## Sets & Relations

Basics of Sets



sets & Relations

#### Relational Database

×	У	Likes(x,y)	
Alice	Alice	TRUE	
	Jabberwock	FALSE	
	Flamingo	TRUE	
Jabberwock	Alice	FALSE	
	Jabberwock	TRUE	
	Flamingo	FALSE	
Flamingo	Alice	FALSE	
	Jabberwock	FALSE	
	Flamingo	TRUE	

Relational DB Table

Likes		
X	y	
Alice	Alice	
Alice	Flamingo	
Jabberwock	Jabberwock	
Flamingo	Flamingo	

- Queries to the DB are set/logical operations
  - SELECT x

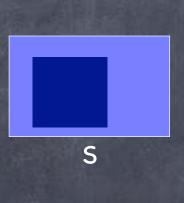
    FROM Likes

    WHERE y='Alice' OR y='Flamingo'

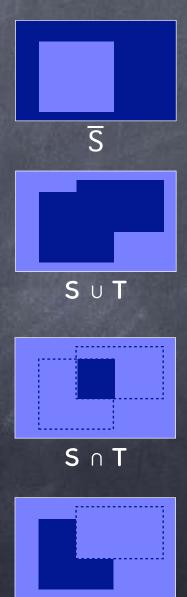
## Sets: Basics

- Unordered collection of "elements"
  - $oe.g.: \mathbb{Z}, \mathbb{R}$  (infinite sets), oeta (empty set), oeta (1, 2, 5), ...
- Will always be given an implicit or explicit universe (universal set) from which the elements come
  - (Aside: In developing the foundations of mathematics, often one starts from "scratch", using only set theory to create the elements themselves)
- Set membership: e.g. 0.5 ∈  $\mathbb{R}$ , 0.5 ∉  $\mathbb{Z}$ ,  $\emptyset$  ∉  $\mathbb{Z}$
- Set inclusion: e.g. Z, ⊆  $\mathbb{R}$ ,  $\emptyset$  ⊆ Z
- Set operations: complement, union, intersection, difference

# Set Operations







#### Sets as Predicates

×	Winged(x)	Flies(x)	Pink(x)	inClub(x)
Alice	FALSE	FALSE	FALSE	TRUE
Jabberwock	TRUE	TRUE	FALSE	FALSE
Flamingo	TRUE	TRUE	TRUE	TRUE

- Given a predicate can define the set of elements for which it holds
  - WingedSet = { x | Winged(x) } = {J'wock, Flamingo}
  - FliesSet = { x | Flies(x) } = {J'wock, Flamingo}
  - PinkSet = { x | Pink(x) } = {Flamingo}
- © Conversely, given a set, can define a membership predicate for it e.g. given set  $Club = \{Alice, Flamingo\}$ . Then, define predicate inClub(x) s.t.  $inClub(x) = True \ iff \ x \in Club$

## Set Operations

Unary operator

Binary operators

Associative

S complement

Symbol: S

 $in\overline{S}(x) = \neg inS(x)$ 

S union T

Symbol: SUT

inS∪T(x)

= inS(x)  $\vee$  inT(x)

S <u>intersection</u> T

Symbol: S ∩ T

inS∩T(x)

= inS(x)  $\wedge$  inT(x)

S difference T

Symbol: S - T

(Alternately: S\T)

|inS-T(x)|

 $\equiv$  inS(x)  $\land$   $\neg$ inT(x)

 $\equiv$  inS(x)  $\leftrightarrow$  inT(x))

 $S-T = S \cap \overline{T}$ 

S symmetric diff. T

Symbol: S \( D \) T

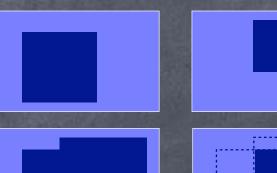
 $inS\Delta T(x)$ 

 $\equiv$  inS(x)  $\oplus$  inT(x)

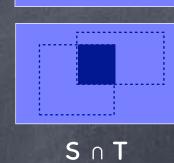
Note: Notation inS(x) used only to explicate the connection with predicate logic. Will always write  $x \in S$  later.

# De Morgan's Laws

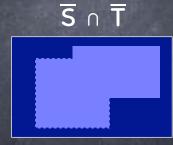
- SUT = S∩T
  - - $\equiv \neg(x \in S \lor x \in T) \equiv \neg(x \in S) \land \neg(x \in T)$
    - $\equiv x \in \overline{S} \land x \in \overline{T} \equiv x \in \overline{S} \cap \overline{T}$



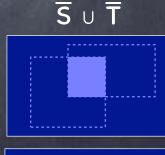
S U T

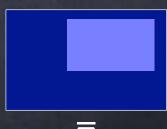


- $S \cap T = \overline{S} \cup \overline{T}$ 
  - - $\equiv \neg(x \in S \land x \in T) \equiv \neg(x \in S) \lor \neg(x \in T)$
    - $\equiv x \in \overline{S} \lor x \in \overline{T} \equiv x \in \overline{S} \cup \overline{T}$









## Distributivity

```
R \cap (S \cup T) = (R \cap S) \cup (R \cap T)
          x \in R \cap (S \cup T) = 
               = x \in R \land (x \in S \lor x \in T) = (x \in R \land x \in S) \lor (x \in R \land x \in T)
               \equiv \mathbf{x} \in (\mathbf{R} \cap \mathbf{S}) \cup (\mathbf{R} \cap \mathbf{T})
  x \in R \cup (S \cap T) = 
               \equiv x \in R \lor (x \in S \land x \in T) \equiv (x \in R \lor x \in S) \land (x \in R \lor x \in T)
               \equiv x \in (R \cup S) \cap (R \cup T)
```

#### Set Inclusion

×	Winged(x)	Flies(x)	Pink(x)
Alice	FALSE	FALSE	FALSE
Jabberwock	TRUE	TRUE	FALSE
Flamingo	TRUE	TRUE	TRUE

- PinkSet ⊆ FliesSet = WingedSet
- S ⊆ T same as the proposition  $\forall x x \in S \rightarrow x \in T$
- **③** S ⊇ T same as the proposition  $\forall x \ x \in S \leftarrow x \in T$
- S = T same as the proposition ∀x x∈S ↔ x∈T

#### Set Inclusion

- S ⊆ T same as the proposition  $\forall x \quad x \in S \rightarrow x \in T$
- If S =  $\emptyset$ , and T any arbitrary set, S ⊆ T
  - $\bullet$   $\forall x$ , vacuously we have  $x \in S \rightarrow x \in T$
- $\odot$  If  $S\subseteq T$  and  $T\subseteq R$ , then  $S\subseteq R$

If no such x, already done

- **②** Consider arbitrary  $x \in S$ . Since  $S \subseteq T$ ,  $x \in T$ . Then since  $T \subseteq R$ ,  $x \in R$ .
- $lackbox{0} S \subseteq T \longleftrightarrow \overline{T} \subseteq \overline{S}$

$$\equiv \ \forall x \ \underline{x \in \overline{T}} \to \underline{x \in \overline{S}}$$

# Proving Set Equality

- To prove S = T, show  $S \subseteq T$  and  $T \subseteq S$
- ø e.g., L(a,b) = { x : ∃u,v ∈  $\mathbb{Z}$  x=au+bv }

  M(a,b) = { x : ( gcd(a,b) | x ) }
- © [Recall] Theorem: L(a,b) = M(a,b)
- Proof in two parts:
  - lacksquare L(a,b)  $\subseteq$  M(a,b) : i.e.,  $\forall x \in \mathbb{Z}$   $x \in L(a,b) \rightarrow x \in M(a,b)$

Let x=au+bv.  $gla, glb \Rightarrow glx$ 

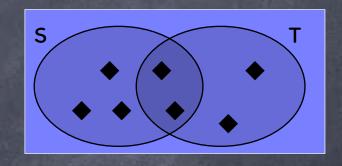
First show that
g∈L(a,b)

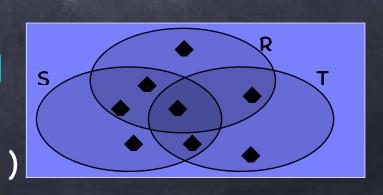
(as the smallest +ve
element in L(a,b))

Let x=ng. But g=au+bv ⇒ x=au'+bv'

## Inclusion-Exclusion

- S| + |T| counts every element that is in S or in T
  - But it double counts the number of elements that are in both:
     i.e., elements in S∩T
- So, |S|+|T| = |S∪T| + |S∩T|
- Or, |S∪T| = |S| + |T| |S∩T|





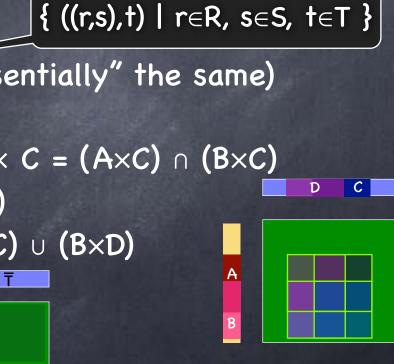
#### Cartesian Product



 $\circ$  Not the same as  $(R \times S) \times T$  (but "essentially" the same)

S

$$\circ$$
 ( $\overline{S} \times \overline{T}$ )  $\cup$  ( $\overline{S} \times T$ )  $\cup$  ( $S \times \overline{T}$ )



# Sets & Relations

Relations



## Relations: Basics

More commonly written as: x Likes y,  $x \sqsubseteq y$ ,  $x \ge y$ ,  $x \sim y$ , xLy, ...

- Informal y, a relation is specified as what is related to what
- Formally, a predicate over the domain SxS
  - e.g. Likes(x,y)
- Alternately, a subset of SxS, namely the pairs for which the relation holds
  - Dikes = { /Alice, Alice),

Homogeneous and binary (the default notion for us) (Alice, Flamingo),
(J'wock,J'wock),
(Flamingo,Flamingo) }

x,y	Likes(x,y)	
Alice, Alice	TRUE	
Alice, Jabberwock	FALSE	
Alice, Flamingo	TRUE	
Jabberwock, Alice	FALSE	
Jabberwock, Jabberwock	TRUE	
Jabberwock, Flamingo	FALSE	
Flamingo, Alice	FALSE	
Flamingo, Jabberwock	FALSE	
Flamingo, Flamingo	TRUE	

## Many ways to look at it!

R ⊆ S × S

a set of

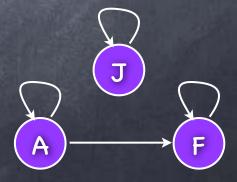
ordered-pairs
{ (a,b) | a □ b }

{ (A,A), (A,F),
 (J,J), (F,F) }

Boolean matrix, M<sub>a,b</sub> = T iff a□b



(directed) graph



## Operations on Relations

- Since a relation is a set, namely R ⊆ S×S, all set operations extend to relations
  - © Complement (with the universe being S×S), Union, Intersection, Symmetric Difference
- © Converse (a.k.a. Transpose)

Composition

"Boolean matrix multiplication"
$$(M \cap M')_{xy} = \bigvee_{w} (M_{xw} \wedge M'_{wy})$$

(Ir)Reflexive Relations

- Reflexive (e.g. Knows, ≤)
  - The kind of relationship that everyone has with All self-loops

themselves

All of diagonal included

None of it

- Irreflexive (e.g. Gave birth to, ≠)
  - The kind that nobody has with themselves
- No self-loops

- Neither (e.g. is a prime factor of)
  - Some, but not all, have this relationship with themselves

## (Anti)Symmetric Relations

- Symmetric (e.g. sits next to)
  - The relationship is reciprocated symmetric matrix
- Anti-symmetric (e.g. parent of, prime factor of, ⊆)
  - No reciprocation (except possibly with self)
- Neither (e.g. likes)
  - Reciprocated in some pairs (with distinct members) and only one-way in other pairs
    Some bidirectional,
- Both (e.g., =)
  - Each one related only to self (if at all)

self-loops & bidirectional edges only

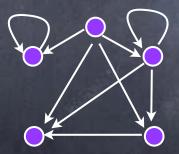
no bidirectional edges

no edges except self-loops

some unidirectional

#### Transitive Relations

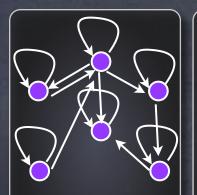
- Transitive (e.g., Ancestor of, subset of, divides, ≤)
  - if a is related to b and b is related to c, then a is related to c
- $\odot$  R is transitive  $\leftrightarrow$  R $\circ$ R  $\subseteq$  R



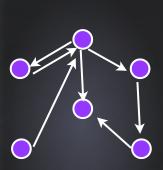
 $\leftrightarrow \forall k > 1 \ R^k \subseteq R$ 

if there is a "path" from a to z, then there is edge (a,z)

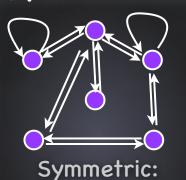
Intransitive: Not transitive



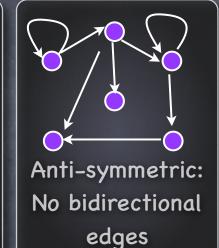
Reflexive: All self-loops

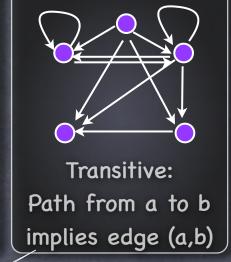


Irreflexive:
No self-loops



Only self-loops & bidirectional edges



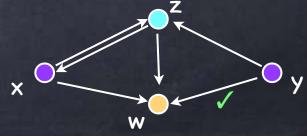


The complete relation  $R = S \times S$  is reflexive, symmetric and transitive

Reflexive closure of R: Minimal relation  $R' \supseteq R$  s.t. R' is reflexive Symmetric closure of R: Minimal relation  $R' \supseteq R$  s.t. R' is symmetric Transitive closure of R: Minimal relation  $R' \supseteq R$  s.t. R' is transitive Each of these is unique [Why?]

## Equivalence Relation

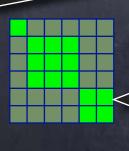
- A relation that is reflexive, symmetric and transitive
  - e.g. is a relative, has the same last digit, is congruent mod 7, ...
- Equivalence class of x:  $Eq(x) = \{y | x \sim y\}$ .
- Every element is in its own equivalence class (by reflexivity)
- Oclaim: If Eq(x)  $\cap$  Eq(y)  $\neq$  Ø, then Eq(x) = Eq(y).
  - The second in the property of the proper
  - Consider an arbitrary  $w \in Eq(x)$ : i.e.,  $x \sim w$ .
  - **3** By symmetry,  $z\sim x$ . Then, by transitivity,  $z\sim w$ . Then,  $y\sim w$ .
  - Thus,  $w \in Eq(y)$ . i.e.,  $Eq(x) \subseteq Eq(y)$ .



## Equivalence Relation

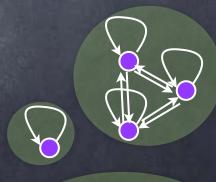
- A relation that is reflexive, symmetric and transitive
  - e.g. is a relative, has the same last digit, is congruent mod 7, ...
- Equivalence class of x: Eq(x)  $\triangleq$  {y|x ~ y}.
- Every element is in its own equivalence class (by reflexivity)
- Oclaim: If Eq(x)  $\cap$  Eq(y)  $\neq$  Ø, then Eq(x) = Eq(y).
- The equivalence classes partition the domain

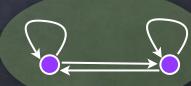
 $P_1,...,P_t \subseteq S$  s.t.  $P_1 \cup ... \cup P_t = S$  $P_i \cap P_j = \emptyset$ 

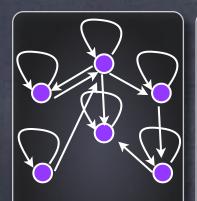


Square blocks along the diagonal, after sorting the elements by equivalence class

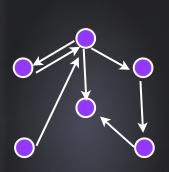
"Cliques" for each class



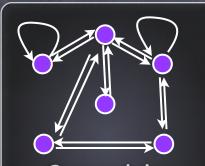




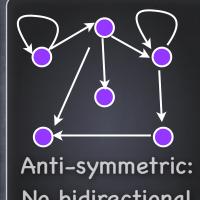
Reflexive: All self-loops



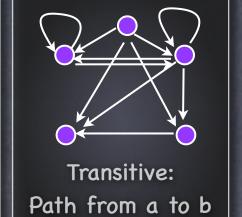
Irreflexive: No self-loops



Symmetric:
Only self-loops &
bidirectional edges



Anti-symmetric: No bidirectional edges



implies edge (a,b)

Equivalence:
Cliques, disconnected

Reflexive closure of R: Small from each other R s.t. R' is reflexive

Symmetric closure of R: Smallest relation  $R' \supseteq R$  s.t. R' is symmetric

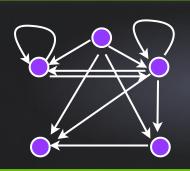
<u>Transitive closure of R:</u> Smallest relation R' ⊇ R s.t. R' is transitive

An equivalence relation R is its own reflexive, symmetric and transitive closure

## Sets & Relations

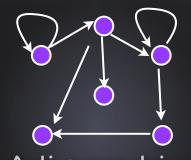
Posets





Transitive: Path from a to b implies edge (a,b)

#### Landscape of Transitive Relations



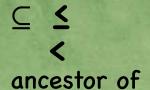
Anti-symmetric: No bidirectional edges



Symmetric: Only self-loops & bidirectional edges

#### Acyclic

Cannot follow a sequence of non-self-loop edges and get back to where you started from



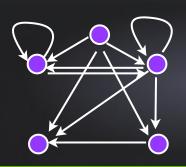
has same last name as

**Anti-Symmetric** 





Symmetric

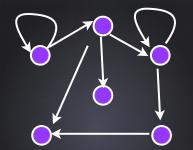


Transitive:
Path from a to b
implies edge (a,b)

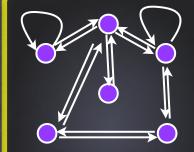
# Landscape of Transitive Relations

Strict

Partial Orders Irreflexive

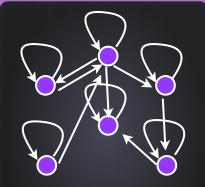


Anti-symmetric: No bidirectional edges

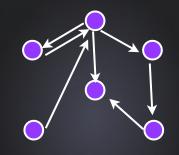


Symmetric:
Only self-loops &
bidirectional edges





Reflexive: All self-loops



Irreflexive:
No self-loops



Reflexive Partial Orders

Anti-Symmetric

has same last name as

Equivalences

**Symmetric** 

# Partial Order

Strict partial order: irreflexive, rather than reflexive

- A transitive, anti-symmetric and reflexive relation
  - ② e.g. ≤ for integers, divides for integers, ⊆ for sets, "containment" for line-segments
- Equivalently, transitive and acyclic (and ir/reflexive) (a pair of bidirectional edges is a "cycle")
  - Order" refers to these properties
- "Partial": not every two elements need be "comparable"
  - i.e., {a,b} s.t. neither a b nor b a
    - $\odot$  e.g., neither  $A \subseteq B$  nor  $B \subseteq A$

## Posets

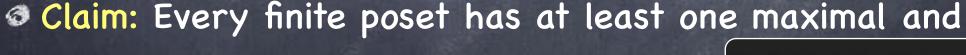
- Partially ordered set (a.k.a Poset)
  - A non-empty set and a partial order over it
  - Denoted like (S, ≤)
- e.g.  $S = \{S_1, S_2, S_3, S_4, S_5\}$  where  $S_1 = \{0, 1, 2, 3\}, S_2 = \{1, 2, 3, 4\}, S_3 = \{1, 2, 3\}, S_4 = \{3, 4\}, \text{ and } S_5 = \{2\}. \text{ Poset } (S, \subseteq)$

#### Check:

- Anti-symmetric
   (no bidirectional edges),
- Transitive,
- Reflexive (all self-loops)
- More generally, (S, ⊆) where S is any set of sets
- Verify: P⊆P; P⊆Q ∧ Q⊆R → P⊆R; P⊆Q ∧ Q⊆P → P=Q
- @ e.g. Divisibility poset: (I+, |)
  - $\odot$  Verify: ala; alb  $\wedge$  blc  $\rightarrow$  alc; alb  $\wedge$  bla  $\rightarrow$  a=b

## Extremal & Extremum

- Maximal & minimal elements of a poset (S, ≤)
  - x∈S is maximal if ∃y∈S-{x} s.t. x≤y
  - $x \in S$  is minimal if  $\exists y \in S \{x\}$  s.t.  $y \leq x$ 
    - Need not exist (e.g., in (ℤ,≤)).
    - Need not be unique when it exists (e.g., divisibility poset restricted to integers > 1)



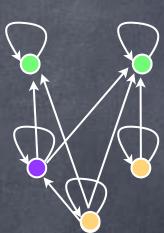
one minimal element

Proof by induction on |S| [Exercise]

 $x \in S$  is the greatest element if  $\forall y \in S, y \leq x$   $x \in S$  is the least element if  $\forall y \in S, x \leq y$ 

Useful in induction proofs about finite posets

Need not exist.
Unique when one exists.



## Other Relations from a Poset

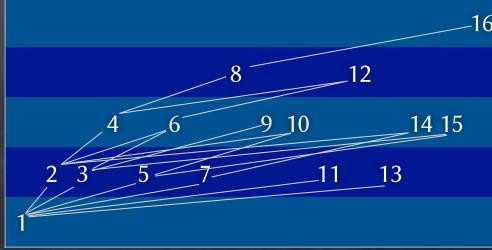
- Consider partial order ≤
- - a < b iff a ≠ b and a ≤ b
    </pre>
- - Well-defined for finite posets: Define  $a \sqsubseteq b$  iff  $a \leqslant b$ and  $∄m ∉ {a,b}$ ,  $a \leqslant m \leqslant b$ . [Prove by induction]

## Running Example

Divisibility poset: (I+, |)

- Claim: | is the transitive closure of the reflexive closure of □ [Verify]
- © Claim: 

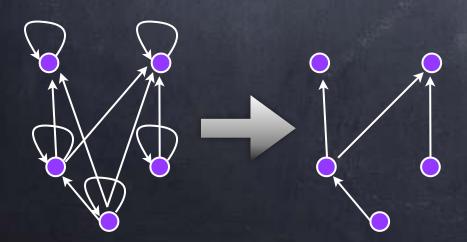
  is the transitive reduction of the reflexive reduction of | [Verify]
  - Note: Divisibility poset has a transitive reduction even though it is infinite

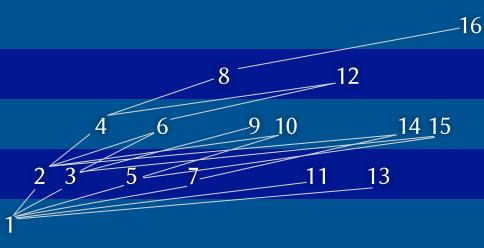


## Hasse Diagram

- For a poset (S, ≤), the transitive reduction of the reflexive reduction of ≤, if it exists, has all the information about the poset
  - Recall: For finite posets, guaranteed to exist

Hasse Diagram: the graph of this relation (with arrowheads implicit)





## Bounding Elements

- Ø Given a poset (S, ≤) and T ⊆ S
- Need not exist.

  Need not be unique when one exists.

Do exist in finite posets

- Maximal element in T :  $x \in T$  s.t.  $\forall y \in T$ ,  $\underline{x \leq y \rightarrow y = x}$ Minimal element in T :  $x \in T$  s.t.  $\forall y \in T$ ,  $\underline{y \leq x \rightarrow y = x}$
- Greatest element in T :  $x \in T$  s.t.  $\forall y \in T$   $y \leq x$ Least element in T :  $x \in T$  s.t.  $\forall y \in T$ ,  $x \leq y$

Need not exist.
Unique when one exists.

- Oupper Bound for T:  $x \in S$  s.t.  $\forall y \in T$ ,  $y \leq x$ Lower Bound for T:  $x \in S$  s.t.  $\forall y \in T$ ,  $x \leq y$
- Least Upper Bound for T: Least in {x| x u.b. for T}
  Greatest Lower Bound for T: Greatest in {x| x l.b. for T}

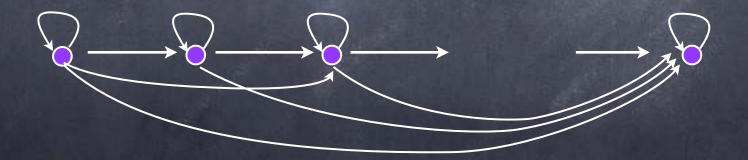
## Running Example

Divisibility poset: (I+, |)

- Lower bound
  - When is c a lower bound for T={a,b}?
    - $\circ$  cla and clb  $\Rightarrow$  c is a common divisor for  $\{a,b\}$
  - Greatest lower bound for {a,b} = gcd(a,b)
- Upper bound
  - $\ensuremath{\mathfrak{G}}$  d is an upper bound for  $\{a,b\} \Rightarrow a|d$ ,  $b|d \Rightarrow d$  a common multiple for  $\{a,b\}$
  - Least upper bound for {a,b} = lcm(a,b)

## Total/Linear Order

- In some posets every two elements are "comparable": for {a,b}, either a b or b a
  - © Can arrange all the elements in a line, with <u>all</u> <u>possible</u> right-pointing edges (plus, all self-loops)



If finite, has <u>unique</u> maximal and <u>unique</u> minimal elements (left and right ends)

#### Order Extension

- $\lozenge$  A poset P'=(S, $\le$ ) is an extension of a poset P=(S, $\le$ ) if  $\forall a,b \in S$ ,  $a \le b \rightarrow a \le b$
- Any finite poset can be extended to a total ordering (this is called topological sorting)
  - Prove by induction on |S|
    - Induction step: Remove a minimal element, extend to a total ordering, reintroduce the removed element as the minimum in the total ordering.
  - For infinite posets? The "Order Extension Principle" is typically taken as an axiom! (Unless an even stronger axiom called the "Axiom of Choice" is used)

## Running Example

#### Divisibility poset: (I+, |)

- The totally ordered set  $(\mathbb{Z}^+, \leq)$ , where  $\leq$  is the standard "less-than-or-equals" relation, is an extension of the divisibility poset
  - $\odot$  Because a|b  $\rightarrow$  a  $\leq$  b
- - $\circ$  For any  $(a,b) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ ,  $a \sqsubseteq b$  iff:
    - @ a=1, or
    - a,b both prime or both composite, and a ≤ b, or
    - a prime and b composite

## Sets & Relations

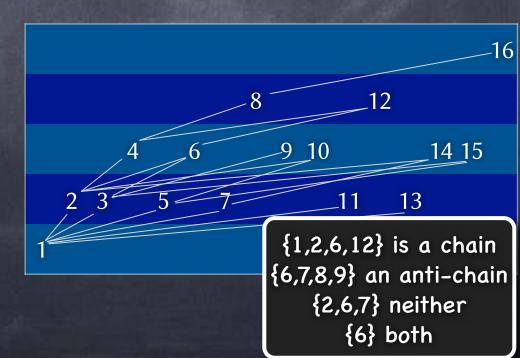
Chains and Anti-Chains



#### Chains & Anti-Chains

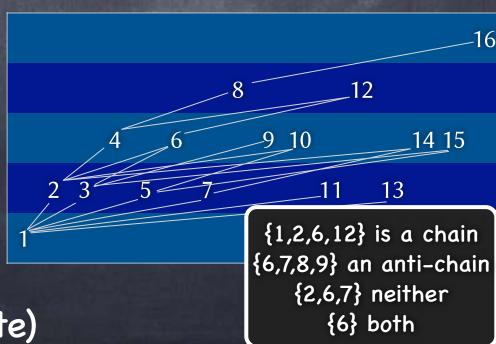
In a poset (S,≤)

- $\odot$  C  $\subseteq$  S is said to be a chain:  $\odot$  A  $\subseteq$  S is an anti-chain if  $\forall a,b \in C$ , either  $a \leq b$  or b≤a
  - if ∀a,b∈A, neither a≤b nor b≤a, unless a=b
- i.e., (C,≤) is a total order : ⊙ (A,≤) is same as (A,=)
- Subset of a chain is a chain. Similarly for anti-chains.
- A singleton set is both a chain and an anti-chain
- For any chain C and antichain A,  $|C \cap A| \leq 1$  (Why?)



## Height in a Poset

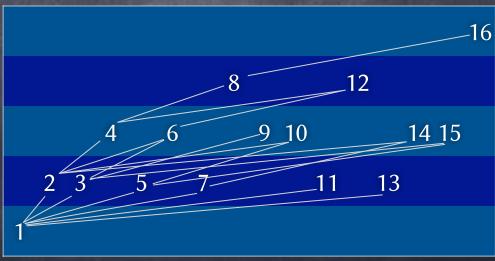
- In a poset  $(S, \le)$ , for any  $a \in S$ , we define Finite if S is finite height(a) = max size of a chain with a as the maximum
  - Note: every a has {a} as such a chain
- © E.g., In ( $\mathbb{Z}^+$ , | ), height(1)=1, height(p)=2 for all primes p. For m=p<sub>1</sub>d<sub>1</sub>·...·p<sub>t</sub>d<sub>t</sub> (p<sub>i</sub> primes), height(m) = 1+ $\Sigma_i$  d<sub>i</sub>
- Height of the poset (S,≤)
  = max { height(a) | a∈S}
  = max { |C| | chain C}
  - Size of the largest chain in the poset
  - Possibly ∞ (only if S infinite)



## Anti-Chains from Height

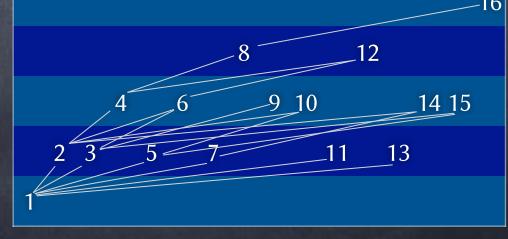
- © Let  $A_h = \{ a \mid height(a) = h \}$
- For every finite h, Ah is an anti-chain (possibly empty)
  - Otherwise, ∃a≠b, a≤b with height(a) = height(b) = h. height(a) = h ⇒ ∃chain C s.t. a=max(C) and |C|=h
    - $\Rightarrow$  b $\notin$ C and C'=C $\cup$ {b} is a chain with b=max(C')

How?  $\Rightarrow$  height(b)  $\geq$  h+1!



## Anti-Chains from Height

- © Let  $A_h = \{ a \mid height(a) = h \}$
- For every finite h, Ah is an anti-chain (possibly empty)
- In a finite poset, since every element has a finite height, every element appears in some A<sub>h</sub>: i.e., A<sub>h</sub>'s partition S
- Mirsky's Theorem: The least number of anti-chains needed to partition S is exactly the size of a largest chain



For chain C⊆S, need ≥ |C| anti-chains to cover C, as |C∩A| ≤ 1 for anti-chain A

## Partitioning with (Anti)Chains

Mirsky's Theorem: The least number of anti-chains needed to partition S is exactly the size of a largest chain

Later

Dilworth's Theorem: The least number of chains needed to partition S is exactly the size of a largest anti-chain

