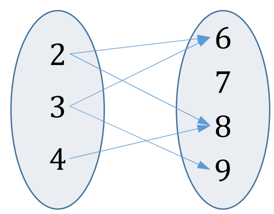
Relations



[Wikipedia:](https://en.wikipedia.org/wiki/Relation_(mathematics))
https://en.wikipedia.org/wiki/Relation_(mathematics)
https://en.wikipedia.org/wiki/Binary_relation

1

A relational database

Books:

id	title	Area	type	author id	translator id
012	Intro to Python	CS	original	11	33
200	Discrete Math	Math	translated	14	–
053	Algorithms	CS	original	14	32
437	Deep Learning	CS	original	11	–
525	Robotics	CS	translated	12	31
363	Intro to Java	CS	translated	15	32
277	Database Systems	CS	translated	11	31

Authors:

id	First_name	Last_name
11	Ellen	Writer
12	Yao	Dou
14	Mario	Lopez
15	Olga	Solver

Translators:

id	First_name	Last_name
31	Ira	Davies
32	Barbara	Jones
33	Roman	Edwards

2

2

Training data in machine learning

User ID	Gender	Age	Salary	Purchased
15624510	Male	19	19000	0
15810944	Male	35	20000	1
15668575	Female	26	43000	0
15603246	Female	27	57000	0
15804002	Male	19	76000	1
15728773	Male	27	58000	1
15598044	Female	27	84000	0
15694829	Female	32	150000	1
15600575	Male	25	33000	1
15727311	Female	35	65000	0
15570769	Female	26	80000	1
15606274	Female	26	52000	0
15746139	Male	20	86000	1
15704987	Male	32	18000	0
15628972	Male	18	82000	0
15697686	Male	29	80000	0
15733883	Male	47	25000	1

Temperature	Pressure	Relative Humidity	Wind Direction	Wind Speed
10.69261758	986.882019	54.19337313	195.7150879	3.278597116
13.59184184	987.8729248	48.0648859	189.2951202	2.909167767
17.70494885	988.1119385	39.11965597	192.9273834	2.973036289
20.95430404	987.8500366	30.66273218	202.0752869	2.965289593
22.9278274	987.2833862	26.06723423	210.6589203	2.798230886
24.04233986	986.2907104	23.46918024	221.1188507	2.627005816
24.41475295	985.2338867	22.25082295	233.7911987	2.448749781
23.93361956	984.8914795	22.35178837	244.3504333	2.454271793
22.68800023	984.8461304	23.7538641	253.0864716	2.418341875
20.56425726	984.8380737	27.07867944	264.5071106	2.318677425
17.76400389	985.4262085	33.54900114	280.7827454	2.343950987
11.25680746	988.9386597	53.74139903	68.15406036	1.650191426
14.37810685	989.6819458	40.70884681	72.62069702	1.553469896
18.45114201	990.2960205	30.85038484	71.70604706	1.005017161
22.54895853	989.9562988	22.81738811	44.66042709	0.264133632
24.23155922	988.796875	19.74790765	318.3214111	0.329656571

Applicant Information	
age	25 year
gender	Male (0)
annual salary	$70,000
years in home	15 months
years in job	8 months
current debt	$27,489

Proteind	Rice	Dairy	Beans	Sauce	Cust.ID
beef	*white*	*sour cream*	*black*	*mild*	7854
beef	*cilantro*	*cheese*	*black*	*medium*	7854
chicken	*brown*	*none*	*black*	*hot*	3512
pork	*cilantro*	*cheese*	*pinto*	*hot*	8701
chicken	*white*	*cheese*	*black*	*medium*	3512

3

3

Relations

- A k***-ary relation*** over sets $A_1, A_2 \dots, A_k$ is a subset of $A_1 \times A_2 \times \cdots \times A_k$
  - Main data abstraction in relational database systems, machine learning, and more

Example. BETWEENNESS $\subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, is a ternary relation that includes (3,4,7) and $(3.14, \pi, 3.15)$, but not (4,3,7)

/usr/share/X11/rgb.txt

```
220  20  60  crimson
75    0 130  indigo
128 128   0  olive
102  51 153  rebecca purple
102  51 153  RebeccaPurple
192 192 192  silver
  0 128 128  teal
```

Example. RGBCOLOR is a 4-ary relation on $\{0,1,\dots,255\}^3 \times$ NAMES. The tuple $(255, 255, 0, \text{yellow}) \in$ RGBCOLOR

Example. COURSES is a 8-ary relation on DEPARTMENT, COURSE-NUMBER, $\mathbb{N}$ (# of credits), CRN, INSTRUCTOR, BUILDING, ROOM, and TIMES. The tuple (COMP, 3200, 4, 1371, Lopez, ECS, 301, MW 2pm) $\in$ COURSES

- We will focus on binary relations, as many of the important results for $k = 2$ generalize to $k \geq 3$

4

4

Binary relations

- When $k = 2$, relations generalize the concept of a function
  - All functions are relations but not all relations are functions

Definition. If A and B are sets, a ***binary relation*** R from A to B, denoted $R : A \to B$, is a subset of the Cartesian product $A \times B$.

– The relation R is a set of ordered pairs (a, b) with $a \in A$ and $b \in B$

– If $(a, b) \in R$ we also write aRb and say that a ***is related to*** b

– A is the *domain* of R and B is the *codomain* of R

*Example***.** Given $P =$ set of people and $C =$ set of countries, then *citizen-of* is a relation from P to C

*Example***.** The relation *written-in* ⊂ *apps*×*programming-languages* includes tuples (Instagram, Python), (Uber, Go), (Gmail, Java), (Winamp, C++)

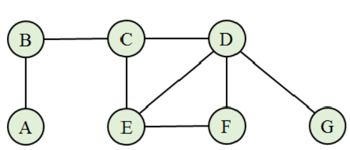
*Example***.** An online retailer is interested in the relation *purchased*, a subset of *Customers* × *Products*

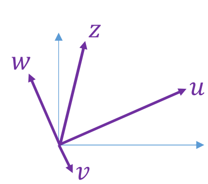
5

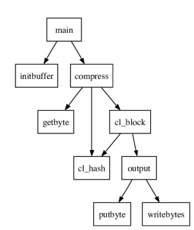
5

Binary relations…

- Binary relations are used in many branches of math and CS to model a wide variety of useful concepts:
  - *is-adjacent-to* is a relation in graph theory
  - *is-greater-than*, *is-equal-to*, etc. are relations in arithmetic
  - *is-orthogonal-to* is a relation in vector algebra
  - *is-nearest-to* is a relation widely used in machine learning
  - *calls* is a relation among functions in a software system
  - *is-a-friend-of* is a relation commonly used in social networks







6

6

Relations on a single set

- It is allowed (and common) for $A = B$ and, in this case, we simply say that *R is a relation on A*
- This includes various arithmetic relations defined on $\mathbb{N}$, such as $<$, $\leq$, $>$, $\geq$, $=$, $\neq$. For example, $(3,7) \in \leq$, while $(3,7) \notin =$

  *Example***.** Given P = set of people, *friend-of* is a relation on P

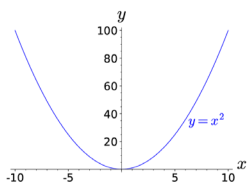
  *Example***.** Given C = set of courses and S = set of students, *prerequisite-of* is a relation on C. A student $s \in S$ can register for course c if for every course p such that $(p, c) \in$ *prerequisite-of* , s has taken p ⇒ DU stores the relation *prerequisite-of* as well as the relation *taken* $\subseteq S \times C$

Examples

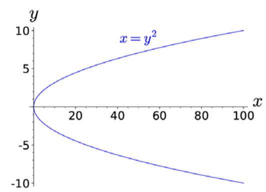
- Given relations: *Prereq-of* ⊆ *Courses* × *Courses* and *Passed* ⊆ *Students* × *Courses,* we may want to compute:
  1. Classes whose prerequisites you have satisfied but have not passed
  2. People with whom you have taken at least one class and have passed a class that satisfies (1)
- Given the relation *Purchased* ⊆ *Customer* × *Products*, a retailer may want to find for each user u all products not purchased by u that have been purchased by other uses that share at least k purchases with u

Relation vs. function

- You may have noticed that a function X is simply a special type of relation, one where each member of X appears *at most once* as the first element of a pair (if every member of X appears *exactly* once, the function is total)
- Unlike functions, which may only be *one-to-one* or *many-to-one*, relations can be *one-to-one*, *many-to-one*, or *one-to-many*



Function f: $(a, b) \in f$ if $b = a^2$



Relation R: $(a, b) \in R$ if $a = b^2$

Example. Given X = set of people and Y = set of countries. Then BORN-IN is a relation which is a function, while CITIZEN-OF is a relation but not a function.

9

9

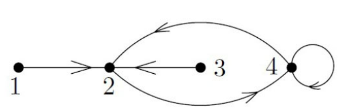
Computer representation

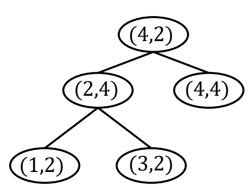
- How do you define and represent a relation $R: A \to B$ in a computer program? What operations would be useful to support?

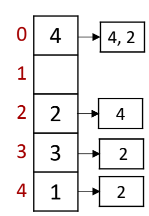
1. Set representation (e.g., AVL tree, hash table)
2. Boolean matrix
3. Graph (directed if $A = B$, bipartite if $A \neq B$)
4. Procedural representation: $A \times B \mapsto$ *Boolean*

Example. Representation of $\{(1,2), (2,4), (3,2), (4,2), (4,4)\}$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$







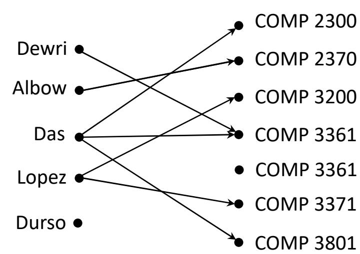
10

10

Example

- *Teaches* is a relation from *Faculty* to *Classes* that includes the following tuples: (Das,COMP2300), (Dewri,COMP3361), (Albow,COMP2370), (Das,COMP3361), (Lopez,COMP3371), (Das,COMP3801), (Lopez,COMP3200)

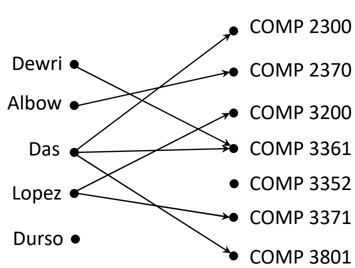
$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$



Terminology

Let $R: X \to Y$ be a relation and $Z \subseteq X$

- The ***image*** of Z under R, denoted $R(Z)$, is $R(Z) = \{y \in Y \mid zRy \text{ for some } z \in Z\}$
- The ***range*** of R is the image of X
- The ***inverse*** R^{-1} is the relation from Y to X such that $yR^{-1}x$ iff xRy



Example. The image of COMP3361 under TEACHES^{-1} is $\{\text{Dewri, Das}\}$

Exercise. Let R be the relation on $\mathbb{R}$ defined as follows: $(a, b) \in R$ iff $b = |a|$. Draw the locus of points $(a, b) \in R^{-1}$

Basic relation properties

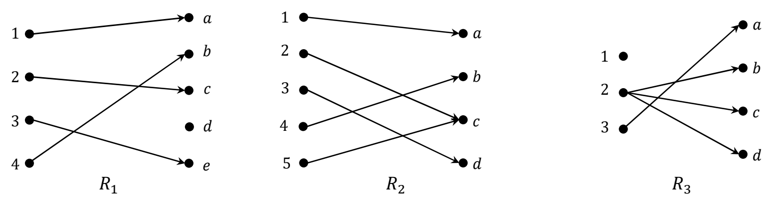
- A relation $R : X \to Y$ is:

surjective or ***onto*** if $R(X) = Y$, i.e., each $y \in Y$ is incident with at least one $x \in X$

injective or **1-1** if every $y \in Y$ is incident with at most one $x \in X$

total if $R^{-1}(Y) = X$, i.e., every $x \in X$ is related to at least one $y \in Y$

bijective if R is total, injective, and surjective



- A database schema may allow you to enforce a subset of these

13

13

Exercise

- Fill in the right column in the table below, using the basic properties that relations may satisfy (*surjective*, *injective*, *bijective*, *total*, or *partial function*)
  - Recall that "function" includes both partial and total functions

R is	iff R^{-1} is
total	
surjective	
a function	
injective	
bijective	

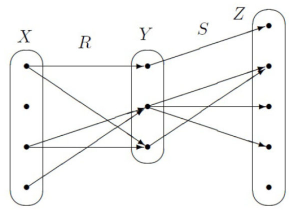
Hint: think about what is going on in terms of arrows from the domain X to the codomain Y

14

14

Composition

Relations can be composed in the same way as functions. Let $R \subseteq X \times Y$ and $S \subseteq Y \times Z$. Then $R \cdot S$ (or simply RS) is the relation $T \subseteq X \times Z$ such that xTz iff there is $y \in Y$ such that xRy and ySz.



*Example***.** The US Senate has two senators per state, each affiliated with zero or one political party. Given relations $S \subset$ *States* $\times$ *Senators* and $T \subset$ *Senators* $\times$ *Parties*, then (*state, party*) $\in S \cdot T$ iff $\exists$ senator *s* that represents *state* such that *s* is affiliated with *party*

Note. If R and S are functions, then $R \cdot S \equiv S \circ R$

15

15

Exercise

1. Let S be the relation *is-a-sibling-of* and C the relation *is-a-child-of*, both on the set of people. In other words, $(x, y) \in S$ if x is a sibling of y, and $(a, b) \in C$ if a is a child of b.
2. What is the natural interpretation of $C \cdot S$ and of $S \cdot C$?
3. Let M_R be the matrix representation of relation $R : A \rightarrow B$ and M_S the matrix representation of relation $S : B \rightarrow C$. How do you compute the matrix representation of $R \cdot S \subseteq A \times C$ using M_R and M_S?

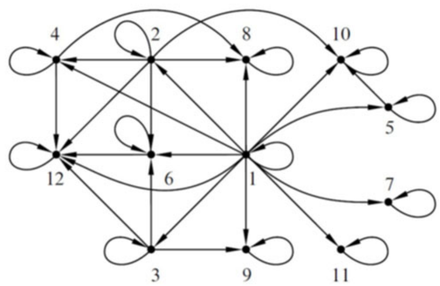
16

16

Relations on the Same Set

- A relation on X is a relation $R : X \to X$
- A relation $R: X \to X$ is often represented with a directed graph

Example. $X = \{1,2,\ldots,12\}$ with the relation “divides” denoted |
Thus 3 | 9 while 3 ∤ 7



17

17

Exercise

Given the following relations R on $\mathbb{Z}^{\geq 1}$:

1. *successor* $= \{(n, n+1) : n \in \mathbb{Z}^{\geq 1}\}$
2. *equals* $= \{(n, n) : n \in \mathbb{Z}^{\geq 1}\}$
3. *relPrime* $= \{(n, m) : \gcd(n, m) = 1, n, m \in \mathbb{Z}^{\geq 1}\}$

Describe in each case the relation $R \cdot R$

18

18

Properties of relations on the same set

Definition. We say that a relation R on X is

reflexive if xRx for all $x \in X$

irreflexive if $\nexists x \in X$ such that xRx

symmetric if $xRy \Rightarrow yRx$, for all $x, y \in X$

antisymmetric if $(x, y) \in R$ and $x \neq y \Rightarrow (y, x) \notin R$, i.e., if $x \neq y$, xRy and yRx don't hold simultaneously

transitive if xRy and $yRz \Rightarrow xRz$ for all $x, y, z \in X$

- In terms of our matrix and graph representations:
  - A relation R is reflexive if the diagonal of its matrix M_R consists entirely of 1's
  - R is symmetric if $M_R = M_R^T$, i.e., if its matrix M_R is equal to its transpose
  - In a symmetric relation the directed graph can be replaced by the simpler undirected graph

19

19

Other relation types

- A relation R on X is an ***equivalence*** if it is reflexive, symmetric, and transitive
- A relation R on X is a ***partial order*** if it is antisymmetric, transitive, and reflexive
  - A partial order corresponds to the notion of $\leq$
- A relation R on X is a ***strict partial order*** if it is irreflexive, antisymmetric and transitive
  - A strict order corresponds to the notion of $<$
- A relation R on X is a ***linear*** or ***total ordering*** if it is a partial or strict partial order and for any distinct $x, y \in X$, either xRy or yRx

Exercise. What type of relation is one that is irreflexive and transitive?

20

20

Equivalence relations

Definition. Let R be an equivalence relation on set X and $x \in X$. The ***equivalence class*** of x, denoted $R[x]$, is defined as $R[x] = \{y \in X : xRy\}$. When R is understood, we simply write $[x]$ instead of $R[x]$.

Claim. For any equivalence R on X, we have:
1. The equivalence class $R[x]$ is nonempty for all $x \in X$.
2. For any two elements $x, y \in X$, either $R[x] = R[y]$ or $R[x] \cap R[y] = \emptyset$.
3. The equivalence classes uniquely determine the relation R.

Note: The equivalence classes of an equivalence relation on X constitute a *partition* of X

Exercise

- For each set and relation, indicate the type of relation:
  1. Set of people, *cousin-of*
  2. Set of people, *parent-of*
  3. Nodes of undirected graph, xRy if there is a path from x to y
  4. Set $\mathbb{N}$ of numbers, xRy if $y \bmod x = 0$
  5. Set $\mathbb{N}$ of numbers, xRy if $x - y$ is even
  6. Set $\mathbb{Q}$ of numbers, xRy if $x \leq y$
  7. Set $\mathbb{Q}$ of numbers, xRy if $x < y$
  8. Set of 2D vectors, uRV if $u \perp v$ ($\perp$ means perpendicular)
  9. Finite set of points in $\mathbb{R}^2$, pRq if $p_x \leq q_x$ and $p_y \leq q_y$

Exercise

- Consider the following relation on $\mathbb{N}$:
  xRy iff $x - y$ is divisible by 5.

1. Prove that R is an equivalence relation
2. What are the equivalence classes of R?

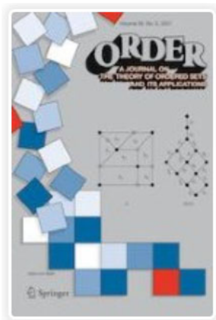
23

Orderings

- A relation R on a set X is a ***partial order*** if R is reflexive, antisymmetric and transitive
- A relation R on a set X is a ***strict order*** if R is irreflexive, antisymmetric and transitive
- Partial orders are commonly denoted by the symbols $\leq$ or $\preccurlyeq$. Once an order $\preccurlyeq$ is given, other order relations such as $\prec$, $\succcurlyeq$, and $\succ$ are readily defined
- Note that if $\preccurlyeq$ is a partial order, then $\prec$ is a strict partial order

24

Springer



Scientific journal devoted to original research on partially ordered sets. It covers all theoretical aspects of the subject and presents applications of order-theoretic methods in the areas of mathematics and computing

Order

A Journal on the Theory of Ordered Sets and its Applications

Total Orders

- Two distinct elements a, b of X are ***comparable*** under R if either aRb or bRa
- A relation R on X is a ***total or linear order*** if it is a partial order and every pair of elements of X are comparable
- A relation R on X is a ***strict total order*** if R is a strict partial order and every pair of elements of X are comparable
- If R is an order (total or not) on a set X, then the pair (X, R) is called a ***partially ordered set*** or ***poset***
- If R is total, we may emphasize this fact by calling (X, R) a ***totally ordered set***

Exercise. Give a non-numeric example of a totally ordered set.

Example

- Let X be a set. Then $(2^X, \subseteq)$ is a poset
- Let $a|b$ denote the relation "a divides b" and Y a set of natural numbers. Then $(Y, |)$ is a poset

Exercise. If $Y = \{1,2,3,4,5,6,7,8,9\}$ what is $(Y, |)$?

Exercise. Let X denote the set of all finite strings of letters. Which of the following relations are partial orders, total orders, strict?

1. xRy if x comes alphabetically no later than y
2. xRy if $|x| \geq |y|$
3. xRy if x contains fewer A's than y

27

DAGs

- The graph representation of a strict poset is a directed graph with no cycles or self-loops, i.e., a *directed acyclic graph* (DAG)
- Is the converse true? In other words, does every DAG correspond to a strict poset?

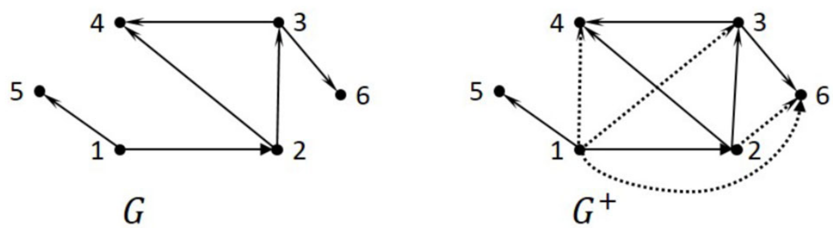
Answer. No, may violate transitivity.

Definition. The ***transitive closure*** of digraph $G = (V, E)$ is the digraph $G^+ = (V, E^+)$, where u is connected to v in G^+ iff there is a directed path from u to v in G. Conversely, if R is the relation corresponding to a digraph G, the ***transitive closure of R*** (denoted R^+) is the relation corresponding to G^+, i.e., the smallest transitive relation containing R.

28

Transitive Closure

- The transitive closure R^+ of a relation R is the relation obtained by adding to R any necessary pairs to satisfy transitivity



29

Example

- Consider the problem of scheduling a project consisting of a set of interdependent activities $A_1, \dots, A_n$
  - $A_i \preccurlyeq A_j$ if A_i needs to be done no later than A_j

Exercise.

- Is $(A, \preccurlyeq)$ a partial order?
- Is $(A, \preccurlyeq)$ total?
- If so, what is the partial order for the "project" of getting out in the morning?

	Activity
1	Choose clothes
2	Dress
3	Eat breakfast
4	Leave house
5	Make coffee
6	Make toast
7	Pour juice
8	Shower
9	Wake up

30

Implications to Computing

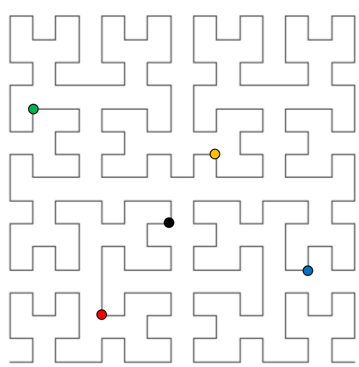
- The relation $\leq$ on $\mathbb{N}$ possesses characteristics that allow us to sort efficiently an array of numbers using algorithms such as QuickSort and MergeSort.
  - What are those characteristics?
  - If we understand these characteristic perhaps we can apply them to sort sets from other domains, such as sets of words, sets of points on the plane, sets of line segments intersecting a line, etc.

31

31

Exercise

- Provide a useful interpretation for $(\mathbb{N}^2, \preccurlyeq)$

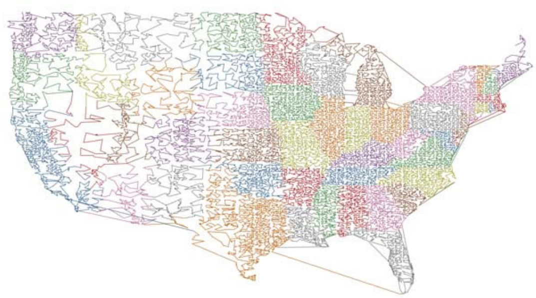


32

32

A Total order on $\mathbb{N}^2$

- How does a politician visit all zip codes while minimizing distance traveled?

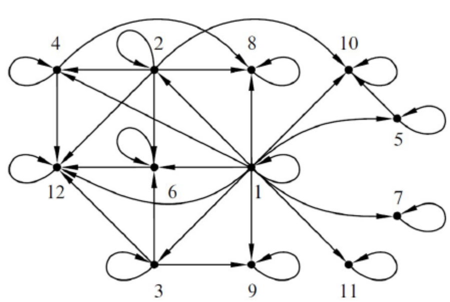


33

33

Poset Representation

- The standard graphical representation of a poset using directed graphs may be difficult to read



34

34

Poset Representation…

- We can omit from the diagram all self-loops and all arrows that can be reconstructed from transitivity (i.e., the transitive closure)
- Only the "immediate predecessors" remain

Definition. Let $(X, \preccurlyeq)$ be an ordered set. We say that $x \in X$ is an ***immediate predecessor*** of $y \in X$, denoted by $x \prec^\star y$, if:

1. $x \prec y$
2. There is no element $t \in X$, such that $x \prec t \prec y$

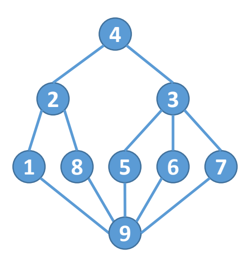
We call y an ***immediate successor*** of x.

Hasse Diagrams

- A ***Hasse diagram*** is a graph of the immediate predecessor relation
- We don't need to draw the direction of the arrows by adopting the convention that arrows are directed upwards
- We denote this relation by $\prec^\star$

- Hasse diagram for $(\{1,2,3,4\}, \leq)$:

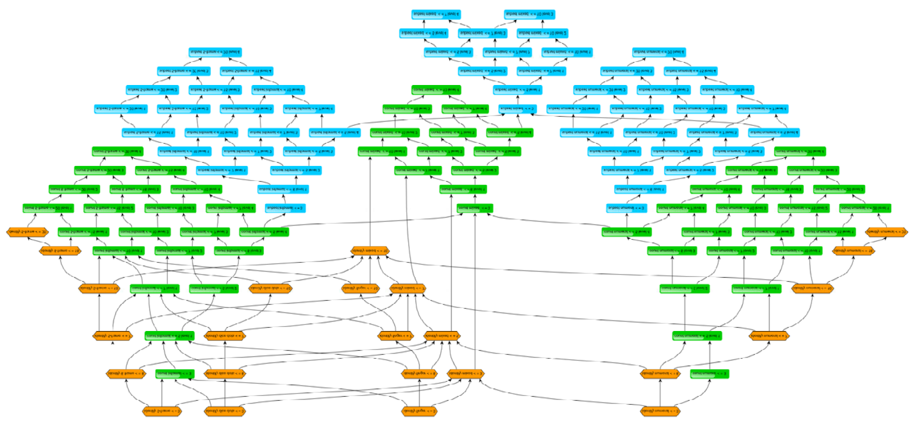
- For getting out in the morning:



	Activity
1	Choose clothes
2	Dress
3	Eat breakfast
4	Leave house
5	Make coffee
6	Make toast
7	Pour juice
8	Shower
9	Wake up

Example: Set of Activities

- Several hundred learning activities organized as a poset

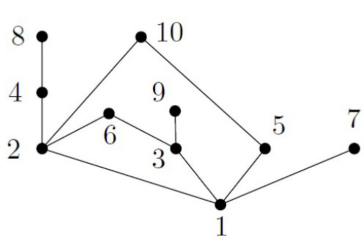


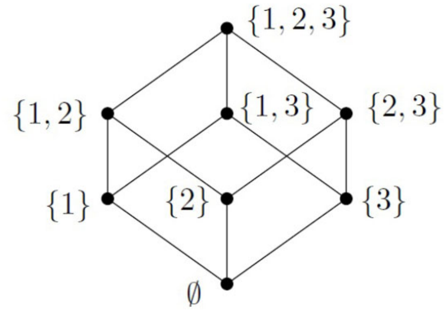
37

37

Exercise

- What do the following Hasse diagrams describe?





38

38

Extreme Elements

Definition. Let $(X, \preccurlyeq)$ be a poset. An element $a \in X$ is called ***minimal*** (resp. ***maximal***) if there is no $b \in X$ such that $b \prec a$ (resp. $a \prec b$).

Theorem. Every finite poset $(X, \preccurlyeq)$ has at least one minimal element.

Proof. Choose an $x \in X$ such that the set $L_x = \{y \in X : y \preccurlyeq x\}$ has the smallest number of elements. If $|L_x| = 1$ we are done and x is minimal; otherwise, the is $y \in L_x$, $y \neq x$. But then $|L_y| < |L_x|$, a contradiction ∎

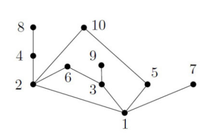
Exercise. Describe and analyze an algorithm to find a minimal element of $(X, \preccurlyeq)$. *Hint*. This can be done very efficiently by choosing the right data structure.

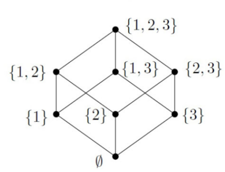
39

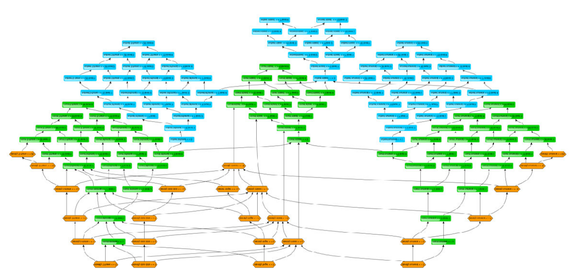
Minimal vs. Minimum

Definition. Let $(X, \preccurlyeq)$ be a poset. An element $a \in X$ is a *minimum* (resp. *maximum*) element of $(X, \preccurlyeq)$ if for every $x \in X$ we have $a \preccurlyeq x$ (resp. $x \preccurlyeq a$).
Note. The term smallest (resp. largest) is often used instead of minimum (resp. maximum).

Exercise. Provide an example of a poset that has a minimal but no minimum element.







40

Down-Sets and Up-Sets

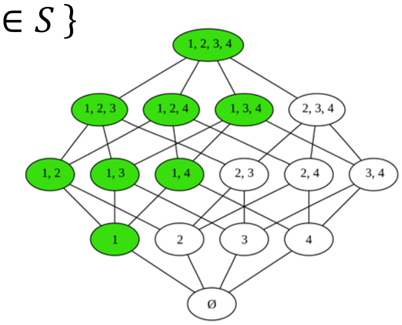
- Let $P = (X, \preccurlyeq)$ be a poset and $S \subseteq X$.
- The ***up-set*** of S (also called *upper set* of S), denoted $\uparrow S$, is defined as

$$\uparrow S = \{x \in X : s \preccurlyeq x \text{ for some } s \in S\}$$

- Similarly, the ***down-set*** (or *lower set*) of S, denoted $\downarrow S$, is defined as

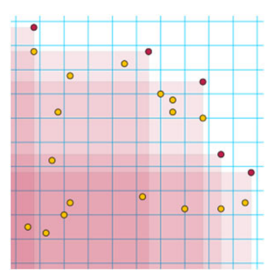
$$\downarrow S = \{x \in X : x \preccurlyeq s \text{ for some } s \in S\}$$

Example. $(\{1,2,3,4\}, \subseteq)$. $\uparrow\{1\}$ is shown in green and $\downarrow\{2,3,4\}$ is shown in white



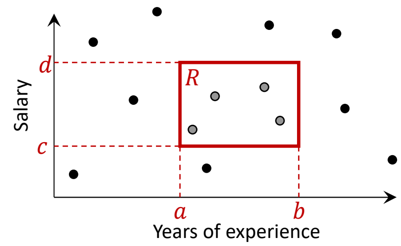
Exercise

- Consider a finite set $P = \{p_1, \ldots, p_n\}$ of points in the plane, with $p_i = (x_i, y_i)$ and define a poset as follows: $p_i \preccurlyeq p_j$ iff $x_i \leq x_j$ and $y_i \leq y_j$
- Describe a practical application of this poset
- Describe an efficient algorithm to find all maximal elements. What do maximal elements represent?



Range Counting

- Given n points in the plane, how many lie inside an upright rectangle $R(a, b, c, d)$?
  - Common query in databases



- Let $D(a, b)$ denote the size of the down set of (a, b). Then,
  $\#R(a, b, c, d) = D(a, b) - D(a, d) - D(b, c) + D(a, c)$

Exercise. Design an efficient algorithm to compute $\#R(a, b, c, d)$. *Hint*. This is easy with the right choice of data structure

Linear Extensions

- Can we sort the elements of a poset in a manner consistent with its partial order?

Theorem. Let $(X, \preccurlyeq)$ be a finite poset. Then there exists a linear ordering $\leq$ on X such that $x \preccurlyeq y$ implies $x \leq y$.

Proof (by induction on $|X|$). The case $|X| = 1$ is trivially true; Consider now the case $|X| > 1$ and let x_0 be minimal. Define $X' \coloneqq X - \{x_0\}$ and $\preccurlyeq' \coloneqq \preccurlyeq$ restricted to X'. Then, by the inductive hypothesis $(X', \preccurlyeq')$ admits a linear ordering $\leq'$.

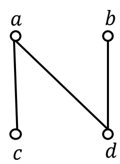
We now define $\leq$ on X: $\begin{cases} x_0 \leq y & \text{for every } y \in X \\ x \leq y & \text{whenever } x \leq' y \end{cases}$

Clearly, $x \preccurlyeq y$ implies $x \leq y$ as required. ∎

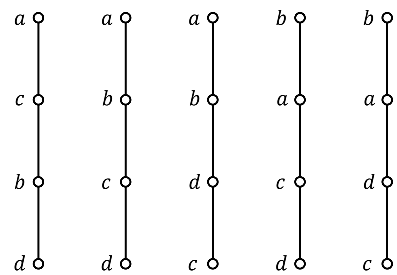
Definition. A linear ordering of a poset P is called a ***linear extension*** of P.

Exercise

- How many linear extensions does the poset $(X, \preccurlyeq)$, shown below, have?
- List all linear extensions of $(X, \preccurlyeq)$



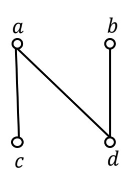
$(X, \preccurlyeq)$



L_1 L_2 L_3 L_4 L_5

Exercise

- Show that the intersection of two linear orders is a poset
- What is the intersection of the five linear extensions of the poset $(X, \preccurlyeq)$ in the previous slide (with $\prec^{\star} = \{(c, a), (d, b), (d, a)\}$)
- What is the intersection of all the linear extensions of a poset



$(X, \preccurlyeq)$

Topological Sorting

- Let $(X, \preccurlyeq)$ denote a poset with $|X| = n$ and $|\prec^*| = m$. An algorithm that computes a linear extension of $(X, \preccurlyeq)$ is called a ***topological sort***
- Below is an algorithm to topologically sort $(X, \preccurlyeq)$
- What is the running time of your algorithm?

EXTEND$(X, \prec^\star)$
Store X in a set S
Store the pairs $\{(x, y) : x \prec^\star y\}$ in a linked list
while S is not empty
    do Find a minimal element $z \in S$.
        Make z the next element in the total order.
        Delete z from S and all pairs of the form (z, y) from $\prec^\star$.

A Fast Linear Extension Algorithm

- We can speed up the linear extension algorithm by associating with each $x \in X$: a list of immediate predecessors and a list of immediate successors of x

FASTEXTEND$(X, \prec^\star)$
For each $x \in X$ use $\prec^\star$ to build a list of immediate predecessors
For each $x \in X$ use $\prec^\star$ to build a list of immediate successors
Initialize an empty queue $\mathcal{Q}$
Enqueue each x whose immediate predecessor list is empty
while $\mathcal{Q}$ is not empty
    do Dequeue the next element z and add it to the total order
        foreach y in z's successor list
            do Delete y from z's immediate successors
                Delete z from y's immediate predecessors
                if y's predecessor list becomes empty
                    then Enqueue y in $\mathcal{Q}$

Chains and Antichains

Definition. Let $P = (X, \preccurlyeq)$ be a finite poset. A subset $A \subseteq X$ is said to be ***independent*** or an ***antichain*** if $x \preccurlyeq y$ does not hold for any pair of distinct $x, y \in A$.

- Each pair $x, y \in A$ is said to be ***incomparable***

Definition. Let $P = (X, \preccurlyeq)$ be a finite poset. A subset $A \subseteq X$ is a ***chain*** if $x \preccurlyeq y$ or $y \preccurlyeq x$ holds for every pair of distinct $x, y \in A$.

- In a sense, chains contain the most order information and antichains contain the least
- Chains can be sorted in only one way while any permutation of its elements is a valid linear extension for an antichain

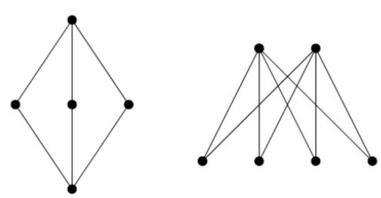
49

How much order does a poset have?

Two parameters of $(X, \preccurlyeq)$ give us an idea of the "amount of order" present in the relation:

- The ***width*** of $(X, \preccurlyeq)$, denoted $w(X, \preccurlyeq)$, is the size of a largest independent set in $(X, \preccurlyeq)$.
- The ***height*** of $(X, \preccurlyeq)$, denoted $h(X, \preccurlyeq)$, is the length of the longest chain in $(X, \preccurlyeq)$.

Example. $w(P_1) = 3$,
$h(P_1) = 3$
$w(P_2) = 4$,
$h(P_2) = 2$



P_1 P_2

50

Large Implies Tall or Wide

Theorem. Every finite poset $P = (X, \preccurlyeq)$ satisfies
$$w(P) \cdot h(P) \geq |X|$$

Proof. First, define subsets $X_1, X_2, \dots, X_t$ inductively:

Base case. X_1 is the set of minimal elements of X.

Inductive step. Once $X_1, \dots, X_k$ have been defined, consider the set $Y_k = X - \bigcup_{i=1}^{k} X_i$ of elements of X not in any of the subsets so far. If Y_k is empty, we are done with $t = k$; otherwise, let $\preccurlyeq'$ represent $\preccurlyeq$ restricted to Y_k and let X_{k+1} be the set of minimal elements in Y_k.

51

51

Proof…

The subsets $X_1, X_2, \dots, X_t$ constitute a partition of X. Furthermore, from above, each X_i is an independent set. We construct a chain $\langle x_1, x_2, \dots, x_t \rangle$ inductively as follows:

Base case. Choose an arbitrary element $x_t \in X_t$.

Inductive step. Suppose x_{i+1} has been chosen. Since $x_{i+1} \notin X_i$, there must be $x_i \in X_i$ with $x_i \preccurlyeq x_{i+1}$.
Since we have a chain of length t then $h(P) \geq t$. Since the X_i's constitute a partition, at least one of them must have $|X|/t$ elements. Since each is an independent set, $w(X) \geq |X|/t$. It follows that

$$w(P) \cdot h(P) \geq |X| \quad \blacksquare$$

52

52

An Application: Scheduling Tasks in Parallel

- If the items of a poset are tasks and the partial order precedence constraints, topological sorting gives an execution schedule consistent with the constraints
- How do you schedule the tasks if you have enough processors? (assume all tasks take unit time)
  - Partition the task into the sets $X_1, X_2, \dots, X_t$ in the previous theorem (relating height and width)
    - A task $a \in X_i$ if the longest chain ending in a has length i
  - Execute the X_i in ascending order, and all the tasks in each X_i in parallel
  - The total (parallel) execution time is t

53

53

Example (Paul Erdős)

Claim. Every sequence $\langle a_1, a_2, \dots, a_{n^2+1} \rangle$ of numbers contains an increasing or decreasing subsequence of length $n+1$

Proof. Suppose there's no *increasing* subsequence of length $n+1$.

We show there must be a *decreasing* sequence of length $n+1$.

For each $k = 1, \dots, n^2+1$, let m_k be the length of the longest subsequence that starts with a_k. Clearly, $1 \le m_k \le n$.

From the *generalized pigeonhole principle*, $n+1$ of the values $m_1, \dots, m_{n^2+1}$ are equal, i.e., $m_{k_1} = m_{k_2} = \cdots = m_{k_{n+1}}$, where $1 \le k_1 < k_2 < \cdots < k_{n+1} \le n^2+1$.

If $a_{k_i} \le a_{k_{i+1}}$ then $m_{k_i} > m_{k_{i+1}}$, not possible. Hence, $a_{k_i} > a_{k_{i+1}}$

Conclusion: either there is an increasing subsequence of length $n+1$ or a decreasing subsequence $a_{k_1} > a_{k_2} > \cdots > a_{k_{n+1}}$

54

54

Erdős-Szekeres Theorem

***Theorem*.** A sequence $(x_1, x_2, \dots, x_{n^2+1})$ of distinct real numbers has an ascending or descending subsequence of length $n+1$.

Proof. Let $X = \{1,2,\dots,n^2+1\}$. Define an ordering $\preccurlyeq$ on a permutation of X as:

$$i \preccurlyeq j \text{ iff } i \leq j \text{ \underline{and} } x_i \leq x_j$$

Since $w(X,\preccurlyeq) \cdot h(X,\preccurlyeq) \geq n^2+1$ then $h(X,\preccurlyeq) > n$ or $w(X,\preccurlyeq) > n$

If $h(X,\preccurlyeq) > n$ then there is an ascending sequence of length $> n$

If $w(X,\preccurlyeq) > n$ then there is a descending sequence of length $> n$ (why?)

55

55

Poset Dimension

- Is it possible to reconstruct a poset $P = (X,\preccurlyeq)$ from a list $\mathcal{E}(P)$ of its linear extensions (total orders)?
  - Yes, as $P = \bigcap_{L\in\mathcal{E}(P)} L$ (why?)
- A set $\mathcal{R}$ of linear extensions of poset P is a ***realizer*** of P if $P = \bigcap_{L\in\mathcal{R}} L$
- The ***dimension*** of a poset P is the size of a smallest realizer of P, i.e., the minimum number of linear extensions whose intersection yields P
- Realizers provide an alternative representation of a poset. Its dimension lower bounds the # of required total orders

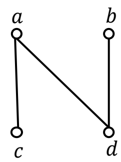
Theorem (Hiraguchi). Let d be the dimension of a poset $P = (X,\preccurlyeq)$. Then, if $|X| \geq 4$, $d \leq \min(w(P), |X|/2)$.

56

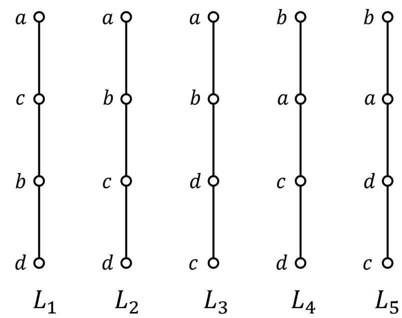
56

Example

- The poset $(X, \preccurlyeq)$ below has 5 linear extensions. How many are needed, i.e., what is the dimension of $(X, \preccurlyeq)$?
- $\dim(X, \preccurlyeq) \geq 2$, why?



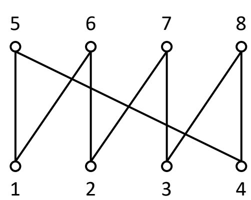
$(X, \preccurlyeq)$



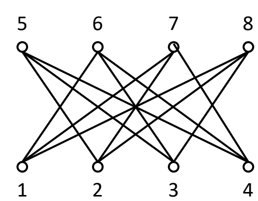
- $\mathcal{R} = \{L_1, L_5\}$ is a realizer $\Rightarrow \dim(X, \preccurlyeq) = 2$

Exercise

- Find small realizers of the following posets. Can you infer the dimension of each poset?



$(X, \leq)$

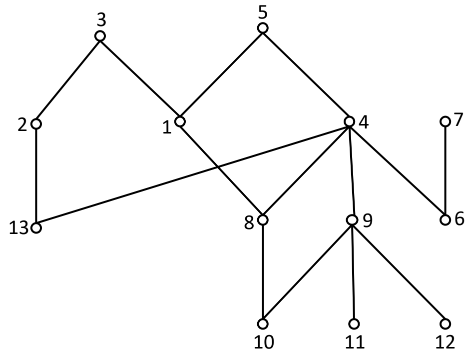


$(X, \preccurlyeq)$

$X = \{1,2,3,4,5,6,7,8\}$

Exercise

- Show that $\dim(P) \leq 3$ for the poset P below:

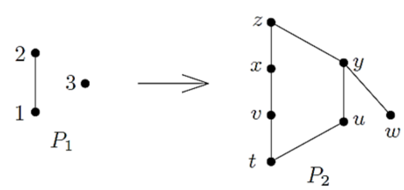


59

Embedding of Posets

Definition. Let $(X, \preccurlyeq)$ and $(X', \preccurlyeq')$ be finite posets. A mapping $f: X \to X'$ is an ***embedding*** of $(X, \preccurlyeq)$ into $(X', \preccurlyeq')$ if the following conditions hold:

1. f is one-to-one
2. $x \preccurlyeq y$ iff $f(x) \preccurlyeq' f(y)$

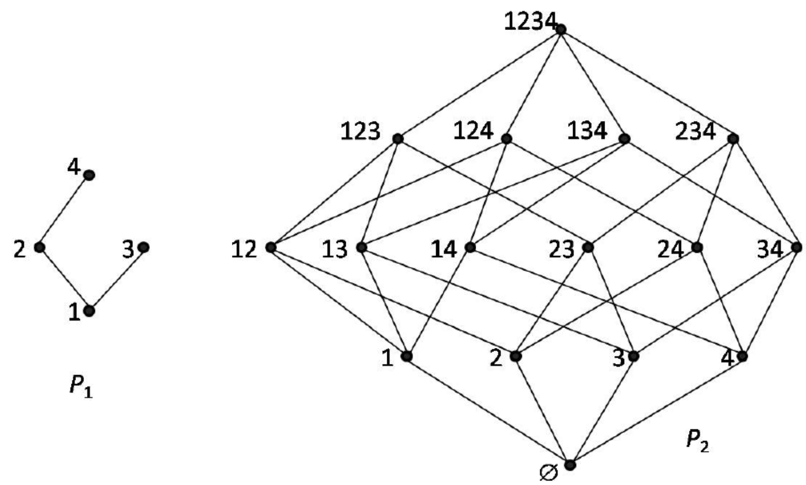


- Is there a kind of “universal poset” that encodes all posets of a given size by containing embeddings of them?

60

Exercise

- Find an embedding of P_1 in P_2



61

61

A universal poset

Theorem. Every poset $(X, \preccurlyeq)$ has an embedding into the poset $(2^X, \subseteq)$.

Proof. Define the mapping $f : X \to 2^X$ by $f(x) = \{y \in X : y \preccurlyeq x\}$.
We now show this is an embedding.

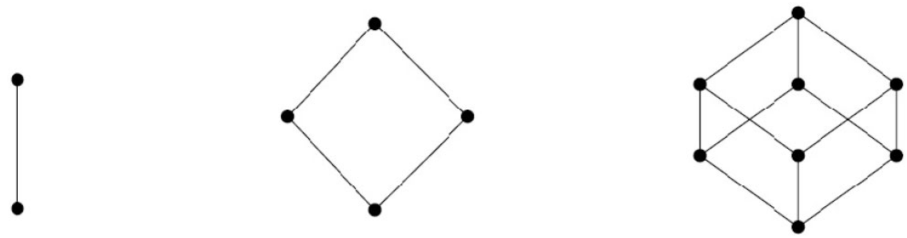
1. f is injective as, otherwise, $f(x) = f(y)$ for two distinct elements x and y. Since $x \in f(x) = f(y)$ and $y \in f(y) = f(x)$, then $x \preccurlyeq y$ and $y \preccurlyeq x$, violating antisymmetry
2. ($\Rightarrow$) Assume $x \preccurlyeq y$ and let $z \in f(x)$. Then $z \preccurlyeq x$ and transitivity implies $z \preccurlyeq y \Rightarrow z \in f(y)$, i.e., $f(x) \subseteq f(y)$
3. ($\Leftarrow$) Assume $f(x) \subseteq f(y)$. Since $x \in f(x)$ then $x \in f(y)$, i.e., $x \preccurlyeq y$

62

62

Example

- From the above we conclude that the posets $B_k = \left(2^{\{1,\ldots,k\}}, \subseteq\right)$, $k = 1, 2, 3, \ldots$ contain "copies" of all possible orderings of finite sets



$\mathcal{B}_1$ $\mathcal{B}_2$ $\mathcal{B}_3$

63

63