posets

A dependency structure has an associated reachable dependency poset (an "unwinding" or "configuration family") which is a subposet of $\mathcal{P}(E)$ generated by possible dependency sets augmented by their "parent" and ordered by inclusion. A rdp (when there is no conflict data) has all joins, and is bounded, so therefore is a lattice.

a depends on b or c



A Sub-Goal

A rdp is *almost* the frame of opens of a topological space. Our aim is to complete it into one so that we can analyze dependency structures by topological means.

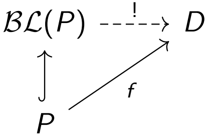
Definitions

- A **Lattice** is a poset with all finite meets (greatest lower bounds) and joins (least upper bounds).
- A **Distributive Lattice** is a lattice such that $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$.
- finite distributive lattices are one to one with finite frames (finite meets, arbitrary joins, distribution), and hence finite sober spaces.
- $\mathcal{J}(P)$ is the subposet of the join-irreducible elements (including nullary joins) of P.
- $\mathcal{O}(P)$ is the distributive lattice generated by the downsets of P under inclusion.

BL for posets, finite case

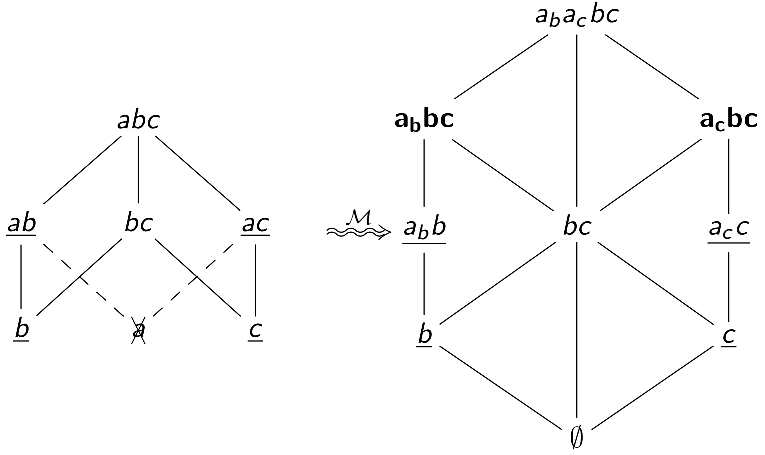
Theorem

(B., Puzio, after MacNeille) For a finite poset, the injective hull may be constructed as $\mathcal{O}(\mathcal{J}(P))$, with an injection that sends join-irreducible elements to their downsets, and composite elements to the union of their join-irreducible basis. Furthermore,



Corollary: (Finite) distributive lattices are a reflective subcategory of (finite) posets with "distributive" morphisms.

Extended BL is idempotent, and meaningful



Definitions

A **Heyting algebra** is a lattice with a unique top and bottom element, and a special "implication" operation called the **relative pseudo-complement** ($a \to b$) which yields the unique greatest element x such that $a \wedge x \leq b$. A **complete Heyting algebra** is a Heyting algebra such that it is also a complete lattice. Finite distributive lattices are one-to-one with finite Heyting algebras.
A **frame** is a complete Heyting algebra, considered in a category where morphisms preserve finite meets and arbitrary joins. A **locale** is a frame, but in a category with morphisms reversed.
$FinLoc = FinFrm^{op} = FinDLat^{op}$

A nucleus:

- endofunction
- preserves meets
- idempotent
- contractive (monotone)

A nucleus on a frame yields a frame surjection to the quotient frame of fixpoints.

Locales let us relate logics and spaces. Nucleii let us relate modal logics (with a “possibility” operator) and topologies.

The BL-topology

Lemma

There exists a **Bruns-Lakser topology** *on a finite locale, which is the least nucleus with* $\mathcal{J}(\mathcal{J}(L)$ *as fixedpoints.*

Lattice presentations of DSCs

Lemma

DSCs are up-to-renaming equivalent to finite lattices. One one side this is the rdp or "unwinding" construction. On the other side, join-irreducibles constitute atoms, as quotiented by unique binary joins.

Localic presentations of DSCs

Theorem

The set of DSCs is one-to-one with the set of finite locales generated by lattice posites, equipped with the Bruns-Lakser topology, i.e. dependency structures can be understood as a particular presentation of a class of locales.

Questions

- What morphisms do we need to define to extend this to an equivalence of categories?
- Where is the relationship between a nucleus of a locale and a coverage on its generating posite recorded?
- What is the correct characterization (logical, topological) of a distributive lattice generated by a lattice of join-irreducibles?

Definition

A **version parameterization** of a DSC is an idempotent endofunction on events, from lower to higher, where for every possible dependency set of every event, there is another possible dependency set of that event where the lower versions have been substituted for higher versions. Idempotency here translates into the condition that no higher event is itself a lower event of something else.

A **poset version parameterization** is the induced endofunction on a reachable dependency poset generated by a version parameterization on a DSC

Lemma

Every poset version parameterization is a nucleus that in addition preserves existing joins.

Theorem

A **ponucleus** *is an idempotent monotone endofunction on a poset which preserves existing finite meets. Given a join-preserving ponucleus j, on a finite poset P, $\mathcal{BL}(j)$ is a nucleus on $\mathcal{BL}(P)$*

Corollary: given a DSC (E, D) and a version parameterization with an induced poset version parameterization p, then $\mathcal{BL}(p)$ is a nucleus on $\mathcal{BL}(E, D)$.

Motivation

Internal modal logic of the locale of a DSC – i.e., $\mathcal{BL}(E, D)$

- Atoms = "events with traces"
- Objects = traced event sets
- $\Diamond$ = "round the version up"
- & = intersection
- $||$ = union

Not the logic we want! (It is a logic "of" states rather than "about" states).

- A set S is *redundant* with regards to T if $T \subset S$.
- The irredundancy quotient is a nucleus, and the double powerset lattice, quotiented by irredundancy, is the free distributive lattice.
- This gives positive formulae in disjunctive normal form, quotiented by all laws of first order intuitionistic logic.
- This is equivalent to the lattice $\mathcal{U}(\mathcal{O}(S))$ (upsets of downsets).
- This construction extends from sets to posets.
- (Lemma. B., Puzio) This construction sends a ponucleus to a nucleus that in addition preserves joins.

The Free Distributive Lattice of a DSC

The internal modal logic of the free locale of a DSC – i.e. $\mathcal{UM}(E, D) = \mathcal{U}(\mathcal{O}(\mathcal{J}(rdp(DSC))))$

- Atoms = "events with traces"
- Objects = **Sets of** traced event sets (as formulae in dnf)
- $\Diamond$ = "round the version up"
- & = actual logical conjunction
- || = actual logical disjunction

This **is** the logic you're looking for.

Definition

A dependency problem in a DSC (E, D) is the pair of a formula ϕ in $\mathcal{UM}(E, D)$ and a monotone increasing (i.e. growing as further elements are added to the source set) objective function of type $\mathcal{P}(E) \to \mathbb{R}$. A solution to such a problem is an event set which satisfies the formula and minimizes the objective function.

Observation: Detecting *compatible* event sets can be formulated as a dependency problem.

The cost of solving a dependency problem over a poset is bounded by the maximum number of disjuncts in a normalized dnf formula over that poset. This is the same as the maximal antichain in the downsets of that poset – i.e. the **width** ($\mathfrak{w}$).

Theorem

(Sperner, 1928): *The powerset of a set with n elements, under inclusion ordering, has a width of* $\binom{n}{\lceil n/2 \rceil}$.

Lemma

Stanley's Width Lemma (2019): *The width of a poset constructed as the product of n chains, each of length l_n is the maximal coefficient of* $(1 + x...x^{l_1}) * (1 + x...x^{l_2})... * (1 + x...x^{l_n})$

Theorem

(B., Puzio): Define $\mathfrak{m}(a, b)$ *as the central (maximal) coefficient of the formal polynomial expansion of* $(1 + x + x^2... + x^a)^b$. *Define* $\mathfrak{h}(P)$ *as the height of a poset, i.e. the length of its longest chain. Given any poset* P*, then:*

$$\mathfrak{w}(\mathcal{O}(P)) \leq \mathfrak{m}(2, \mathfrak{w}(P)) * \mathfrak{m}(\mathfrak{h}(P), \lceil \mathfrak{w}(P)/2 \rceil)$$

Future Work

Mathematical:

- Develop a homology theory over conflicts using this framework. It should look like discrete morse theory.
- Relate this construction to directed topology.

Applied:

- Characterize the difficulty of problems for SAT solvers (and possibly give a general construction of "eager" SMT solving).
- Develop a modal language of dependencies for software packaging (Haskell, Nix, etc.).
- Provide a sound basis for the organization of data in Hales' Formal Abstracts project.

Stray weird questions:

- Is Bruns-Lakser a cubulation functor?
- Completions give reflections give factorizations. Do they yield a model structure on *Pos*?

References I

- Carlo Angiuli, Edward Morehouse, Daniel R Licata, and Robert Harper, *Homotopical patch theory*, ACM SIGPLAN Notices, vol. 49, ACM, 2014, pp. 243–256.
- Günter Bruns and Harry Lakser, *Injective hulls of semilattices*, Canadian Mathematical Bulletin **13** (1970), no. 1, 115–118.
- Richard N Ball, Aleš Pultr, and Joanne Walters Wayland, *The dedekind macneille site completion of a meet semilattice*, Algebra universalis **76** (2016), no. 2, 183–197.
- Lisbeth Fajstrup, Eric Goubault, Emmanuel Haucourt, Samuel Mimram, and Martin Raussen, *Directed algebraic topology and concurrency*, vol. 138, Springer, 2016.
- Robin Forman, *A user's guide to discrete morse theory*, Sém. Lothar. Combin **48** (2002), 35pp.

References II

- Mai Gehrke and Samuel J Van Gool, *Distributive envelopes and topological duality for lattices via canonical extensions*, Order **31** (2014), no. 3, 435–461.
- Peter T Johnstone, *Stone spaces*, vol. 3, Cambridge university press, 1982.
- Holbrook Mann MacNeille, *Partially ordered sets*, Transactions of the American Mathematical Society **42** (1937), no. 3, 416–460.
- Samuel Mimram and Cinzia Di Giusto, *A categorical theory of patches*, Electronic notes in theoretical computer science **298** (2013), 283–307.

References III

- Mogens Nielsen, Gordon Plotkin, and Glynn Winskel, *Petri nets, event structures and domains, part i*, Theoretical Computer Science **13** (1981), no. 1, 85–108.
- Emanuel Sperner, *Ein satz über untermengen einer endlichen menge*, Mathematische Zeitschrift **27** (1928), no. 1, 544–548.
- Richard P Stanley, *Algebraic combinatorics*, Springer **20** (2013), 22.
- Steven Vickers, *Topology via logic*, Cambridge University Press, 1996.