Partially ordered set merging

Elizabeth Sprangel

May 20th, 2022

1 Introduction

Some data is equipped with multiple different inherent orderings. For example, consider the employees for a company. On one hand, it is possible to impose a (partial) order based on company hierarchy, where employee A precedes employee B in the partial order whenever employee A oversees employee B in the company hierarchy. On the other hand, it is possible to impose the interval order obtained from assigning each employee an interval based on their time employed by the company. Is there a way to simultaneously represent multiple orderings on the same data set?

A *partially ordered set* (or *poset*) is a pair $(S, \leq)$ consisting of a set S and a binary relation $\leq$ such that, for all $u, v, w \in S$, the following three axioms hold:

1. $u \leq u$ (reflexivity).
2. If $u \leq v$ and $v \leq u$, then $u = v$ (anti-symmetry).
3. If $u \leq v$ and $v \leq w$, then $u \leq w$ (transitivity).

Throughout this document, we will consider a fixed underlying set S and binary relations $\leq_1, \ldots, \leq_k$, each of which forms a partially ordered set with S. We suppress the notation $(S, \leq_i)$ when the underlying set is clear and instead refer this poset as simply $\leq_i$. The *union* of two orders $\leq_1$, $\leq_2$ over the same underlying set S is the set of relations R where $a \leq_R b$ for $a, b \in S$ whenever $a \leq_1 b$ or $a \leq_2 b$. In general, the *union* of k orders $\leq_1, \ldots, \leq_k$ over the same underlying set S is the set of relations R where $a \leq_R b$ for $a, b \in S$ if $a \leq_i b$ for some $i \in [k]$.

For two binary relations $\leq_1$ and $\leq_2$ on set S, we say $\leq_2$ is an *extension* of $\leq_1$ if the identity map $1 : (S, \leq_1) \to (S, \leq_2)$ is order-preserving. In other words, $\leq_2$ is an extension of $\leq_1$ if $a \leq_1 b$ implies $a \leq_2 b$ for all $a, b \in S$. If $\leq_2$ is a totally ordered set, or a chain, then $\leq_2$ is a *linear extension* of $\leq_1$. Linear extensions are regarded as one of the best measures of complexity for posets, and are hence widely studied. However, this document is concerned with identifying an extension of *minimal* size which simultaneously extends multiple posets. In many cases, we expect this to be a *nonlinear extension*, meaning some elements may remain unrelated.

Question 1.1. *Given a set of orders $\leq_1, \leq_2, \ldots, \leq_t$ on the same underlying set, is there a partial order $\leq$ on the same underlying set which is an extension of $\leq_i$ for each $i = 1, \ldots, t$? If such an order exist, can we identify a minimal such order?*

Liz: Look into whether this is actually well posed or if it is in the literature somewhere...i.e. in the case when posets are compatible what is the minimal extention when is it linear? Liz:

In general, the answer to the question is no. Consider the set $\{a, b, c\}$. Define partial order $\leq_1$ given by the relations $a <_1 b$, $b <_1 c$, and $a <_1 c$, as in Figure 1. Define $\leq_2$ to be given by $a <_1 b$ and $c <_1 b$. The union of these two posets has both $b <_1 c$ and $c <_2 b$. Meaning, no partial order can naturally extend both $\leq_1$ and $\leq_2$. However, both orders preserve the relation $a < b$. That is, $\leq_1$ and $\leq_2$ agree on some relations and conflict on others.

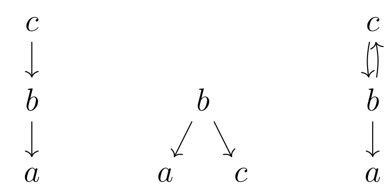


Figure 1: Hasse diagrams for partial orders $\leq_1$ (left), $\leq_2$ (center), and the union of $\leq_1$ and $\leq_2$ (right).

Since finding a common extension is not feasible in general, we instead turn our attention to inspecting the amount of incompatability between multiple partial orders. For example, in Figure 1, the union of the two orders preserves the relation $a < b$. We would like to assign a consistent function to measure this incompatibility.

Question 1.2. *Suppose $\leq_1, \leq_2, \ldots, \leq_t$ are partial orders on the same underlying set such that there is no partial order $\leq$ on the same underlying set which simultaneously extends each $\leq_i$ for $i = 1, \ldots, t$. Is there a consistent way to quantify the amount of conflict between the constituent orders?*

Liz: Define digraph and $E(D)$ and $V(D)$

Given a relation R on S, form a directed graph (digraph) D_R on the vertex set S where (a, b) is an arc from a to b whenever $a <_R b$. We do not include loops. If R is a poset, then D_R is a directed acyclic graph (DAG) because it is anti-symmetric (i.e. D_R has no cycles). To easily visualize poset R, we often consider the Hasse diagram, which is a drawing of the transitive reduction of R. Figure 1 shows two Hasse diagrams obtained from posets on $\{a, b, c\}$, and a DAG representing the union of the first two diagrams.

A *strongly connected component* of a directed graph D is a maximal subgraph which has a directed path between any two vertices. The set of strongly connected components forms a partition of the vertices of the graph. The *condensation*, denoted $\mathcal{C}(D)$, of a digraph D is the digraph obtained from D by contracting all strongly connected components to a single vertex. Because there are no nontrivial strongly connected components in the condensation digraph, the resulting digraph is acyclic and hence a DAG. See Figure 2 for an example of a condensation digraph.

$$
\begin{tikzcd}
c \arrow[d, shift right] & \\
b \arrow[u, shift right] \arrow[d] & \{b, c\} \arrow[d] \\
a & a
\end{tikzcd}
$$

Figure 2: The digraph on the right is the condensation DAG of the digraph on the left.

Getting back to Question 1.2, let $\leq_1, \ldots, \leq_k$ be a list of k posets on the same underlying set S and suppose R is the union of all k posets. Consider the condensation of the associated DAG, or $\mathcal{C}(D_R)$. Let the *edge compatibility* of orders $\leq_1, \ldots, \leq_k$ be defined as

$$\operatorname{comp}_E(\leq_1, \ldots, \leq_k) = \begin{cases} \frac{|E(\mathcal{C}(D_R))|}{|E(D_R)|} & \text{if } E(D_R) > 0, \\ 1 & \text{if } E(D_R) = 0. \end{cases}$$

Similarly, define the *node compatibility* of orders $\leq_1, \ldots, \leq_k$ to be

$$\operatorname{comp}_V(\leq_1, \ldots, \leq_k) = \begin{cases} \frac{|V(\mathcal{C}(D_R))|}{|V(D_R)|} & \text{if } V(D_R) > 0, \\ 1 & \text{if } V(D_R) = 0. \end{cases}$$

Liz: I don't love the term 'compatibility' and the notation comp$_E(\leq_1, \leq_2)$. Open to ideas to change them, and it's easy enough to update in the future...

Loosely speaking, the edge compatibility measures what proportion of the relations in $D_\leq$ are non-contradictory. That is, it is the portion of edges not condensed in $\mathcal{C}(D_R)$

If the union of posets $\leq_1, \ldots, \leq_k$ is also a poset, then, in turn, the condensation graph $\mathcal{C}(D_\leq)$ is isomorphic to $D_\leq$. This is because if the transitive property holds in $\leq$, then the associated digraph $D_\leq$ is acyclic. Hence, every strongly connected component in $D_\leq$ is already a singleton vertex.

2 Examples

2.1 Chains and anti-chains

Let $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ be the underlying set for the following examples. First, we examine the case when $\leq_1$ and $\leq_2$ are both the same chain on S. Since $\leq_1$ and $\leq_2$ are indistinguishable, $D_\leq \cong D_{\leq_1}$ and since $D_{\leq_1}$ is a DAG, so is $D_\leq$. Therefore, $\mathcal{C}(D_\leq) \cong D_\leq$ and so the edge and node compatilibities are both 1. In fact, the edge (node) compatibility for a list of identical posets is always 1.

$D_{\leq_1}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$D_{\leq_2}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$D_{\leq}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$\mathcal{C}(D_{\leq})$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$\mathrm{comp}_E(\leq_1, \leq_2)$	1
$\mathrm{comp}_V(\leq_1, \leq_2)$	1

Next, we compare two chains, where one is the reverse of the other. In this case, if $a \leq_1 b$, then $b \leq_2 a$ and a, b form a 2-cycle in the union. Thus, the condensation is a singleton node.

$D_{\leq_1}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$D_{\leq_2}$	$9 \longleftarrow 8 \longleftarrow 7 \longleftarrow 6 \longleftarrow 5 \longleftarrow 4 \longleftarrow 3 \longleftarrow 2 \longleftarrow 1 \longleftarrow 0$
$D_{\leq}$	$0 \leftrightarrows 1 \leftrightarrows 2 \leftrightarrows 3 \leftrightarrows 4 \leftrightarrows 5 \leftrightarrows 6 \leftrightarrows 7 \leftrightarrows 8 \leftrightarrows 9$
$\mathcal{C}(D_{\leq})$	0
$\mathrm{comp}_E(\leq_1, \leq_2)$	0
$\mathrm{comp}_V(\leq_1, \leq_2)$	0.1

Let $\leq_1$ be a chain on S and let $\leq_2$ be the anti-chain with vertex set S. When performed on the same vertex set, the union of any DAG with an anti-chain results in the original DAG. Thus, the edge and node compatibility are both 1.

$D_{\leq_1}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$D_{\leq_2}$	$0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9$
$D_{\leq}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$\mathcal{C}(D_{\leq})$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$\mathrm{comp}_E(\leq_1, \leq_2)$	1
$\mathrm{comp}_V(\leq_1, \leq_2)$	1

Once again let $\leq_1$ be a chain on S. Let $\leq_2$ be defined on vertex set S with one edge $(0, 9)$. Since $D_{\leq}$ is a cycle running through all vertices in S, the condensation is a singleton node. Hence, the edge compatibility is 0.

$D_{\leq_1}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$D_{\leq_2}$	$0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9$
$D_{\leq}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$\mathcal{C}(D_{\leq})$	0
$\mathrm{comp}_E(\leq_1, \leq_2)$	0
$\mathrm{comp}_V(\leq_1, \leq_2)$	0.1

Let $\leq_1$ be a chain on S. Let $\leq_2$ be defined on S with one edge, $(0, 3)$. Since $1 \to 2 \to 3 \to 1$ forms a cycle, the condensation is a path on 7 nodes. Hence, the edge compatibility is 0.6 and the node compatibility is 0.7.

$D_{\leq_1}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$D_{\leq_2}$	$0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9$
$D_{\leq}$	$0 \longleftarrow 1 \longleftarrow 2 \longleftarrow 3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$\mathcal{C}(D_{\leq})$	$3 \longleftarrow 4 \longleftarrow 5 \longleftarrow 6 \longleftarrow 7 \longleftarrow 8 \longleftarrow 9$
$\text{comp}_E(\leq_1, \leq_2)$	0.6
$\text{comp}_V(\leq_1, \leq_2)$	0.7

Notice, this immediately implies there exists a pair of orders $\leq_1$ and $\leq_2$ such that the edge compatibility is r for any rational in $[0, 1]$. Namely, let $r = p/q \in (0, 1]$ for positive integers p and q. Let $\leq_1$ be the path $1, 2, \ldots, q$. Let $\leq_2$ be defined on $\{1, 2, \ldots, q\}$ with one edge, $(1, p)$.

$D_{\leq_1}$	$1 \longleftarrow \cdots \leftarrow p-1 \leftarrow p \leftarrow p+1 \leftarrow \cdots \leftarrow q-1 \leftarrow q$
$D_{\leq_2}$	$1 \quad \cdots \quad p-1 \quad p \quad p+1 \quad \cdots \quad q-1 \quad q$
$D_{\leq}$	$1 \longleftarrow \cdots \leftarrow p-1 \leftarrow p \leftarrow p+1 \leftarrow \cdots \leftarrow q-1 \leftarrow q$
$\mathcal{C}(D_{\leq})$	$p \leftarrow p+1 \leftarrow \cdots \leftarrow q-1 \leftarrow q$
$\text{comp}_E(\leq_1, \leq_2)$	p/q
$\text{comp}_V(\leq_1, \leq_2)$	$(p+1)/q$

2.2 Boolean lattices

For the following examples, let $S = \mathcal{P}(\{1, 2, 3\}) = \{\{1, 2, 3\}, \{2, 3\}, \{1, 3\}, \{1, 2\}, \{3\}, \{2\}, \{1\}, \emptyset\}$. The *Boolean lattice* for a family of sets S is a DAG obtained by including an edge (A, B) if and only if $B \subseteq A$ for sets $A, B \in S$.

Consider a lexicalgraphical ordering on $S = \mathcal{P}(\{1, 2, 3\})$. That is, the dictionary order on $S = \mathcal{P}(\{1, 2, 3\})$ induced by relations $1 < 2 < 3$. The following table shows the edge compatibility of four different chains with the Boolean lattice. Notice, the Boolean lattice and the lexicalgraphical ordering are not compatible, as they have an edge compatibility of 0. Furthermore, the reverse lexicalgraphical order also has low compatibility with the boolen lattice, at 0.2.

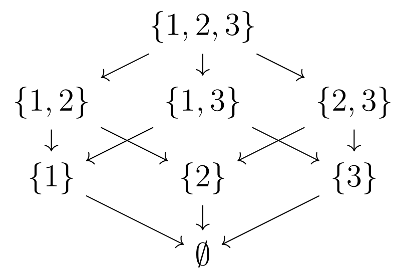


Figure 3: The Boolean lattice on $\mathcal{P}(\{1,2,3\})$.

Depiction	Edge compatibility
$\emptyset \longleftarrow \{1\} \longleftarrow \{2\} \longleftarrow \{3\} \longleftarrow \{1,2\} \leftarrow \{1,3\} \leftarrow \{2,3\} \leftarrow \{1,2,3\}$	1
$\emptyset \longrightarrow \{1\} \longrightarrow \{2\} \longrightarrow \{3\} \longrightarrow \{1,2\} \rightarrow \{1,3\} \rightarrow \{2,3\} \rightarrow \{1,2,3\}$	0
$\emptyset \longrightarrow \{1\} \longrightarrow \{1,2\} \rightarrow \{1,2,3\} \rightarrow \{1,3\} \longrightarrow \{2\} \longrightarrow \{2,3\} \longrightarrow \{3\}$ (Lexicalgraphic)	0
$\emptyset \longleftarrow \{1\} \longleftarrow \{1,2\} \leftarrow \{1,2,3\} \leftarrow \{1,3\} \longleftarrow \{2\} \longleftarrow \{2,3\} \longleftarrow \{3\}$ (Reverse lexical.)	0.2

The Boolean lattice is a highly structured poset, with many symmetries. A natural question is "How close does the Boolean DAG resemble a random DAG?" The following table depicts six pseudo-randomly generated DAGs and their edge-compatibility with the Boolean lattice. The pseudo-random DAGs were generated by taking the condensation of instances of the random graph $G(n,p)$. The average edge compatibility between the Boolean DAG on $\{1,2,3\}$ and $10,000$ interations of the pseudo-random DAG on 8 vertices is 0.131716. Meaning, in general, the edge compatibility between a random DAG and the Boolean lattice is low, at about 0.13.

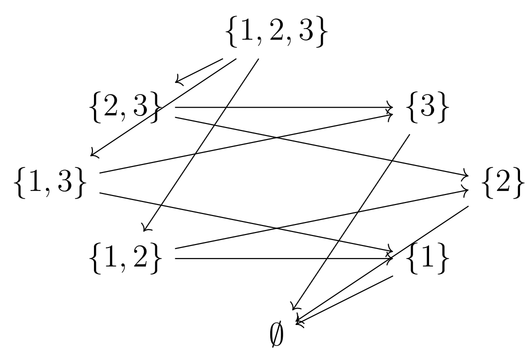
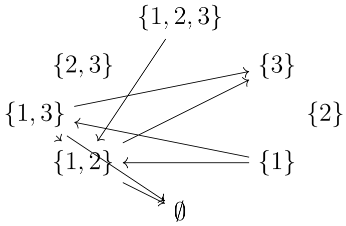
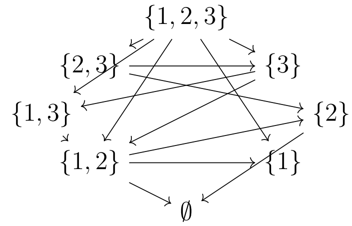
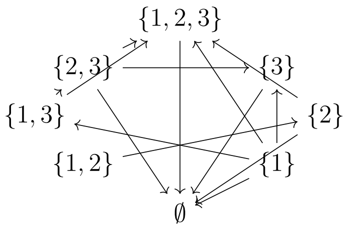
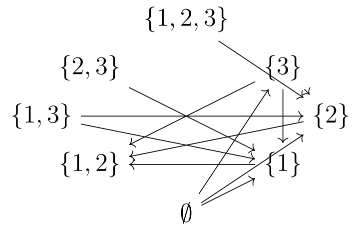
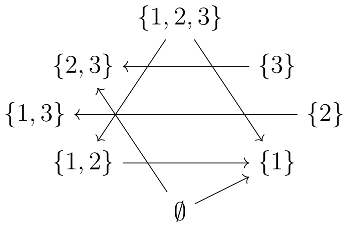
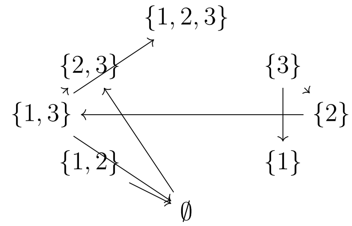


Figure 4: The Boolean lattice on $\mathcal{P}(\{1,2,3\})$ drawn in a circular layout.

Random DAG	Edge compatibility	Random DAG	Edge compatibility
	0.5		0.7778
	0.142857		0.2173913
	0.1764706		0

3 Some observations

Liz: Look into what necessary an sufficient conditions are needed for two posets to have edge and/or vertex compatibility 1 (or 0)... Below is one example (the trivial one)

Observation 3.1. *Let $(S, \leq_1)$ and $(S, \leq_2)$ be partially ordered sets on underlying set S. The following are equivalent:*

1. *$comp_E(\leq_1, \leq_2) = 1$,*
2. *the union of $D_{\leq_1}$ and $D_{\leq_2}$ is a DAG,*
3. *the transitive closure of the union of $(S, \leq_1)$ and $(S, \leq_2)$ is a poset.*

Observation 3.2. *Let $(S, \leq_1), (S, \leq_2), \ldots, (S, \leq_k)$ be partially ordered sets on underlying set S. The following are equivalent:*

1. *$comp_E(\leq_1, \ldots, \leq_k) = 1$,*
2. *the union of $D_{\leq_1}, \ldots, D_{\leq_k}$ is a DAG,*
3. *the transitive closure of the union of $\leq_1, \ldots, \leq_k$ is a poset.*