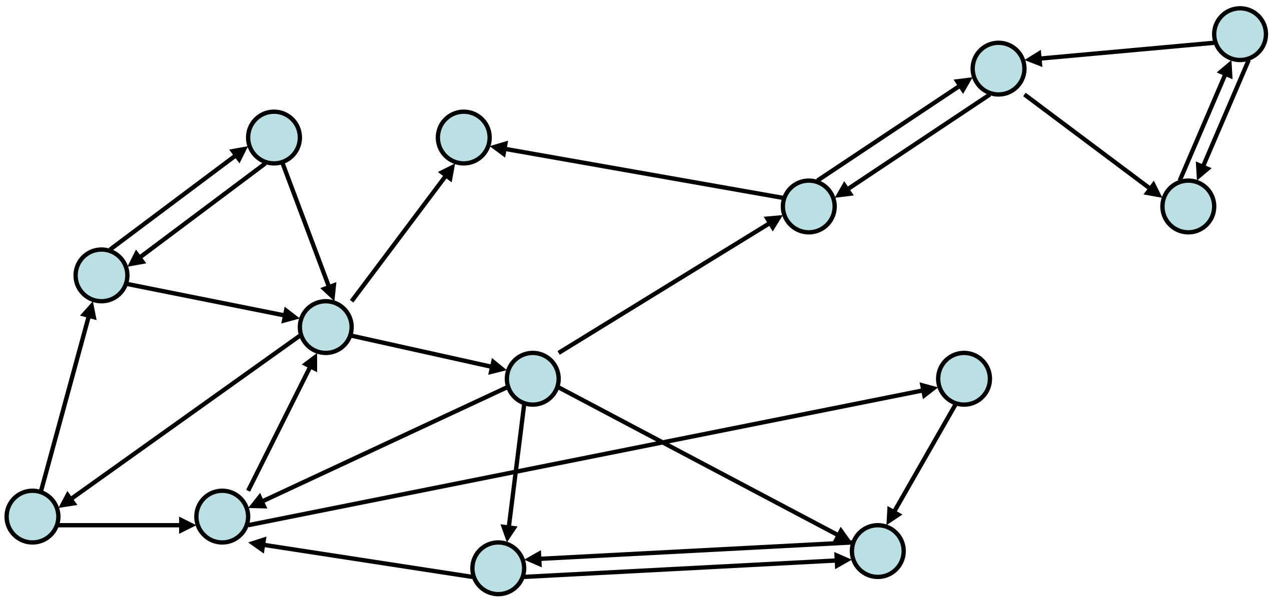
CSE 311: Foundations of Computing

Lecture 20: Context-Free Grammars, Relations and Directed Graphs



Review: Regular Expressions

ε matches the **empty string**

a matches the one character string *a*

A $\cup$ **B** matches all strings that either **A** matches or **B** matches (or both)

AB matches all strings that have a first part that **A** matches followed by a second part that **B** matches

A* matches all strings that have any number of strings (even 0) that **A** matches, one after another

– equivalently, $A^* = \varepsilon \cup A \cup AA \cup AAA \cup \ldots$

Review: Context Free Grammars

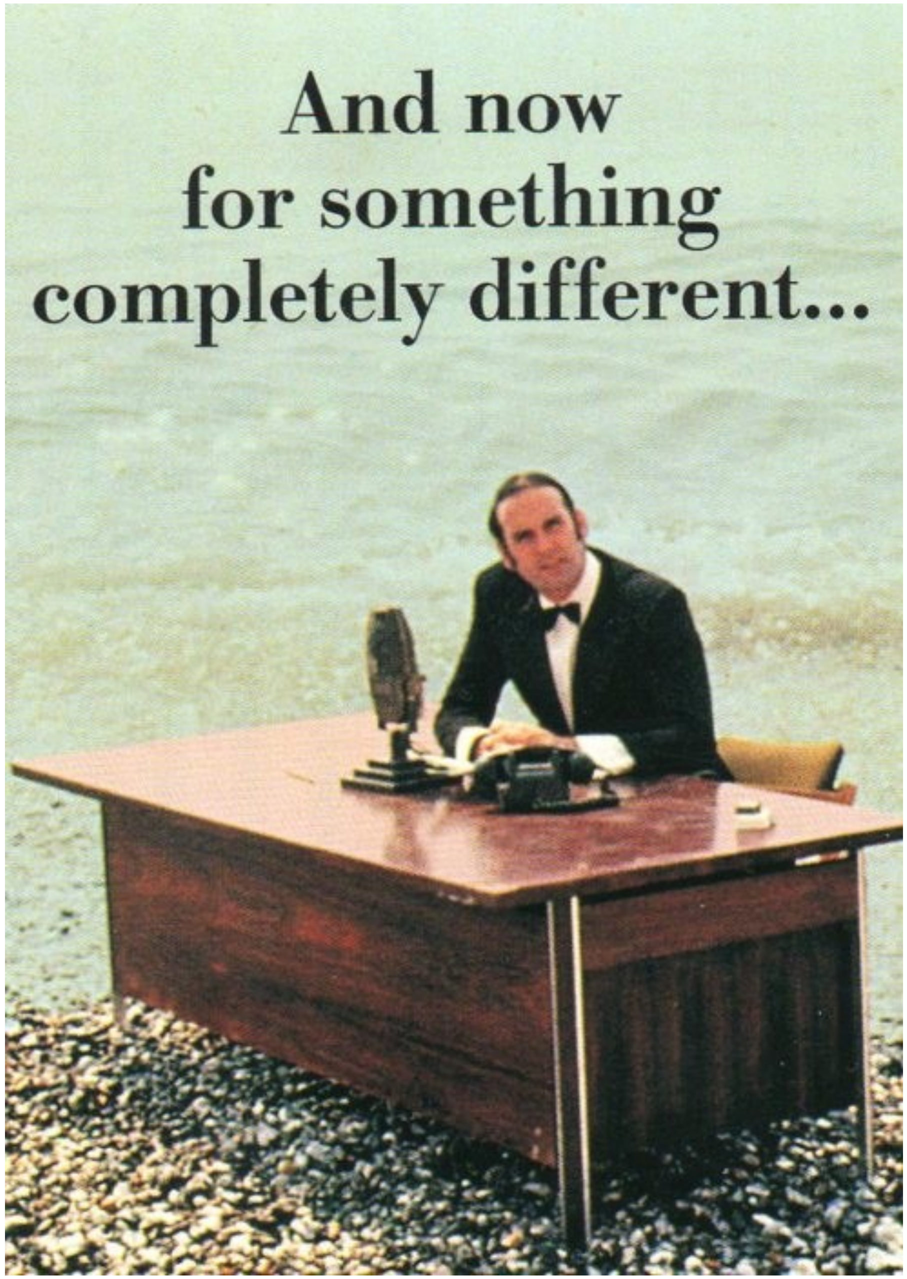
- A Context-Free Grammar (CFG) is given by a finite set of substitution rules involving
  - A finite set **V** of *variables* that can be replaced
  - Alphabet Σ of *terminal symbols* that can't be replaced
  - One variable, usually **S**, is called the *start symbol*

- The rules involving a variable **A** are written as

$$\mathbf{A} \rightarrow w_1 \mid w_2 \mid \cdots \mid w_k$$

where each w_i is a string of variables and terminals – that is $w_i \in (\mathbf{V} \cup \Sigma)^*$

Relations and Directed Graphs



Relations

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 $\mathcal{P}(U)$ for universe U

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

More Relation Examples

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

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

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

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

Properties of Relations

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 $(a, c) \in R$

Which relations have which properties?

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

$<$ **on** $\mathbb{R}$ **:**

$=$ **on** Σ^* **:**

$\subseteq$ **on** $\mathcal{P}$**(U):**

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

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

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 $(a, c) \in R$

Which relations have which properties?

$\geq$ on $\mathbb{N}$: Reflexive, Antisymmetric, Transitive

$<$ on $\mathbb{R}$: Antisymmetric, Transitive

$=$ on Σ^* : Reflexive, Symmetric, Antisymmetric, Transitive

$\subseteq$ on $\mathcal{P}(U)$: Reflexive, Antisymmetric, Transitive

$R_2 = \{(x, y) \mid x \equiv y \pmod 5\}$: Reflexive, Symmetric, Transitive

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

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 $(a, c) \in R$

Combining Relations

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

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

$R \circ S = \{$ (a, c) | $\exists$ b such that (a,b)$\in R$ 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 $(x,y) \in$ Parent $\circ$ Sister?

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

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

Examples

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

Powers of a Relation

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

$\boldsymbol{R^0} = \{(\boldsymbol{a}, \boldsymbol{a}) \mid \boldsymbol{a} \in \boldsymbol{A}\}$ “the equality relation on $\boldsymbol{A}$”

$$\boldsymbol{R^1} = \boldsymbol{R} = \boldsymbol{R^0} \circ \boldsymbol{R}$$

$\boldsymbol{R^{n+1}} = \boldsymbol{R^n} \circ \boldsymbol{R}$ for $\boldsymbol{n} \geq \boldsymbol{0}$

Matrix Representation

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

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

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

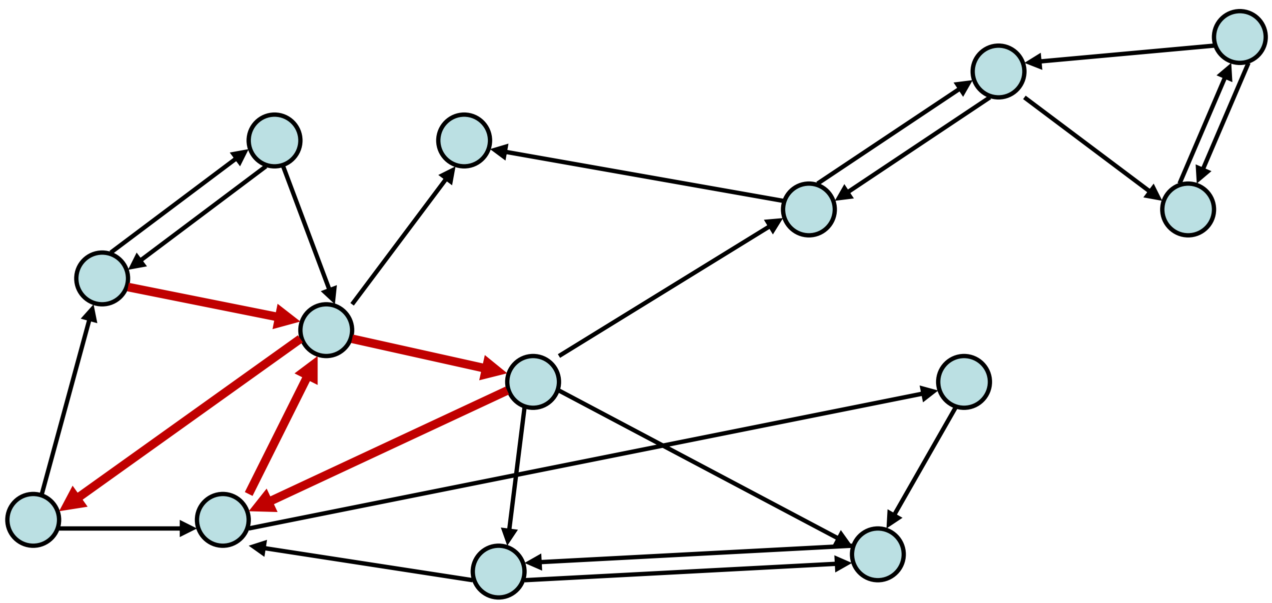
	1	**2**	**3**	**4**
1	1	1	0	1
2	1	0	1	0
3	0	1	1	0
4	0	1	1	0

Directed Graphs

G = (V, E) V – vertices
E – edges, ordered pairs of vertices

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

Simple Path: none of $\mathbf{v_0}, \ldots, \mathbf{v_k}$ repeated
Cycle: $\mathbf{v_0} = \mathbf{v_k}$
Simple Cycle: $\mathbf{v_0} = \mathbf{v_k}$, none of $\mathbf{v_1}, \ldots, \mathbf{v_k}$ repeated

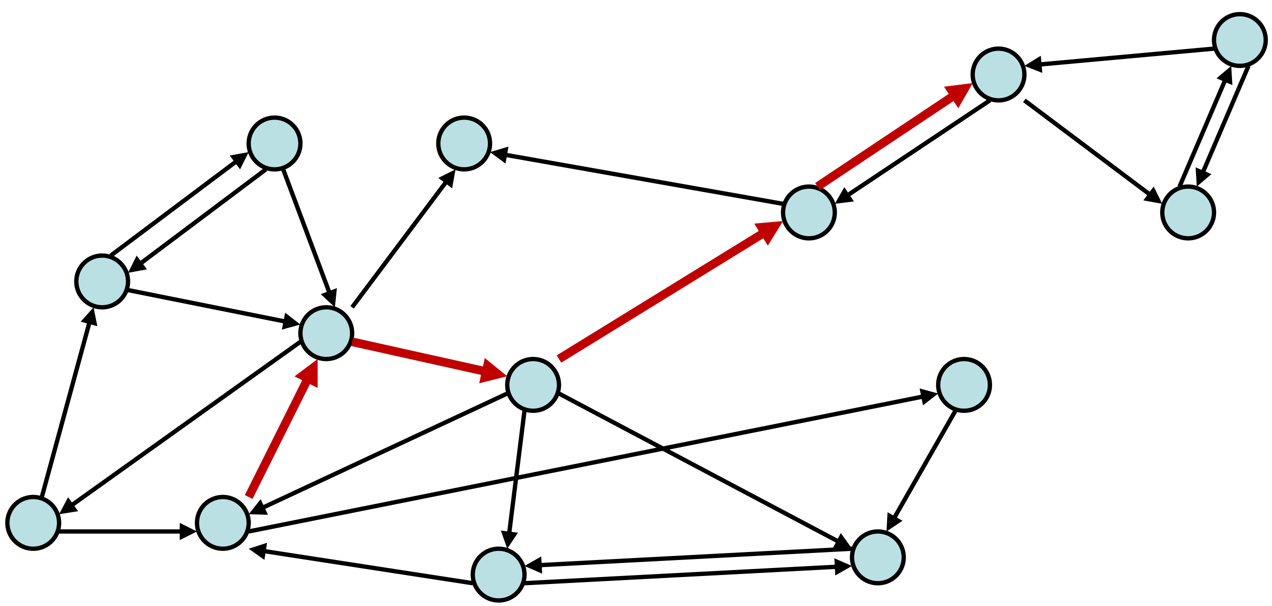


Directed Graphs

$G = (V, E)$ V – vertices
E – edges, ordered pairs of vertices

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

Simple Path: none of $\mathbf{v_0}, \ldots, \mathbf{v_k}$ repeated
Cycle: $\mathbf{v_0} = \mathbf{v_k}$
Simple Cycle: $\mathbf{v_0} = \mathbf{v_k}$, none of $\mathbf{v_1}, \ldots, \mathbf{v_k}$ repeated

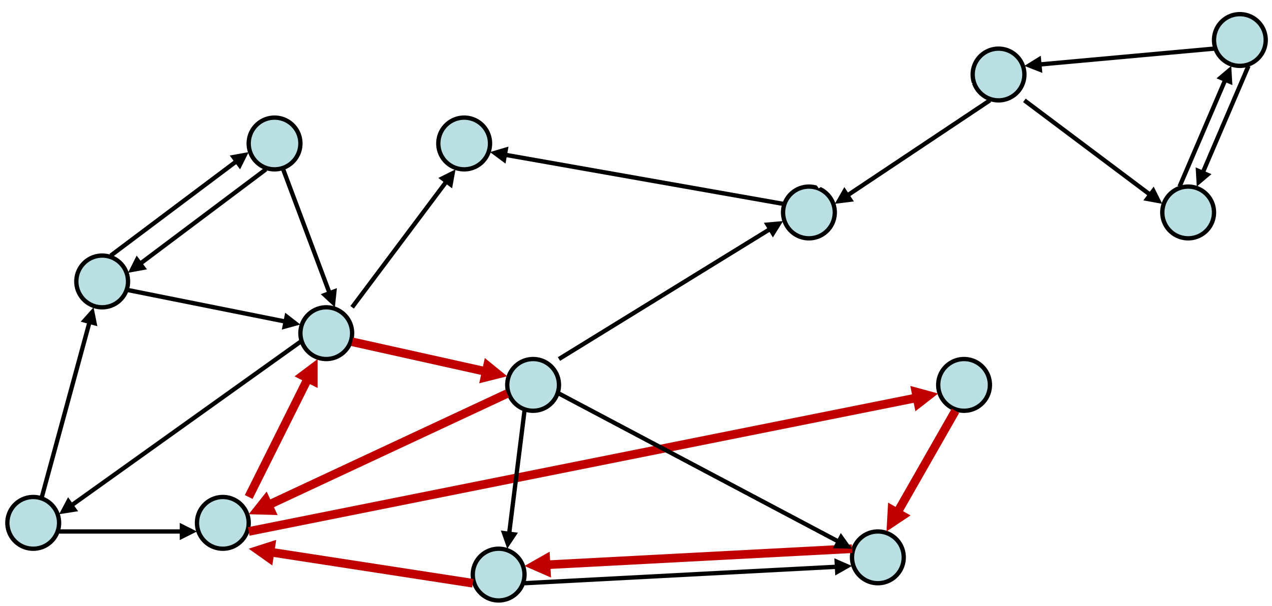


Directed Graphs

$G = (V, E)$ V – vertices
E – edges, ordered pairs of vertices

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

Simple Path: none of $\mathbf{v_0}, \ldots, \mathbf{v_k}$ repeated
Cycle: $\mathbf{v_0} = \mathbf{v_k}$
Simple Cycle: $\mathbf{v_0} = \mathbf{v_k}$, none of $\mathbf{v_1}, \ldots, \mathbf{v_k}$ repeated

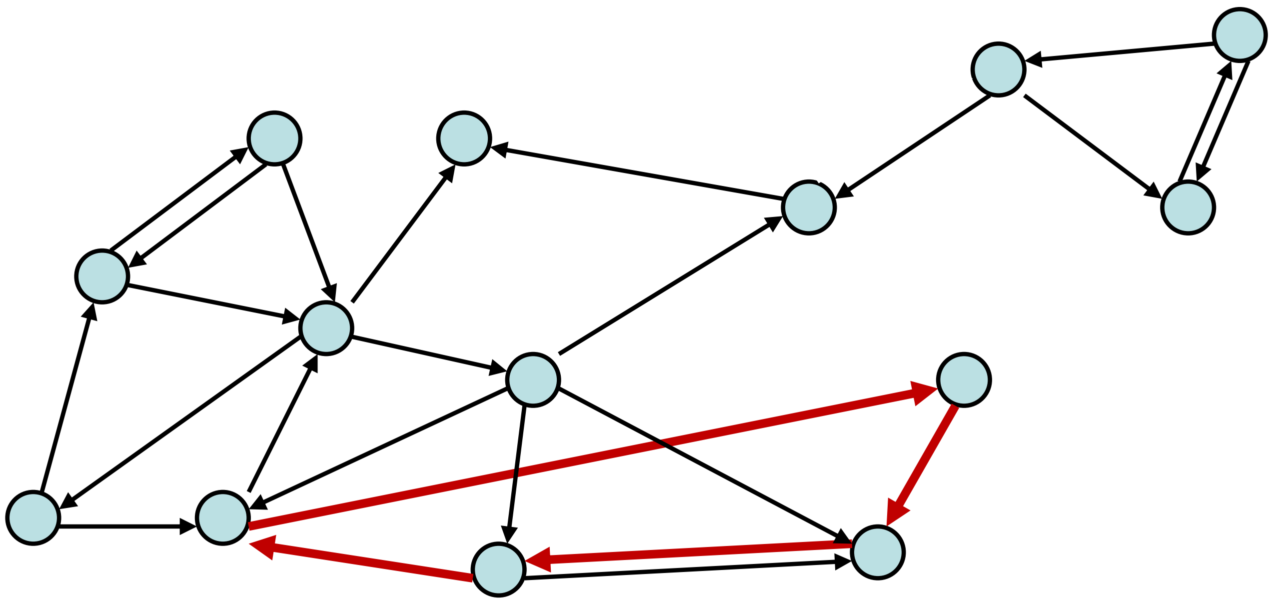


Directed Graphs

$G = (V, E)$ V – vertices
E – edges, ordered pairs of vertices

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

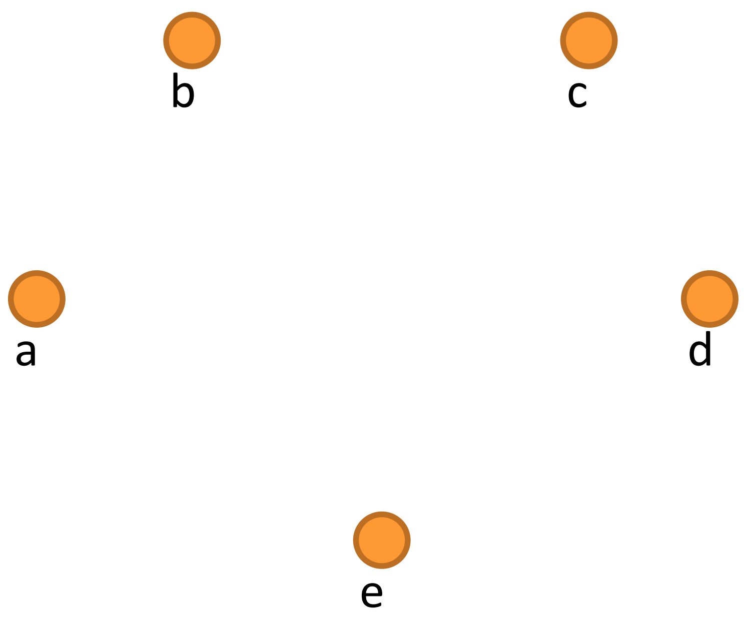
Simple Path: none of $\mathbf{v_0}, \ldots, \mathbf{v_k}$ repeated
Cycle: $\mathbf{v_0} = \mathbf{v_k}$
Simple Cycle: $\mathbf{v_0} = \mathbf{v_k}$, none of $\mathbf{v_1}, \ldots, \mathbf{v_k}$ repeated



Representation of Relations

Directed Graph Representation (Digraph)

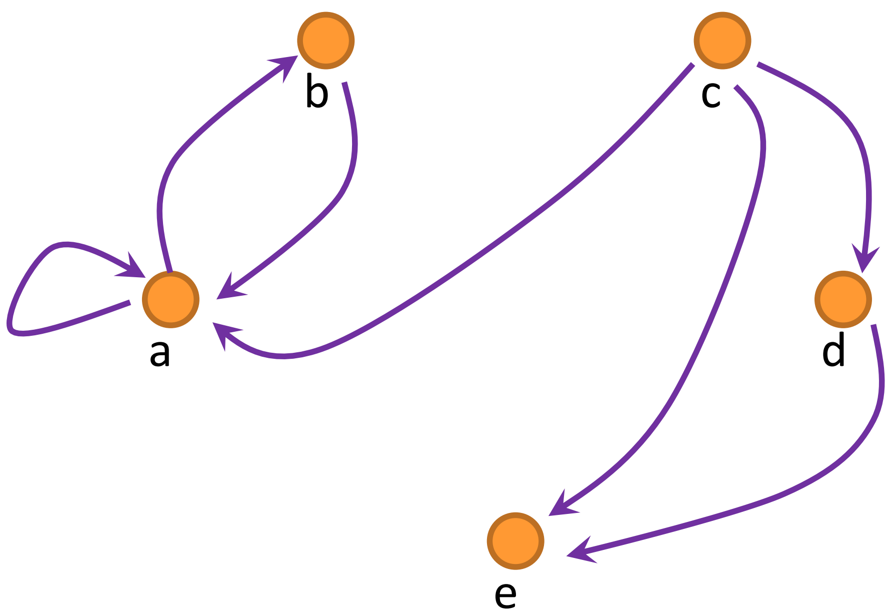
{(a, b), (a, a), (b, a), (c, a), (c, d), (c, e) (d, e) }



Representation of Relations

Directed Graph Representation (Digraph)

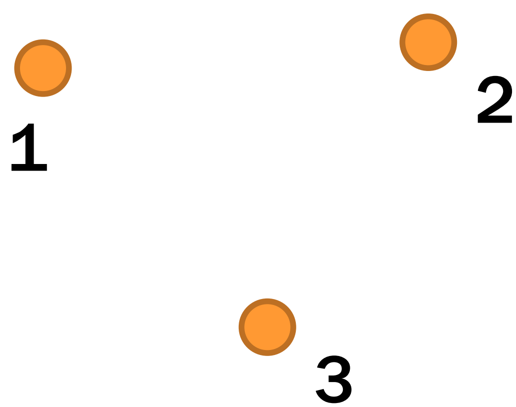
{(a, b), (a, a), (b, a), (c, a), (c, d), (c, e) (d, e) }

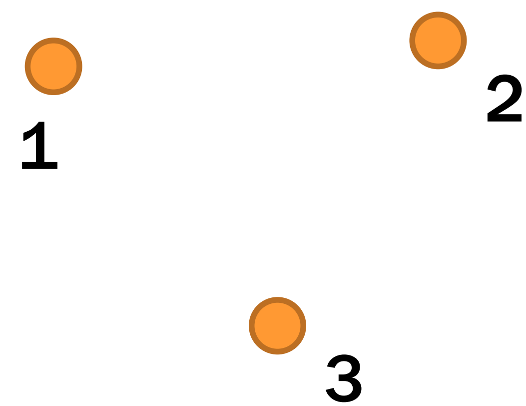


Relational Composition using Digraphs

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

Compute $R \circ S$

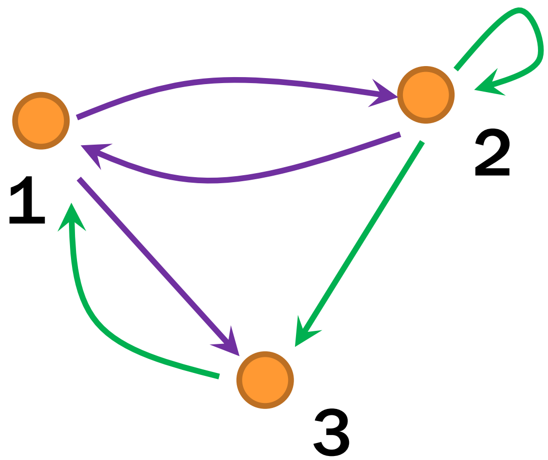


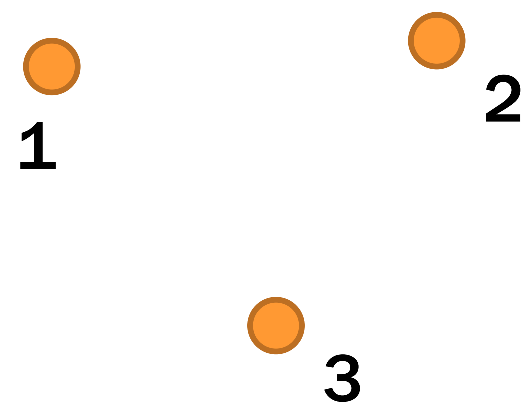


Relational Composition using Digraphs

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

Compute $\boldsymbol{R} \circ \boldsymbol{S}$

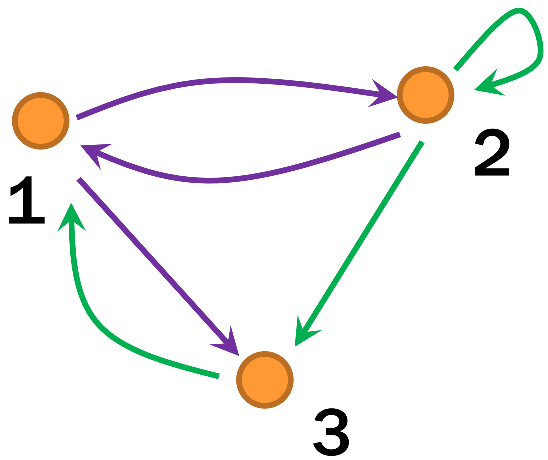


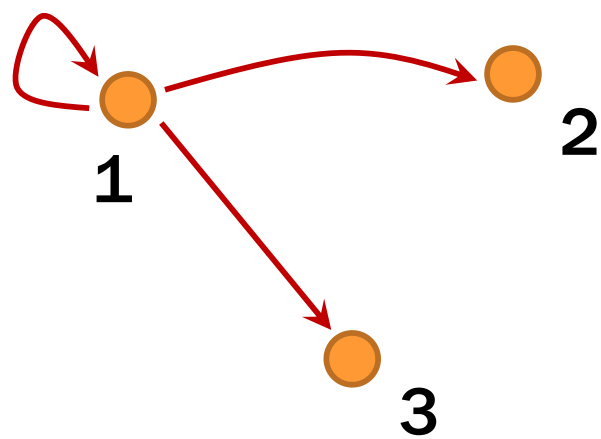


Relational Composition using Digraphs

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

Compute $R \circ S$





Paths in Relations and Graphs

Defn: The **length** of a path in a graph is the number of edges in it (counting repetitions if edge used > once).

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

Connectivity In Graphs

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

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

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

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

How Properties of Relations show up in Graphs

Let R be a relation on A.

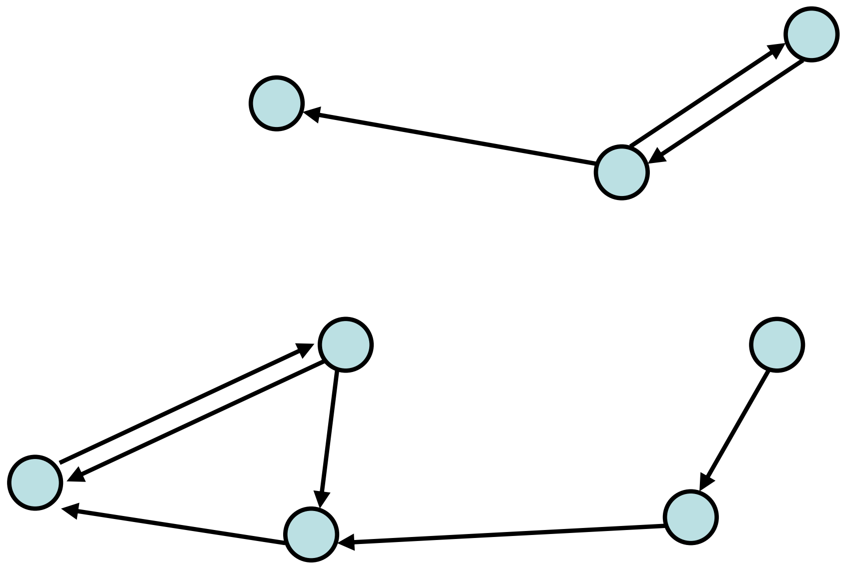
R is **reflexive** iff $(a,a) \in R$ for every $a \in A$

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 $(a, c) \in R$

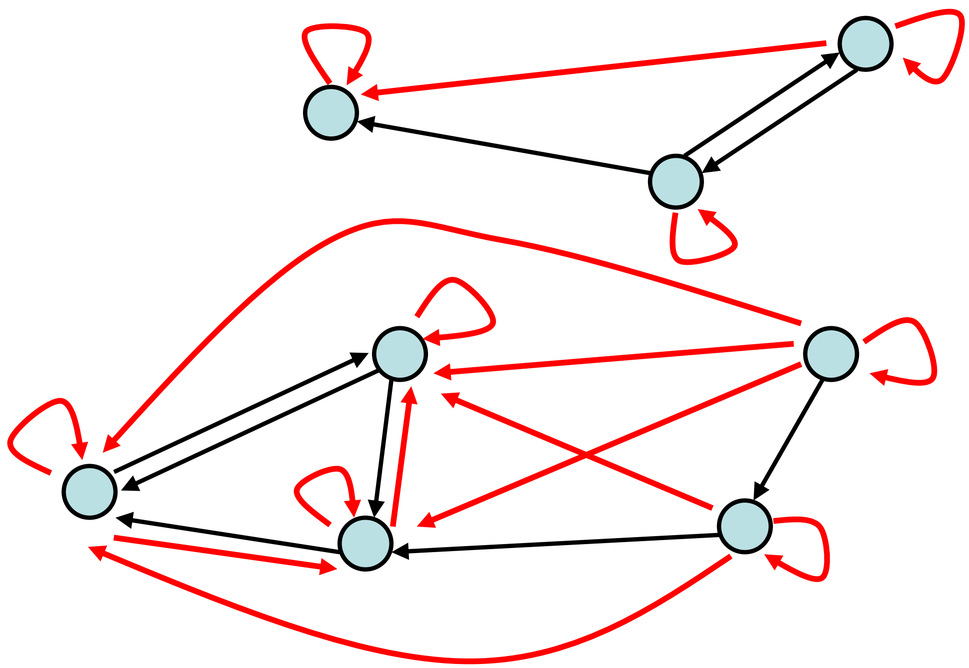
Transitive-Reflexive Closure



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

The **transitive-reflexive closure** of a relation $\boldsymbol{R}$ is the connectivity relation $\boldsymbol{R}^*$

Transitive-Reflexive Closure



Relation with the **minimum possible** number of **extra edges** to make the relation both transitive and reflexive.

The **transitive-reflexive closure** of a relation $\boldsymbol{R}$ is the connectivity relation $\boldsymbol{R}^*$

n-ary Relations

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

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

Relational Databases

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: $\Pi_{\mathbf{Office}}(\mathrm{STUDENT})$

Office
022
555
333

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

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 : $\boldsymbol{\sigma}_{\mathbf{GPA>3.9}}$(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:

$\boldsymbol{\Pi}_{\mathbf{Student\_Name,GPA}}(\boldsymbol{\sigma}_{\mathbf{GPA>3.9}}(\text{STUDENT}))$

Student_Name	GPA
Knuth	4.00
Karp	3.98

Database Operations: Natural Join

Student $\bowtie$ 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