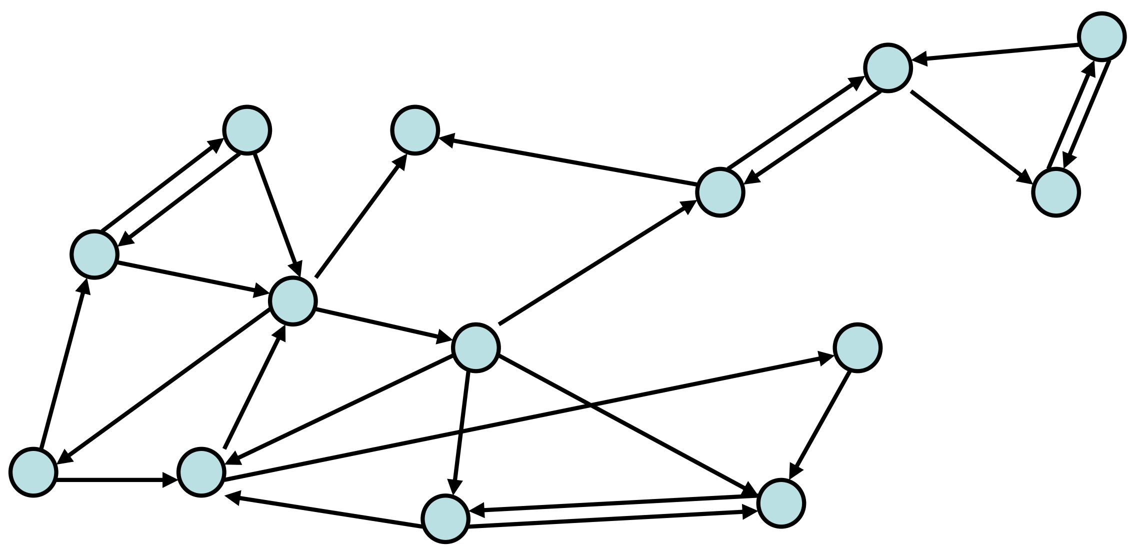
Spring 2015

Lecture 26: Relations and directed graphs



Let A and B be sets.
A binary relation from A to B is a subset of $A \times B$

Let A be a set.
A binary relation on A is a subset of $A \times A$

relations you already know!

$\geq$ on $\mathbb{N}$

That is: $\{(x,y) : x \geq y \text{ and } x, y \in \mathbb{N}\}$

$<$ on $\mathbb{R}$

That is: $\{(x,y) : x < y \text{ and } x, y \in \mathbb{R}\}$

$=$ on Σ^*

That is: $\{(x,y) : x = y \text{ and } x, y \in \Sigma^*\}$

$\subseteq$ on $P(U)$ for universe U

That is: $\{(A,B) : A \subseteq B \text{ and } A, B \in P(U)\}$

$R_1 = \{(a, 1), (a, 2), (b, 1), (b, 3), (c, 3)\}$

$R_2 = \{(x, y) \mid x \equiv y \pmod{5}\}$

$R_3 = \{(c_1, c_2) \mid c_1 \text{ is a prerequisite of } c_2\}$

$R_4 = \{(s, c) \mid \text{student } s \text{ had taken course } c\}$

Let R be a relation on A.

R is reflexive iff $(a,a) \in R$ for every $a \in A$

R is symmetric iff $(a,b) \in R$ implies $(b, a) \in R$

R is antisymmetric iff $(a,b) \in R$ and $a \neq b$ implies $(b,a) \notin R$

R is transitive iff $(a,b) \in R$ and $(b, c) \in R$ implies $(a, c) \in R$

Let R be a relation from A to B.
Let S be a relation from B to C.

The **composition** of R and S, $S \circ R$ is the relation from A to C defined by:

$$S \circ R = \{(a, c) \mid \exists\, b \text{ such that } (a,b) \in R \text{ and } (b,c) \in S\}$$

Intuitively, a pair is in the composition if there is a "connection" from the first to the second.

examples

$(a,b) \in$ Parent iff b is a parent of a

$(a,b) \in$ Sister iff b is a sister of a

When is $(x,y) \in$ Sister $\circ$ Parent?

When is $(x,y) \in$ Parent $\circ$ Sister?

$S \circ R = \{(a, c) \mid \exists$ b such that $(a,b) \in R$ and $(b,c) \in S\}$

Using the relations: Parent, Child, Brother, Sister, Sibling, Father, Mother, Husband, Wife express:

Uncle: b is an uncle of a

Cousin: b is a cousin of a

Using the relations: Parent, Child, Brother, Sister, Sibling, Father, Mother, Husband, Wife express:

Uncle: b is an uncle of a

$$\text{Uncle} = \text{Brother} \circ \text{Parent}$$

Cousin: b is a cousin of a

$$\text{Cousin} = \text{Child} \circ \text{Sibling} \circ \text{Parent}$$

powers of a relation

$$R^2 = R \circ R$$
$$= \{(a, c) \mid \exists b \text{ such that } (a, b) \in R \text{ and } (b, c) \in R\ \}$$

$$R^0 = \{(a, a) \mid a \in A\}$$

$$R^1 = R$$

$$R^{n+1} = R^n \circ R$$

powers of a relation

$$R^2 = R \circ R$$
$$= \{(a, c) \mid \exists b \text{ such that } (a, b) \in R \text{ and } (b, c) \in R\}$$

$$\text{Parent}^2 = \text{GrandParent}$$

$$R^0 = \{(a, a) \mid a \in A\}$$

R^0 is always equality

$$R^1 = R$$

$$R^{n+1} = R^n \circ R$$

matrix represenation

Relation R on $A = \{a_1, \ldots, a_p\}$

$$m_{ij} = \begin{cases} 1 \text{ if } (a_i, a_j) \in R, \\ 0 \text{ if } (a_i, a_j) \notin R. \end{cases}$$

{(1, 1), (1, 2), (1, 4), (2,1), (2,3), (3,2), (3, 3), (4,2), (4,3)}

	1	2	3	4
1	1	1	0	1
2	1	0	1	0
3	0	1	1	0
4	0	1	1	0

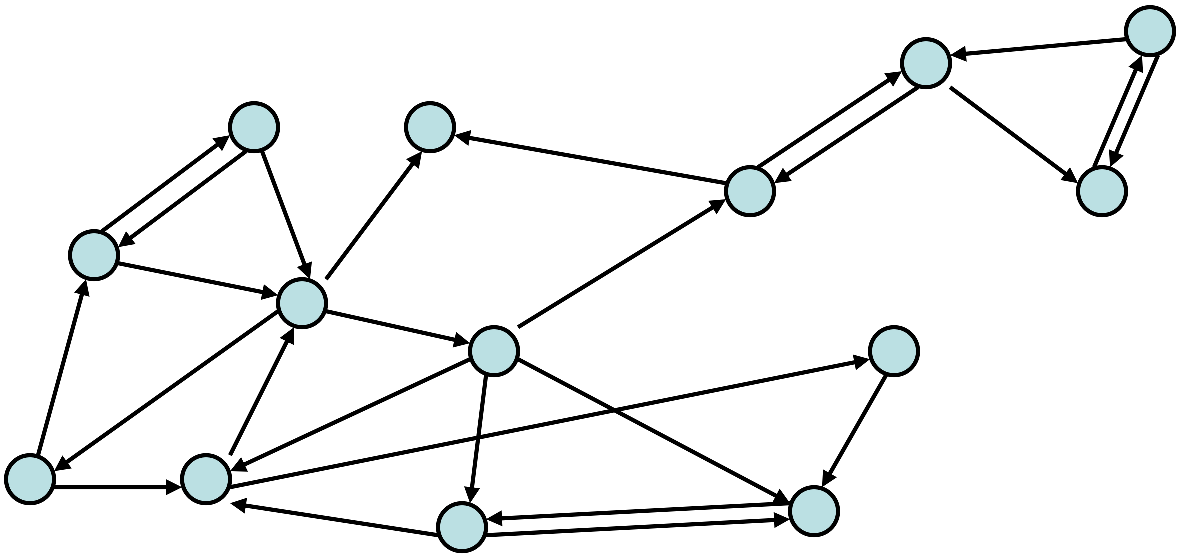
directed graphs

G = (V, E)

V – vertices
E – edges, ordered pairs of vertices

Path: $v_0, v_1, \ldots, v_k$, with (v_i, v_{i+1}) in E

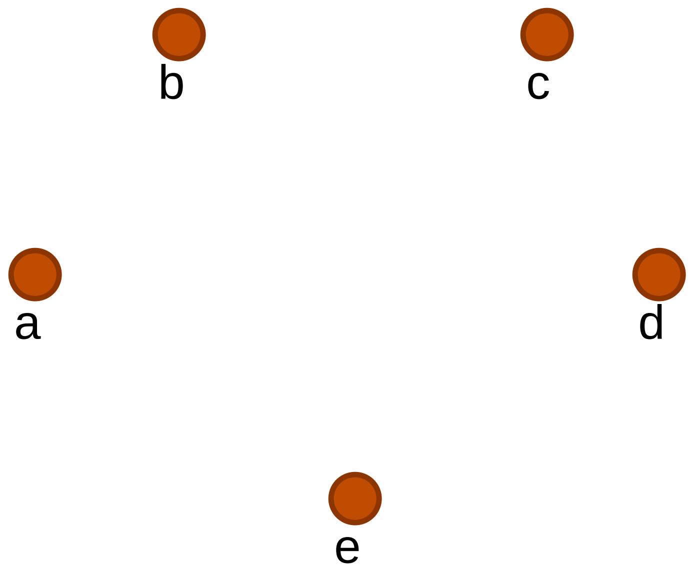
Simple Path
Cycle
Simple Cycle



representation of relations

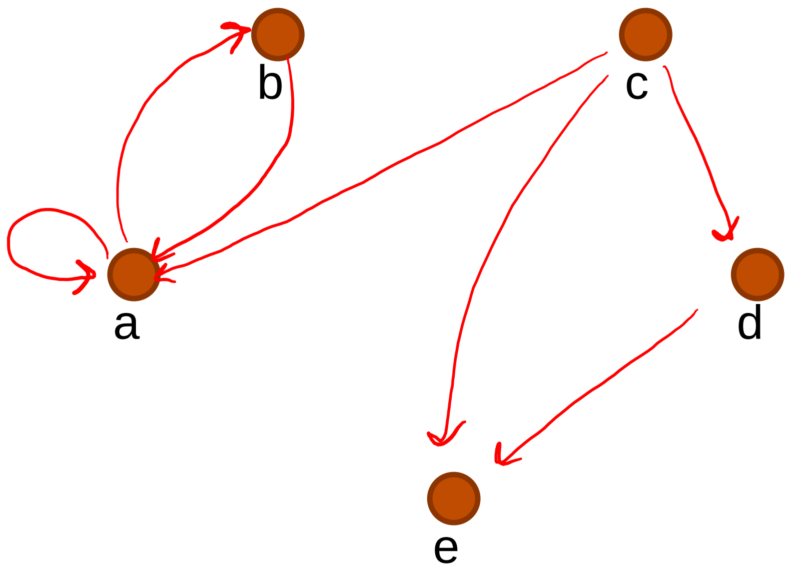
Directed Graph Representation (Digraph)

{(a, b), (a, a), (b, a), (c, a), (c, d), (c, e) (d, e) }



Directed Graph Representation (Digraph)

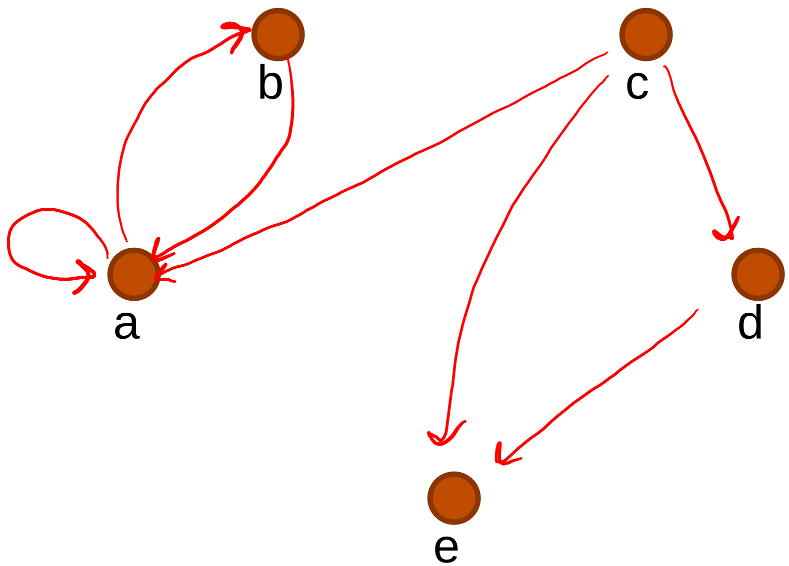
{(a, b), (a, a), (b, a), (c, a), (c, d), (c, e) (d, e) }

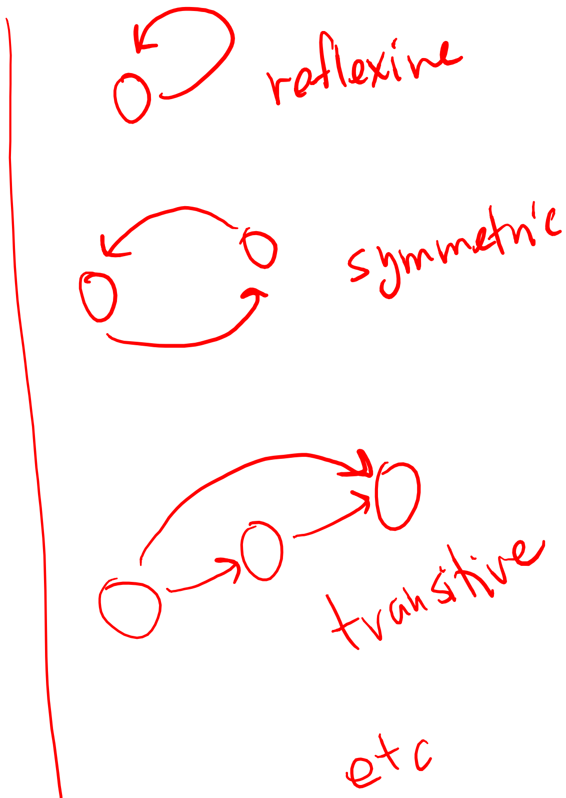


representation of relations

Directed Graph Representation (Digraph)

{(a, b), (a, a), (b, a), (c, a), (c, d), (c, e) (d, e) }





relational composition using digraphs

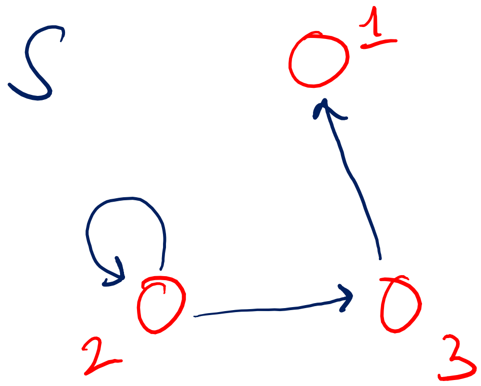
If $S = \{(2,2),(2,3),(3,1)\}$ and $R = \{(1,2),(2,1),(1,3)\}$

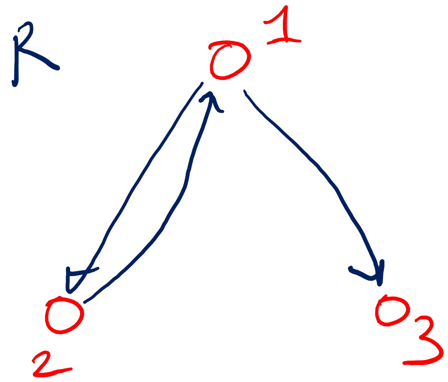
Compute $S \circ R$

relational composition using digraphs

If $S = \{(2,2), (2,3), (3,1)\}$ and $R = \{(1,2), (2,1), (1,3)\}$

Compute $S \circ R$

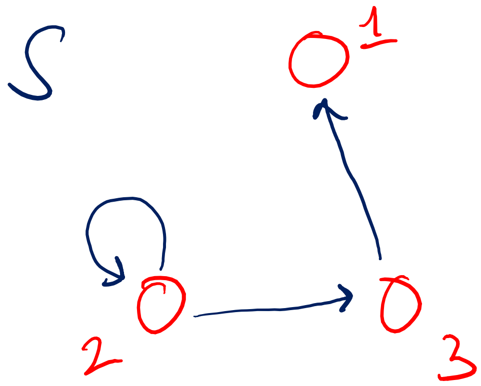


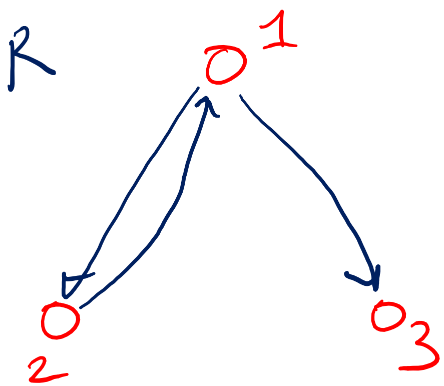


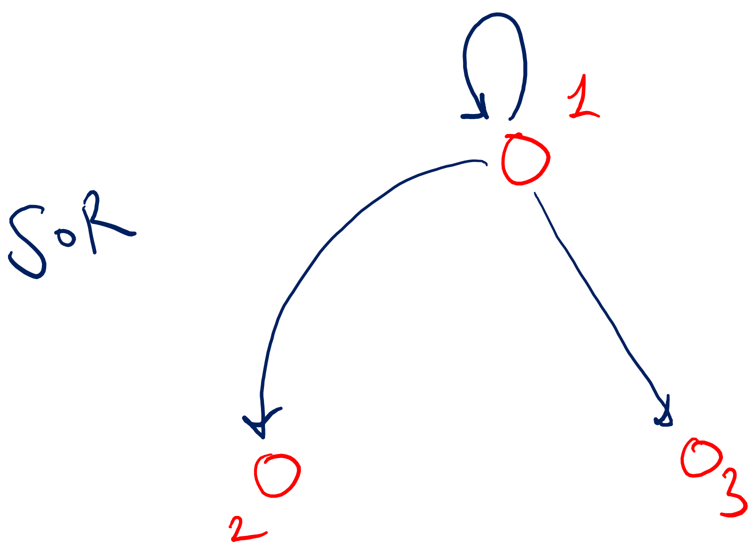
relational composition using digraphs

If $S = \{(2,2), (2,3), (3,1)\}$ and $R = \{(1,2), (2,1), (1,3)\}$

Compute $S \circ R$

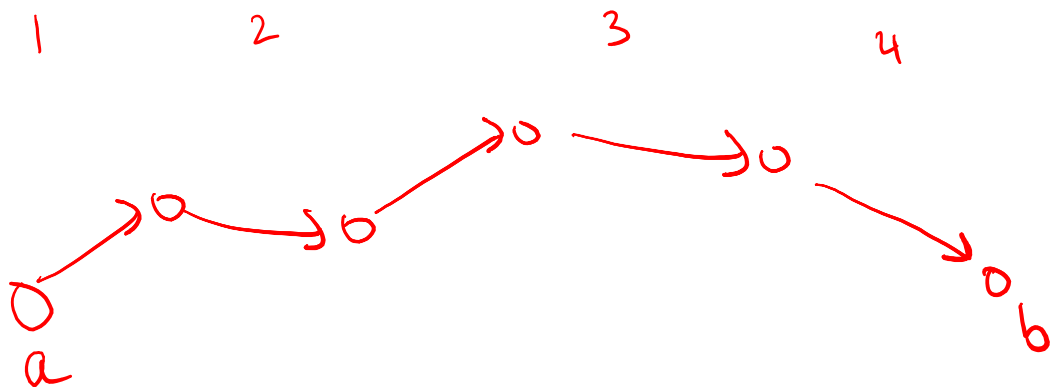






A **path** in a graph of length n is a list of edges with vertices next to each other.

Let R be a relation on a set A. There is a path of length n from a to b if and only if $(a,b) \in R^n$



$(a,b) \in R^4$

connectivity in graphs

Two vertices in a graph are **connected** iff there is a path between them.

Let R be a relation on a set A. The connectivity relation R* consists of the pairs (a,b) such that there is a path from a to b in R.

$$R^* = \bigcup_{k=0}^{\infty} R^k$$

Note: The book uses the wrong definition of this quantity.
What the text defines (ignoring k=0) is usually called R^+

properties of relations (repeated)

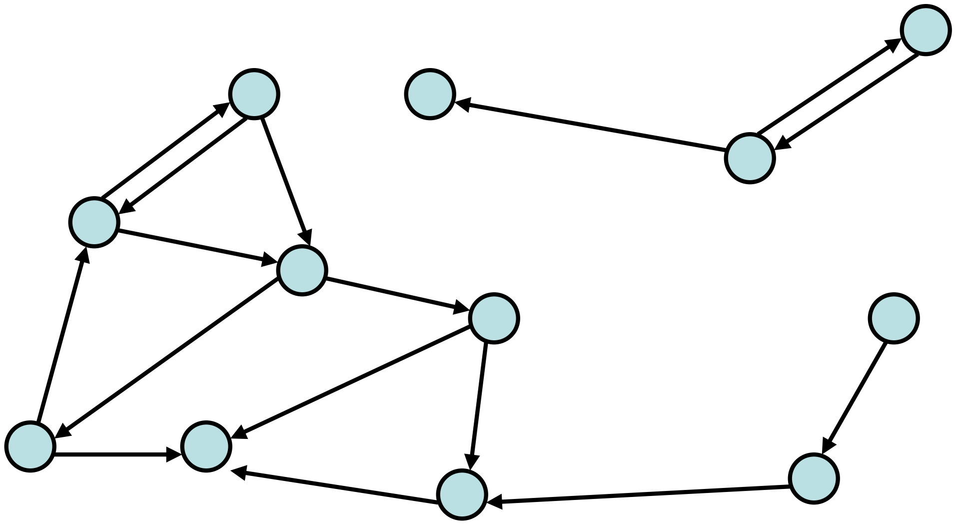
Let R be a relation on A.

R is reflexive iff $(a,a) \in R$ for every $a \in A$

R is symmetric iff $(a,b) \in R$ implies $(b, a) \in R$

R is transitive iff $(a,b) \in R$ and $(b, c) \in R$ implies $(a, c) \in R$

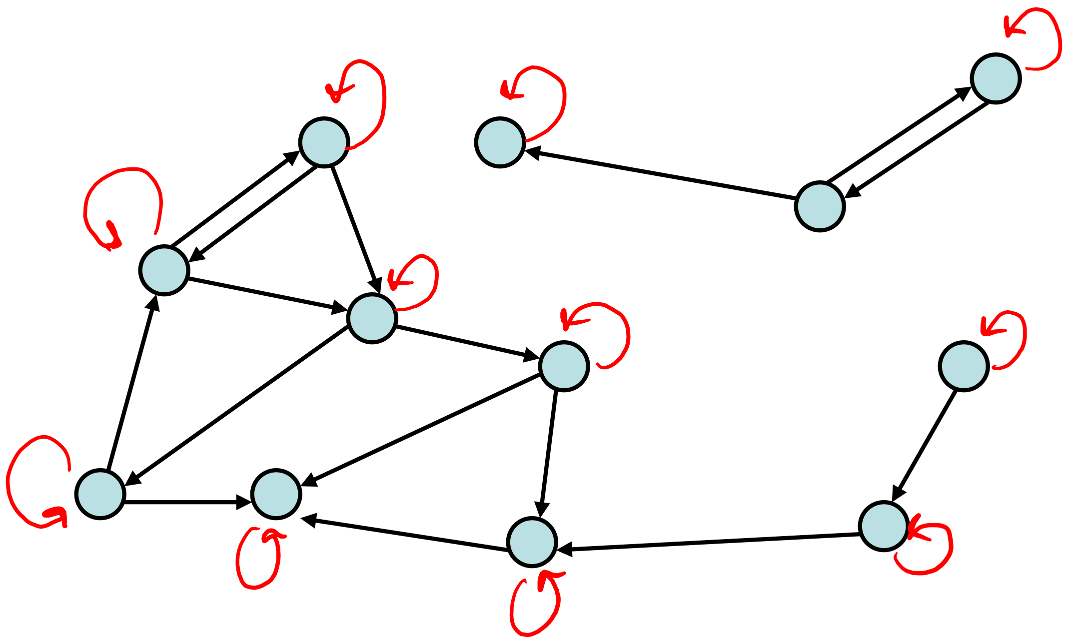
transitive-reflexive closure



Add the minimum possible number of edges to make the relation transitive and reflexive.

The transitive-reflexive closure of a relation R is the connectivity relation R*

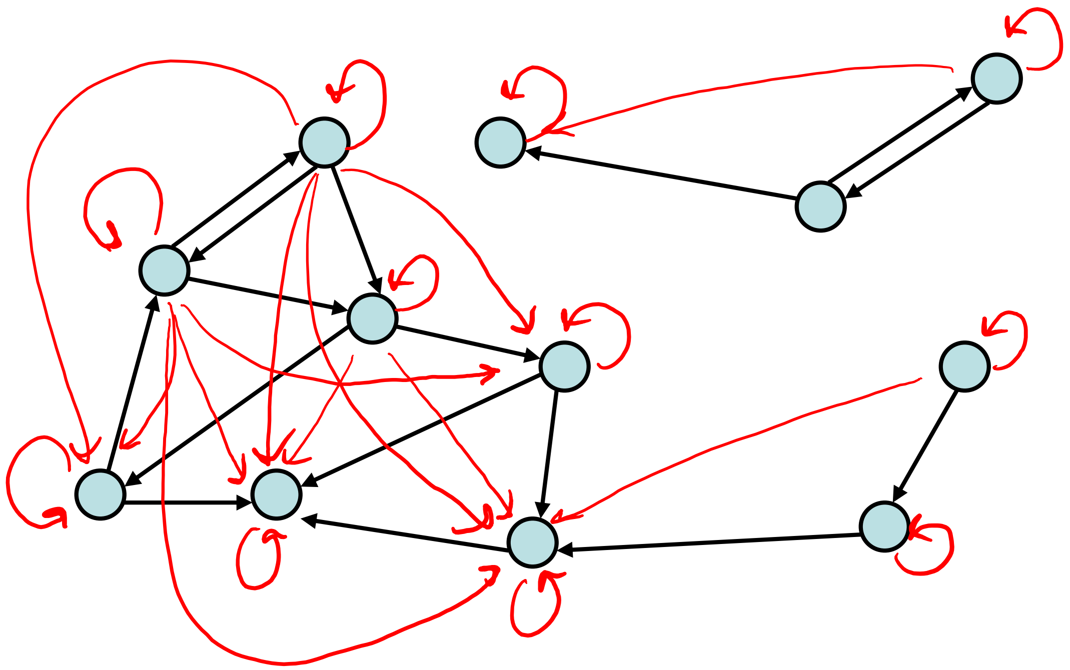
transitive-reflexive closure



Add the minimum possible number of edges to make the relation transitive and reflexive.

The transitive-reflexive closure of a relation R is the connectivity relation R*

transitive-reflexive closure



Add the minimum possible number of edges to make the relation transitive and reflexive.

The transitive-reflexive closure of a relation R is the connectivity relation R*

n-ary relations

Let A_1, A_2, …, A_n be sets. An **n-ary** relation on these sets is a subset of $A_1 \times A_2 \times \cdots \times A_n$.

$(a_1, a_2, \ldots, a_n) \in A_1 \times A_2 \times \cdots \times A_n$

relational databases

STUDENT

Student_Name	ID_Number	Office	GPA
Knuth	328012098	022	4.00
Von Neuman	481080220	555	3.78
Russell	238082388	022	3.85
Einstein	238001920	022	2.11
Newton	1727017	333	3.61
Karp	348882811	022	3.98
Bernoulli	2921938	022	3.21

STUDENT

Student_Name	ID_Number	Office	GPA	Course
Knuth	328012098	022	4.00	CSE311
Knuth	328012098	022	4.00	CSE351
Von Neuman	481080220	555	3.78	CSE311
Russell	238082388	022	3.85	CSE312
Russell	238082388	022	3.85	CSE344
Russell	238082388	022	3.85	CSE351
Newton	1727017	333	3.61	CSE312
Karp	348882811	022	3.98	CSE311
Karp	348882811	022	3.98	CSE312
Karp	348882811	022	3.98	CSE344
Karp	348882811	022	3.98	CSE351
Bernoulli	2921938	022	3.21	CSE351

What’s not so nice?

relational databases

STUDENT

Student_Name	ID_Number	Office	GPA
Knuth	328012098	022	4.00
Von Neuman	481080220	555	3.78
Russell	238082388	022	3.85
Einstein	238001920	022	2.11
Newton	1727017	333	3.61
Karp	348882811	022	3.98
Bernoulli	2921938	022	3.21

TAKES

ID_Number	Course
328012098	CSE311
328012098	CSE351
481080220	CSE311
238082388	CSE312
238082388	CSE344
238082388	CSE351
1727017	CSE312
348882811	CSE311
348882811	CSE312
348882811	CSE344
348882811	CSE351
2921938	CSE351

Better

database operations: projection

Find all offices: $\Pi_{\text{Office}}(\text{STUDENT})$

Office
022
555
333

Find offices and GPAs: $\Pi_{\text{Office,GPA}}(\text{STUDENT})$

Office	GPA
022	4.00
555	3.78
022	3.85
022	2.11
333	3.61
022	3.98
022	3.21

database operations: selection

Find students with GPA > 3.9 : $\sigma_{GPA>3.9}(STUDENT)$

Student_Name	ID_Number	Office	GPA
Knuth	328012098	022	4.00
Karp	348882811	022	3.98

Retrieve the name and GPA for students with GPA > 3.9:
$\Pi_{Student\_Name,GPA}(\sigma_{GPA>3.9}(STUDENT))$

Student_Name	GPA
Knuth	4.00
Karp	3.98

database operations: natural join

Student ⋈ Takes

Student_Name	ID_Number	Office	GPA	Course
Knuth	328012098	022	4.00	CSE311
Knuth	328012098	022	4.00	CSE351
Von Neuman	481080220	555	3.78	CSE311
Russell	238082388	022	3.85	CSE312
Russell	238082388	022	3.85	CSE344
Russell	238082388	022	3.85	CSE351
Newton	1727017	333	3.61	CSE312
Karp	348882811	022	3.98	CSE311
Karp	348882811	022	3.98	CSE312
Karp	348882811	022	3.98	CSE344
Karp	348882811	022	3.98	CSE351
Bernoulli	2921938	022	3.21	CSE351