

A lattice-theoretic approach for representing and managing hypothesis-driven research

Bernardo Gonçalves, Fabio Porto

Extreme Data Lab (DEXL Lab)
National Laboratory for Scientific Computing (LNCC),
Av. Getulio Vargas 333, Petrópolis, Brazil
{bgonc, fporto}@lncc.br

Abstract. As the problems of interest in large-scale science are ever more complex or intertwined, scientists can benefit from machinery to manage hypotheses and their interconnection in the context of a large-scale research project. A partial order is assumed to model hypothesis interconnectivity, meaning that hypotheses may be ‘based on or equal to’ one another. Then we investigate lattices as a mathematical abstraction for representing and managing hypothesis-driven research. We present queries designed to exploit lattice-theoretic topological features of a hypothesis-driven research and then discuss their complexity.

Keywords: Scientific Hypotheses, Large-scale Science, Data Management, Lattice Theory.

1 Introduction

Large-scale science relies on data management technology in order to deal with the increasing amount of observational and experimentally-produced data. As the problems of interest are ever more complex or intertwined, scientists can also benefit from machinery to manage hypotheses in the context of a large-scale research project. Hypotheses appear *interconnected* in a research, playing roles of assumption or proposition to explain phenomena [11]. We say that a research is *hypothesis-driven* if hypotheses are central elements for its conceptual organization and key to its fulfillment.

Problem Statement. *Representation and management of scientific hypotheses within the context of a large-scale research project.*

We address this problem by considering that hypothesis interconnectivity is order-theoretic in nature. We write $h_\alpha \leq h_\beta$, read hypothesis h_α ‘is based on or equal to’ hypothesis h_β . We aim at enabling a scientist end-user to query topological features of a research, such as hypothesis connectivity, paths and neighborhoods. There are, nonetheless, some special topological features that seem to be of particular interest in a hypothesis-driven research: given any two hypotheses in a research, (i) what is their *infimum* hypothesis (“weakest stronger claim”), and (ii) what is their *supremum* hypothesis (“strongest weaker claim”). Both are warranted to exist for all h_α, h_β in a partially-ordered set \mathcal{R} , if \mathcal{R} is a *lattice* [6,10]. In this paper we investigate lattices as a mathematical abstraction for representing and managing hypothesis-driven research.

Section 2 introduces notation and concepts from Lattice Theory (Order) for normalization and easy reference. Section 3 introduces the (hypothesis-driven) research lattice abstraction. Section 4 considers querying features on a research lattice. Section 5 briefly discusses related work on hypothesis data modeling. Section 6 concludes the paper.

2 Lattice Theory: Notation and Basic Concepts

Order-theoretic properties of sets can be expressed in terms of an ordering relation \leq . Let A be an arbitrary set. Then, for all $a, b, c \in A$, we consider:

- | | | |
|------|---|-----------------|
| (P1) | $a \leq a$ | (reflexivity) |
| (P2) | $a \leq b$ and $b \leq a$ implies that $a = b$ | (antisymmetry) |
| (P3) | $a \leq b$ and $b \leq c$ implies that $a \leq c$ | (transitivity). |

A *partially ordered set*, or *poset* $\langle A; r \rangle$, \mathcal{A} for short, consists of a non-empty set A and a binary relation r on A , such that r satisfies properties (P1)–(P3). Note that $a \geq b$ means $b \leq a$, i.e., \leq has an inverse; and that $<$ is defined by $\neg(\text{P1})$ and (P3). If \mathcal{A} is a poset, $a, b \in A$, then a and b are *comparable* if $a \leq b$ or $b \leq a$. Otherwise, a and b are *incomparable*, in notation, $a \parallel b$. A *zero* of a poset \mathcal{P} is an element 0 with $0 \leq x$ for all $x \in P$. A *unit*, 1 , satisfies $x \leq 1$ for all $x \in P$. There are at most one zero and at most one unit in a poset. A *bounded* poset is one that has both 0 and 1 . A poset \mathcal{P} can be represented by a list of the pairs of elements satisfying the partial order. It can also be represented by means of the covering relation: a is *covered by* b , in notation, $a \prec b$, if $a < b$ and for no x , $a < x < b$. This relation is enough to determine a finite poset [10], with $\langle P; \prec \rangle$ for \mathcal{P} , and it is used to build its Hasse diagram.¹

Given two hypotheses $h_\alpha, h_\beta \in R$, we shall write: $h_\alpha \leq h_\beta$ (h_α ‘is based on or equal to’ h_β); $h_\alpha < h_\beta$ (h_α ‘is based on’ h_β); $h_\alpha \prec h_\beta$ (h_α ‘is directly based on’ h_β); and $h_\alpha \parallel h_\beta$ (h_α ‘is incomparable to’ h_β).

Let $H \subseteq P$, $a \in P$, for an arbitrary poset $\langle P; \leq \rangle$. Then a is an *upper bound* of H , if $h \leq a$ for all $h \in H$. An upper bound a of H is the *least upper bound* of H or *supremum* of H if, for any upper bound b of H , we have $a \leq b$. It is written $a = \sup H$, and the uniqueness of the supremum can be shown straightforwardly.² The concepts of *lower bound* and *greatest lower bound* or *infimum* are similarly defined. The latter is denoted by $\inf H$, and uniqueness of the infimum is verified likewise. The set of upper (lower) bounds of an element h is denoted $\uparrow h$ ($\downarrow h$).

Def. 1. A poset \mathcal{L} is a **lattice** if $\sup\{a, b\}$ and $\inf\{a, b\}$ exist for all $a, b \in L$.

We can refer to lattices as algebraic structures by the notation:

$$\begin{aligned} a \vee b &\equiv \sup\{a, b\} \\ a \wedge b &\equiv \inf\{a, b\} \end{aligned}$$

read \vee the **join**, and \wedge the **meet**. In lattices, these are binary operations, i.e., they can be applied to any pair of elements $a, b \in L$ to yield again an element of L . For all $a, b \in L$, these operations satisfy idempotency, commutativity, associativity and absorption. Lattice Theory has singled out a special kind of posets for detailed investigation. It is a vast mathematical theory with plenty of interesting results. In Theoretical Computer Science, for instance, lattices are used in the study of semantic properties of programs (e.g., confluence, termination) [6]. In the sequel we refer to lattices as a mathematical abstraction for hypothesis-driven research.

The *inf-sup* existence property of lattices as special posets may translate into neat properties of a hypothesis-driven research as a scientific theory.

3 The Research Lattice Abstraction

We characterize a hypothesis-driven research as a poset, further constrained to be a lattice.

Def. 2. Let H be a set of hypotheses. The **top**, \top , is a special hypothesis such that $h \leq \top$, for all $h \in H$. The **bottom**, \perp , is another special hypothesis, dual of \top , such that $\perp \leq h$, for all $h \in H$.

Def. 3. A hypothesis-driven **research poset** \mathcal{R} is a finite, bounded poset, whose elements are hypotheses, $\top \in R$ is unit, and partial order \leq reads ‘is based on or equal to’.³

Def. 4. Let \mathcal{R} be a research poset. We say that \mathcal{R} is **lifted**, written \mathcal{R}_\perp , if $\perp \in R$. Then \perp is zero of \mathcal{R}_\perp .

¹ http://en.wikipedia.org/wiki/Hasse_diagram.

² In fact, if a_0 and a_1 are both suprema of H , then $a_0 \leq a_1$, since a_1 is an upper bound, and a_0 is a supremum. Similarly, $a_1 \leq a_0$, since a_0 is an upper bound, and a_1 is a supremum. Thus, by (P2), $a_0 = a_1$. \square

³ N.B.: since \mathcal{R} is a finite poset, we can infer the set of partial order pairs \leq from the set of covering pairs \prec . At a logical level, we shall refer to both somewhat interchangeably. At the user level, only \prec is seen.

The top hypothesis, \top , serves as a root element under which a research poset is built. It assumes nothing, leaving everything open for consideration. The bottom hypothesis, \perp , is meant to indicate that the research is yet unfinished, either because it is simply ongoing (without a final hypothesis), or because there are two or more incomparable hypotheses that can be considered to be competing.

If there exists a hypothesis $h_\nu \neq \perp$ that is directly based on two others, i.e., $h_\nu \prec h_\alpha$ and $h_\nu \prec h_\beta$, then it is intuitive to say that h_α and h_β are confluent to h_ν as trains of thought. As a stronger notion, consider a mental operation in which a scientist/practitioner takes two (or more) hypotheses h_α and h_β as input and (non-deterministically) generates another hypothesis h_ν as output. Def. 5 gives a precise meaning for that conceptual operation, which is founded in Conceptual Blending [7], an established theory in cognitive sciences.

Def. 5. Let \mathcal{R} be a research poset, and let $S \subset R$ such that $|S| > 1$ and $x \parallel y$ for all $x, y \in S$. Then $h \in R$ is a **blend** of S , written $h \triangleleft S$, if $h \prec s$ for all $s \in S$ and $h \prec s'$ implies $s' \in S$.

Examples of research posets are presented in Fig. 1. They have been derived from authoritative sources, i.e., taken from their (historical) research context and organized according to our conceptual framework—see Appendix B; note, in contrast to Fig. 1, that hypotheses lack their inherent interconnectivity when presented in tabular form. Those research posets are all lattices. Nonetheless, we need to investigate research posets in general w.r.t. the *inf-sup* existence property. In a poset \mathcal{R} , the greatest lower bound $\inf\{x, y\}$ of whatever pair $\{x, y\} \subset R$ may fail to exist for two different reasons [6]: (i) because x and y have no common lower bound, i.e., $\{x, y\}^\ell = \emptyset$; or (ii) because they have no *greatest* lower bound, i.e., $|\text{Max}\{x, y\}^\ell| > 1$. The rationale for $\sup\{x, y\}$ non-existence is dually the same.

Lemma 1. Let \mathcal{R} be a research poset. Then $\{x, y\}^\ell \neq \emptyset$ and $\{x, y\}^u \neq \emptyset$ for all $x, y \in R$.

Proof. By Def. 3, \mathcal{R} is bounded. Then it has both 0 and 1. Since it has 0, and $0 \leq x$ for all $x \in R$, we have $0 \in \{x, y\}^\ell$. Thus $\{x, y\}^\ell \neq \emptyset$. Similarly, since it has 1 and $x \leq 1$ for all $x \in R$, we have $1 \in \{x, y\}^u$. Thus $\{x, y\}^u \neq \emptyset$. \square

By Lemma 1, we know that condition (i) above is avoided. We still need to ensure that condition (ii) is avoided as well. For that, we shall constrain our application of Conceptual Blending specifically for scientific thinking (Axiom 1). This axiom can be interpreted as a parsimony principle.

Axiom 1. Let \mathcal{R} be a research poset, and let $h \triangleleft S$ such that $\{x, y\} \subseteq S$, for $h \in R$ and $S \subset R$. Then there can be no $h' \neq h$ in R such that $h' \triangleleft S'$ and $\{x, y\} \subseteq S'$.

Theorem 1. Let $\mathcal{R} = \langle R; \prec \rangle$ be a research poset, initialized $\prec := \{(\perp, \top)\}$, and let it be freely manipulated by Algorithms 1, 2 or 3 (cf. Appendix A) to derive an arbitrary poset \mathcal{R}' . We claim that \mathcal{R}' is a lattice.

Proof. The proof is given in Appendix A. \square

4 Querying over a Research Lattice

We can treat querying features related to $\inf\{a, b\}$ and $\sup\{a, b\}$ from an algebraic perspective through the operators meet and join (cf. Section 2). They are among the main querying capabilities we have in mind in view of a scientist end-user thinking over his/her research lattice. Let us consider some query examples (see Fig. 1).

- (Q1). Given (h_2) *Polynucleotide structure* and (h_3) *DNA storage of genetic information*, find their join (“strongest weaker claim”): $\{h \in R \mid h = h_2 \vee h_3\}$. $[h_1]$.
- (Q2). Given (h_2) *Intra-species competition* and (h_5) *Predator-prey coupling*, find their meet (“weakest stronger claim”): $\{h \in R \mid h = h_2 \wedge h_5\}$. $[\perp]$.
- (Q3). Given (h_3) *Logistic growth* and (h_5) *Predator-prey coupling*, find the disjoint union of their complementary upper bounds: $(\uparrow h_3 \setminus \uparrow h_5) \cup^* (\uparrow h_5 \setminus \uparrow h_3)$. $[\{(h_3, 1), (h_2, 1), (h_5, 2), (h_4, 2)\}]$.

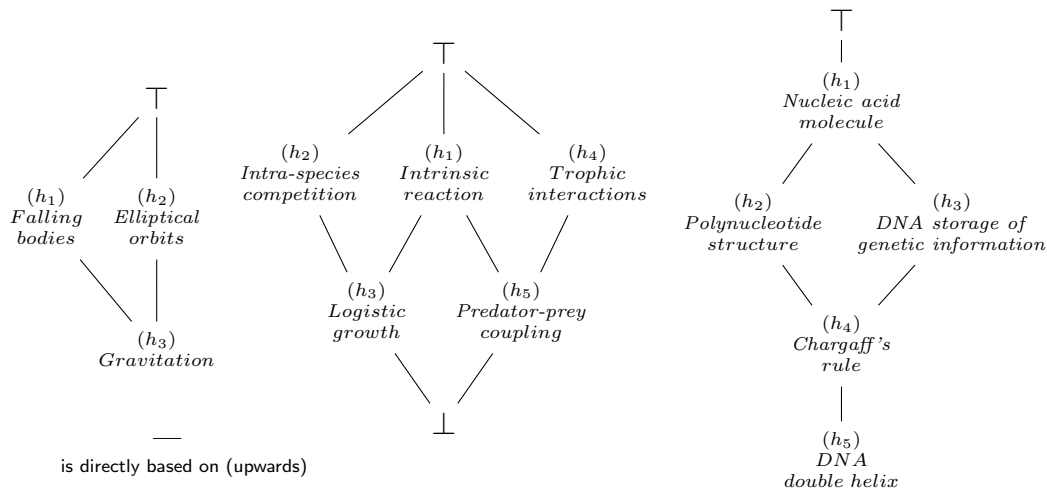


Fig. 1. Hasse diagrams of research posets (lattices) from diverse fields of science. Hypotheses (the elements) are connected by line segments (the covering relation \prec) that go upward from h_α to h_β whenever h_α 'is directly based on' h_β . The 'ID' and 'name' properties for each element are presented (cf. Appendix B).

These queries illustrate a potential interest (Q1) in the strongest claim once dropping h_2 and h_3 ; or (Q2) in the weakest claim once assuming h_2 and h_5 ; or (Q3) in clustering sets of competing hypotheses (one set competing with the other) for inspection. Retrieving the up-set $\uparrow h$ of a hypothesis (reflexive transitive closure of \leq) in a research lattice can also be relevant. For brevity, we are presenting small examples, but the research lattice abstraction is meant to be used in large-scale science. Furthermore, we can think of research lattices being further merged on the web (*merge* as an additional operation), such that large lattices are to be formed.

The complexity of queries related to the *meet* and *join* operators (Q1-Q2) or to computing reflexive transitive closures (Q3) is affordable. Efficient methods for the two lattice operations that also involve computing the closures are available in the literature [1,15,8]. They have been motivated by theoretical concerns [15], multiple inheritance in programming language design [1], or querying over trees and directed acyclic graphs (DAGs) [8]. Taking advantage of pre-processing on specific data structures, they can be computed in almost constant time [1]. Besides, a connectivity-based *search* procedure for research lattices does not have to differ from one for DAGs. Depth-first search (DFS) [13], for example, is a classical graph-traversal algorithm that has a linear complexity on the size of its input, $O(|V| + |E|)$, where V stand for vertices and E for edges. The adaptation of DFS for research lattices (seen as special DAGs to be traversed) must be straightforward. A systematic study of the complexity of operations and queries on a research lattice data structure shall be object of future work.

5 Related Work

Most of the current work on hypothesis modeling is found embodied in the RDF data model.⁴ We point out the initiatives on the Robot Scientist [12], the SWAN Knowledge Base [9], and the Hypothesis Browser (or HyQue) [5]. The choice for RDF (with OWL on top), however, is presumably more related to its reasoning capabilities and web compliance, and less because the RDF triple structure is an accurate fit for the data (the hypotheses). Graph data models [3], in general, are in fact candidates to frame hypothesis-driven research. In this sense, the RDF query language still lacks support for topological features (connectivity, neighborhoods, etc) as standard abstractions [2]. This paper contributes to the investigation of a theoretical basis for managing hypothesis-driven research (Problem 1). Seen as graphs, lattices would have the special features of: (i) being hierarchical (DAGs) and bounded; (ii) having their edge labels defined by a notion of order (the covering relation); and (iii) the *inf-sup* existence property for each pair of nodes.

This work can be complementary to initiatives on the (semi-)automatic retrieval of hypotheses (or claims) from the narrative structure of scientific documents. Primarily, we aim at providing scientists with a data management framework to *operate* their large-scale hypothesis-driven research.

⁴ <http://www.w3.org/TR/rdf-primer/#rdfmodel>.

6 Conclusions

In this paper we introduced a lattice-theoretic approach for managing hypothesis-driven research. To our knowledge, this is the first study on scientific hypotheses from a data management point of view. In this approach, scientists themselves are enabled with operations to insert/delete hypotheses into/from their research lattices and query them further. This research vision can be made into useful data management technology for hypothesis-driven large-scale science.

References

1. H. Ait-Kaci, R. Boyer, P. Lincoln, and R. Nasr. Efficient implementation of lattice operations. *ACM T Progr Lang Sys*, 11(1):115–46, 1989.
2. R. Angles and C. Gutierrez. Querying RDF data from a graph database perspective. In *Proc. of ESWC*, 2005.
3. R. Angles and C. Gutierrez. Survey of graph database models. *ACM Computing Surveys*, 40(1):1–39, 2008.
4. N. Bacaer. *A Short History of Mathematical Population Dynamics*. Springer, 2011.
5. A. Callahan, M. Dumontier, and N. H. Shah. HyQue: Evaluating hypotheses using semantic web technologies. *Journal of Biomedical Semantics*, 2(Suppl 2):S3, 2011.
6. B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, 2nd edition, 2002.
7. G. Fauconnier and M. Turner. *The way we think: Conceptual blending and the mind’s hidden complexities*. Basic Books, 2002.
8. J. Fischer and D. H. Huson. New common ancestor problems in trees and directed acyclic graphs. *Information Processing Letters*, 110(8-9):331–5, 2010.
9. Y. Gao et al. SWAN: A distributed knowledge infrastructure for alzheimer disease research. *J. Web Semantics*, 4(3):222–8, 2006.
10. G. A. Gratzler. *General Lattice Theory*. Birkhauser, 2nd edition, 1998.
11. N. R. Hanson. *Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science*. Cambridge University Press, 1958.
12. R. D. King et al. The automation of science. *Science*, 324(5923):85–9, 2009.
13. J. Kleinberg and E. Tardos. *Algorithm Design*. Addison Wesley, 2005.
14. L. A. Pray. Discovery of DNA structure and function: Watson and Crick. *Nature Education*, 1(1):1–8, 2008.
15. M. Talamo and P. Vocca. A data structure for lattice representation. *Theor. Comput. Sci.*, 175(2):373–92, 1997.

Appendix A: Proof of Theorem 1

Proof. The proof roadmap is as follows. We refer to the insertion and deletion operations to accomplish it by induction. So, consider research poset $R_{\perp} = \langle R; \prec \rangle$, with $\prec = \{(\perp, \top)\}$, as given. We have $\inf\{\top, \perp\} = \perp$ and $\sup\{\top, \perp\} = \top$, therefore \mathcal{R}_{\perp} is a lattice. Now, assume by the induction hypothesis that \mathcal{R} of size $|R| = k$ elements, $k \geq 2$, is a lattice, either lifted or not.⁵ We must show that $\mathcal{R}' := \text{insert}(\mathcal{R}, h, ?)$, of size $|R'| = k + 1$, is also a lattice, either after simple insertion, $? = s$, or blending insertion, $? = S$ of a new element h . Furthermore, suppose the insertion operations have been shown to preserve the property such that R can be arbitrarily large, say $|R| = M$. Then, by reverse induction, assume that \mathcal{R} of size $|R| = k$ elements, $2 < k \leq M$, is a lattice, either lifted or not. We must show that $\mathcal{R}' := \text{delete}(\mathcal{R}, h)$, of size $|R'| = k - 1$ is a lattice idem. Before we proceed with the induction argument for each operation, it is worth highlighting some general features of posets. If $a \leq b$, then $\inf\{a, b\} = a$ and $\sup\{a, b\} = b$; hence, to show that a poset is a lattice, it suffices to check the incomparable pairs, i.e., $\{a, b\}$ such that $a \parallel b$ [6]. Consider also Def. 3, Lemma 1, to recall that after manipulation the poset must be kept bounded. In particular, after insertion of a new element h , we have to verify for all new pairs $\{h, x\} \subset R' \setminus R$ that both $\inf\{h, x\}$ and $\sup\{h, x\}$ exist; and after deletion of element h , we have to verify that it does not violate that property w.r.t. the remaining pairs.

Simple insertion (Algorithm 1). New element $h \prec s$ is connected to its one cover $s \in R$. Then, if $s = \mathcal{R}(0)$, we have $h = \mathcal{R}'(0)$. Thus $\inf\{h, x\} = h$ and $\sup\{h, x\} = x$ for all $x \in R'$. Otherwise, we should ensure that \mathcal{R}' is bounded (cf. Def. 3). If \mathcal{R} was already \mathcal{R}_{\perp} , then just connect \perp to h making $\perp \prec h$ in \mathcal{R}'_{\perp} . Else, lift \mathcal{R}' to connect \perp to both h and $\mathcal{R}(0)$. In any case, then, we have $\inf\{h, s\} = h$ and $\sup\{h, s\} = s$, and $\inf\{\perp, h\} = \perp$ and $\sup\{\perp, h\} = h$. Now, it remains to check all pairs $\{h, x\}$ such that $x \in R'$ and $h \parallel x$. Note, however, that $\downarrow h = \{h, \perp\}$ and $\perp \leq x$ for all $x \in R'$. Thus $\{h, x\}^{\ell} = \{\perp\}$, i.e., $\inf\{h, x\} = \perp$.

⁵ As a special element, \perp is not counted for the purpose of the induction step, i.e., its insertion into R (“lifting”) or its removal from R (“unlifting”) is automated and not seen.

Besides, note that $x, s \in R$ and \mathcal{R} is a lattice by the induction hypothesis, therefore $\sup\{s, x\}$ must be well defined in \mathcal{R} for all $x \in R$. Since $\{s, x\}^u \subset R$ does not change in R' with the insertion of h , then $\sup\{s, x\}$ is well defined in \mathcal{R}' for all $x \in R'$ idem. Finally, since $h \prec s$ for exactly one $s \in R'$, then $\sup\{h, x\}$ exists and is the same as $\sup\{s, x\}$.

Blending insertion (Algorithm 2). New element $h \triangleleft S$ is connected to its many covers $s \in S$. For blending insertion, it is required that \mathcal{R} is already lifted \mathcal{R}_\perp . If h is replacing \perp as $\mathcal{R}'(0)$, then there is nothing to verify because \mathcal{R} is a lattice by hypothesis and the connections $h \prec s$ in \mathcal{R}' just replace $\perp \prec s$ in \mathcal{R} for all $s \in S$ (the blending leads to an order-isomorphism $\varphi : \mathcal{R} \rightarrow \mathcal{R}'$). Otherwise, connect \perp to h , and disconnect \perp from all $p \subseteq S$ such that $\perp \prec p$ in \mathcal{R} . Then, for all pairs $\{x, y\} \subseteq S$, we have $\sup\{x, y\}$ in \mathcal{R}' the same as in \mathcal{R} (which, by the induction hypothesis, are well defined), and $\inf\{x, y\} = h$. Finally, for all pairs $\{h, z\}$ such that $z \in R'$ and $h \parallel z$, we have $\inf\{h, z\} = \perp$. By the induction hypothesis, we know that $\sup\{x, z\}$ is well defined for all $x \in S \subset R$. Therefore $\sup\{h, z\}$ must exist for each z such that $z \parallel h$ and coincide to $\sup\{x, z\}$ for some $x \in S$.

Deletion (Algorithm 3). If element $h \in R$, such that $h \neq \top$ and $h \neq \perp$, is $\mathcal{R}(0)$, then just lift \mathcal{R} to \mathcal{R}'_\perp replacing h by \perp in \mathcal{R}' and there is nothing to verify. Else, it has associated non-empty sets $Q' = \{q' \in R \mid q' \prec h\}$, and $P' = \{p' \in R \mid h \prec p'\}$. Consider, further, subsets $Q \subseteq Q'$ such that $Q = \{q \in R \mid q \prec h \text{ and } q \prec x \text{ implies } x = h\}$, and $P \subseteq P'$ such that $P = \{p \in R \mid h \prec p \text{ and } x \prec p \text{ implies } x = h\}$. Then we can verify that \mathcal{R}' is a lattice by means of a twofold analysis, first based on Q and P , and then broadened to their supersets Q' and P' . If $|Q| = 1$ and $|P| \geq 1$, connect the element $q \in Q$ to all $p \in P$ and then remove h from R' . When $|P| > 1$, we have $\inf\{x, y\} = q$ for all $x, y \in P \subset R'$. Now, consider the superset P' of P . For all $w \in P' \setminus P$, (by construction) there exists some $z \in R$ such that $z \neq h$ and $z \prec w$. Take any $p \in P$, and note that in \mathcal{R} we had $\inf\{p, w\} = h$. Moreover, by the induction hypothesis, $\inf\{q, z\} = \ell$ for some $\ell \in R$ (which has not been affected by the deletion of h). Now, in \mathcal{R}' , $\inf\{p, w\}$ is the same as $\inf\{q, z\}$ and, therefore, exists. For $q \in Q$, the analysis of $\sup\{q, r\}$ for all $r \in Q' \setminus Q$ is similar to what we just have seen for the elements of P and P' . Besides, if $|Q| > 1$ and $|P| = 1$, connect all $q \in Q$ to the only element $p \in P$ and then remove h from R' . That is the dual problem, thus the reasoning for the *inf-sup* existence is analogous. Otherwise, the logic follows idem. If $|Q| > 0$ then connect all $q \in Q$ to \top ; if $|P| > 0$ then connect all $p \in P$ to \perp (lifting \mathcal{R} to \mathcal{R}'_\perp in case $\perp \notin R$). Remove h from R' . It is straightforward to verify in \mathcal{R}' that, if $|P| \geq 2$, then $\inf\{x, y\} = \perp$ for all $x, y \in P$; and if $|Q| \geq 2$, then $\sup\{x, y\} = \top$ for all $x, y \in Q$. \square

Algorithm 1 Simple insertion on $\mathcal{R} = \langle R; \prec \rangle$.⁶

```

procedure INSERT( $\mathcal{R} : \text{poset}, h : \text{element}, s : \text{element}$ ) ▷ returns: poset
Require:  $h \notin R, h \neq \perp$ , and  $s \in R, s \neq \perp$ 
Ensure: If  $\mathcal{R}$  given is a lattice, then  $\mathcal{R}$  returned after the insertion of  $h$  is idem
3:  $R \leftarrow R \cup \{h\}$ 
    $\text{zero} \leftarrow \mathcal{R}(0)$  ▷ holds zero of  $\mathcal{R}$ 
    $\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(h, s)\}$  ▷ connects  $h$  to  $s$ 
   if  $\text{zero} = s$  then ▷  $h$  is new zero of  $\mathcal{R}$ 
6:   return  $\mathcal{R}$ 
   else if  $\perp \in R$  then ▷  $\mathcal{R}$  is already lifted  $\mathcal{R}_\perp$ 
    $\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(\perp, h)\}$ 
9:   if  $(\perp, s) \in \mathcal{R}_\prec$  then
      $\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \setminus \{(\perp, s)\}$  ▷ disconnects  $\perp$ 
   else ▷ lifts  $\mathcal{R}$  to  $\mathcal{R}_\perp$ 
12:   $R \leftarrow R \cup \{\perp\}$ 
      $\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(\perp, h), (\perp, \text{zero})\}$ 
return  $\mathcal{R}$ 

```

Algorithm 2 Blending insertion on $\mathcal{R} = \langle R; \prec \rangle$.

```

procedure INSERT( $\mathcal{R} : \text{poset}, h : \text{element}, S : \text{set}$ ) ▷ returns: poset
Require:  $h \notin R, \perp \in R$ , and  $S \subset R$  is parsimonious (cf. Axiom 1)
Ensure: If  $\mathcal{R}$  given is a lattice, then  $\mathcal{R}$  returned after the insertion of  $h$  is idem
3: if  $\perp \triangleleft S$  then ▷  $h$  is replacing  $\perp$  as  $\mathcal{R}(0)$ 
   for all  $s \in S$  do
      $\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(h, s)\}$ 
6:    $\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \setminus \{(\perp, s)\}$ 
    $R \leftarrow R \setminus \{\perp\}$ 
else ▷  $h$  is set  $\perp \prec h \prec s$  for all  $s \in S$ 
9:   for all  $s \in S$  do
      $\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(h, s)\}$ 
     if  $(\perp, s) \in \mathcal{R}_\prec$  then
12:     $\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \setminus \{(\perp, s)\}$ 
      $\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(\perp, h)\}$ 
return  $\mathcal{R}$ 

```

⁶ By a slight abuse of notation, we write \mathcal{R}_\prec instead of \prec to refer to the set of covering pairs of $\mathcal{R} = \langle R; \prec \rangle$.

Algorithm 3 Deletion on $\mathcal{R} = \langle R; \prec \rangle$.

	procedure DELETE($\mathcal{R} : poset, h : element$)	\triangleright returns: <i>poset</i>
	Require: $h \neq \top$ and $h \neq \perp$	
	Ensure: If \mathcal{R} given is a lattice, then \mathcal{R} returned after the deletion of h is idem	
3:	if $h = \mathcal{R}(0)$ then	\triangleright lifts \mathcal{R} to \mathcal{R}_\perp
	$R \leftarrow R \cup \{\perp\}$	
	$S \leftarrow \{s \in R \mid h \prec s\}$	
6:	for all $s \in S$ do	
	$\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(\perp, s)\}$	
	$\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \setminus \{(h, s)\}$	\triangleright disconnects h
	else	
9:	$C \leftarrow \{c \in R \mid c \prec h \text{ or } h \prec c\}$	
	$Q \leftarrow \{q \in C \mid q \prec h \text{ and } q \prec x \text{ implies } x = h\}$	
	$P \leftarrow \{p \in C \mid h \prec p \text{ and } x \prec p \text{ implies } x = h\}$	
12:	if $ Q = 1$ and $ P \geq 1$ then	
	for all $p \in P$ do	
	$\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(q, p)\}$	
15:	else if $ Q > 1$ and $ P = 1$ then	
	for all $q \in Q$ do	
	$\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(q, p)\}$	
18:	else	
	for all $q \in Q$ do	
	$\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(q, \top)\}$	
21:	if $\perp \notin R$ and $ P > 0$ then	\triangleright lifts \mathcal{R} to \mathcal{R}_\perp
	$R \leftarrow R \cup \{\perp\}$	
	for all $p \in P$ do	
24:	$\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \cup \{(\perp, p)\}$	
	for all $c \in C$ do	
	if $c \prec h$ then	\triangleright disconnects h
27:	$\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \setminus \{(c, h)\}$	
	else	
	$\mathcal{R}_\prec \leftarrow \mathcal{R}_\prec \setminus \{(h, c)\}$	
30:	$R \leftarrow R \setminus \{h\}$	\triangleright removes h
	return \mathcal{R}	

Appendix B: Hypothesis in Tabular Form

Table 1. Hypotheses in Celestial Mechanics that explain phenomena of *motion of planets in the solar system*. Adapted from: [11].

ID	Author Name	Statement
h_1	Galileo	Falling bodies "The velocities of a freely falling body are proportional to the times."
h_2	Kepler	Elliptical orbits "Each planet moves in an ellipse with the sun in one focus."
h_3	Newton	Gravitation "The gravitational force exerted by the sun on each planet is proportional to its mass and inversely proportional to the square of its distance to the sun."

Table 2. Hypotheses in Population Dynamics that explain phenomena of *population growth and decay*. Adapted from: [4].

ID	Author	Name	Statement
h_1	Malthus	Intrinsic reaction	"Population changes by intrinsic reaction measured as natality and mortality rates."
h_2	Verhulst	Intra-species competition	"Population growth is limited by competition for resources among individuals of the same species."
h_3	Verhulst	Logistic growth	"Population grows by intrinsic reaction, but limited by competition."
h_4	Lotka, Volterra	Trophic interactions	"Population growth/decay is limited by inter-species trophic interactions."
h_5	Lotka, Volterra	Predator-prey coupling	"Population grows/decays by intrinsic reaction and is regulated by predator-prey interactions."

Table 3. Hypotheses in Molecular Biology that explain phenomena of *genetic information storage in living cells*. Adapted from: [14].

ID	Author	Name	Statement
h_1	Miescher	Nucleic acid molecule	"The nucleic acid molecule is a new substance that is in some sense equivalent to a protein."
h_2	Levene	Polynucleotide structure	"Nucleic acids are composed of a series of nucleotides, with each nucleotide composed of just one of four nitrogen-containing bases, a sugar molecule, and a phosphate group."
h_3	Avery et al.	DNA storage of genetic information	"DNA is the exclusive chromosomal component bearing the genetic information of living cells."
h_4	Chargaff et al.	Chargaff's rule	"The total amount of purines and the total amount of pyrimidines are nearly equal."
h_5	Watson, Crick	DNA double helix	"DNA has a three-dimensional double-helical structure."