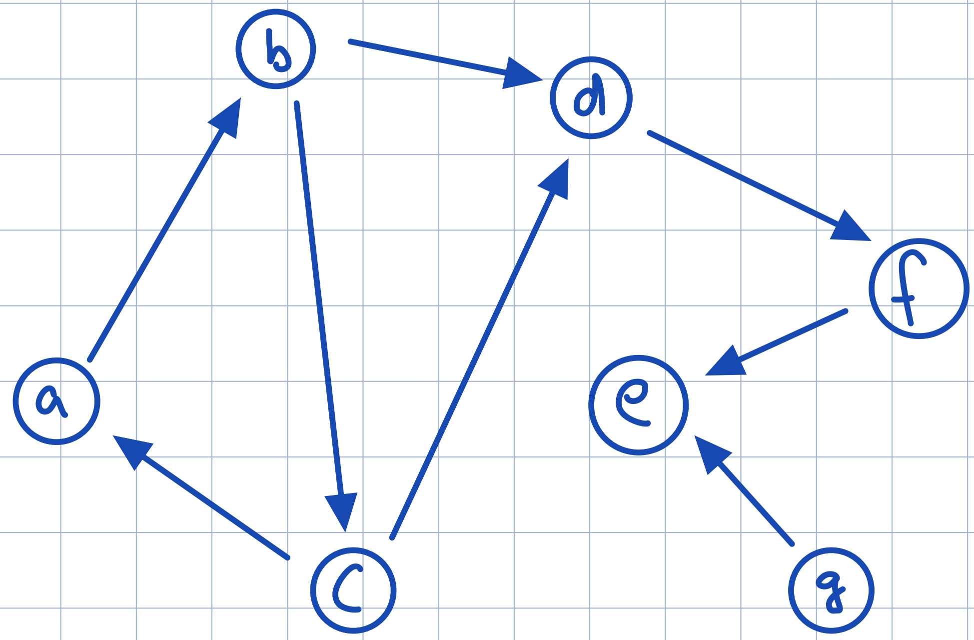
Fixed-point Computations

Neel Krishnaswami
University of Cambridge

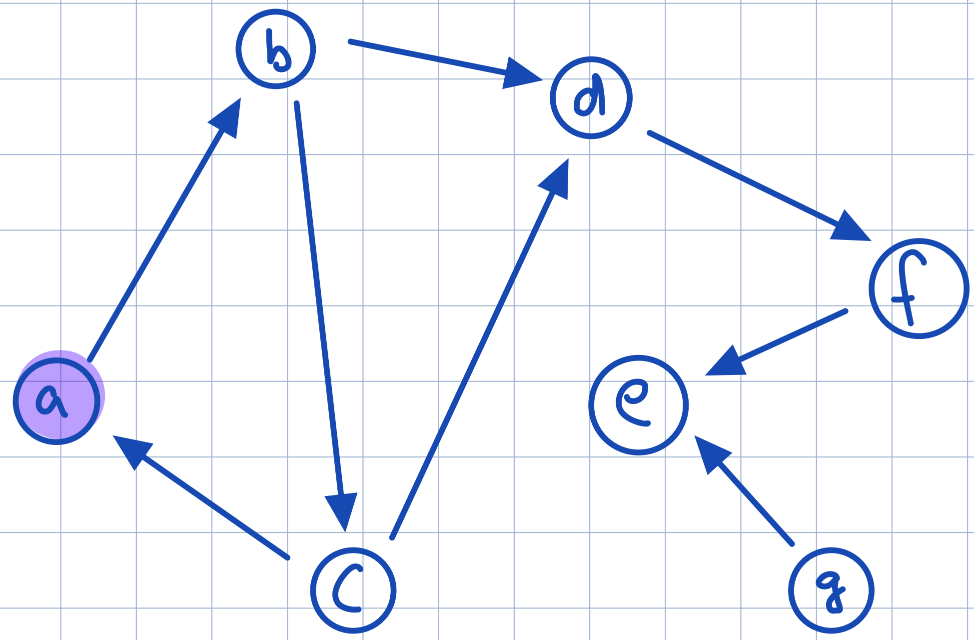
FLOLAC 2024
Taipei, Taiwan

A Graph



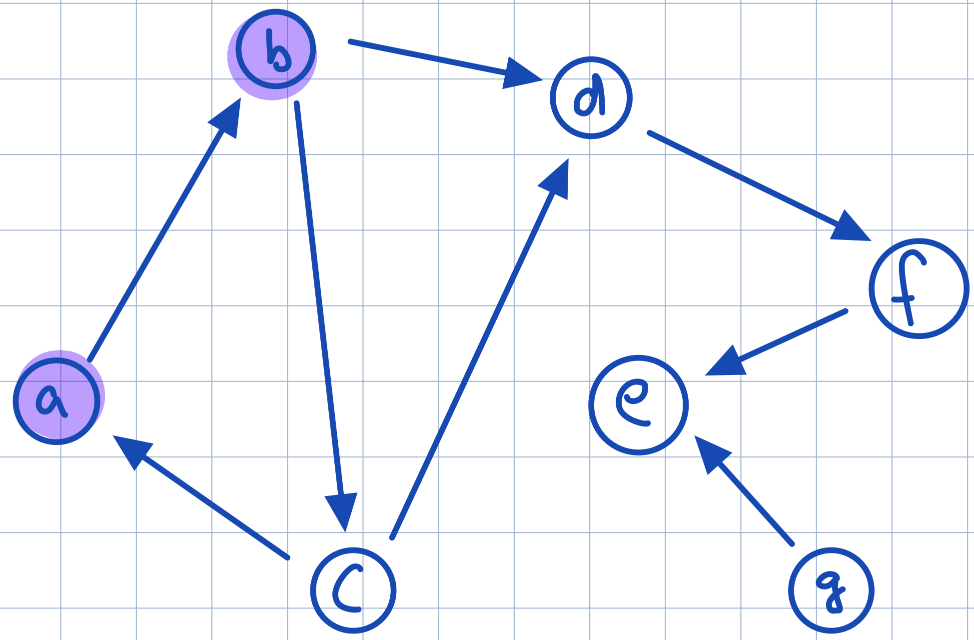
Q: Which nodes are reachable from a?

A Graph



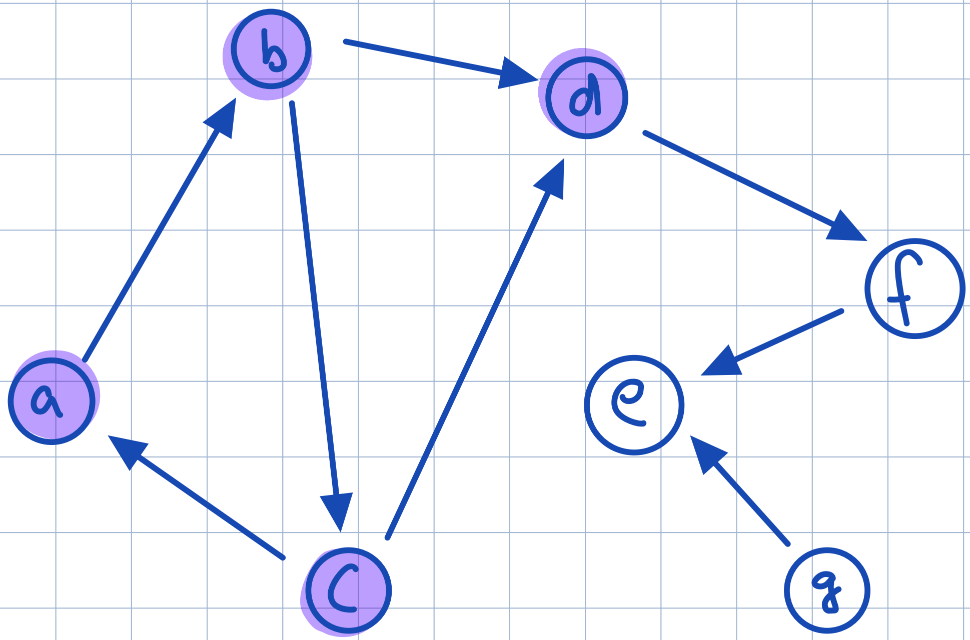
Q: Which nodes are reachable from a?

A Graph



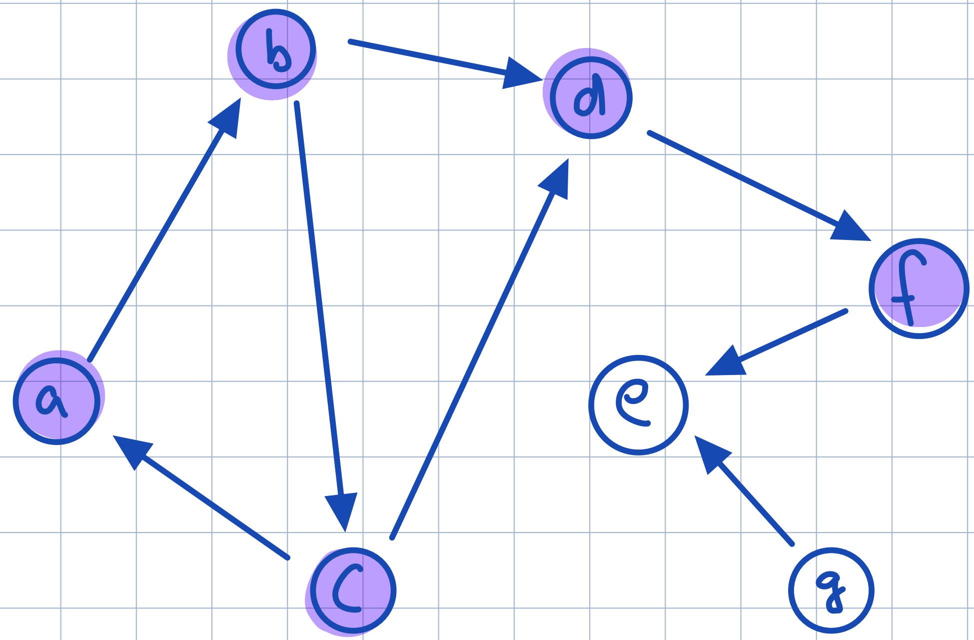
Q: Which nodes are reachable from a?

A Graph



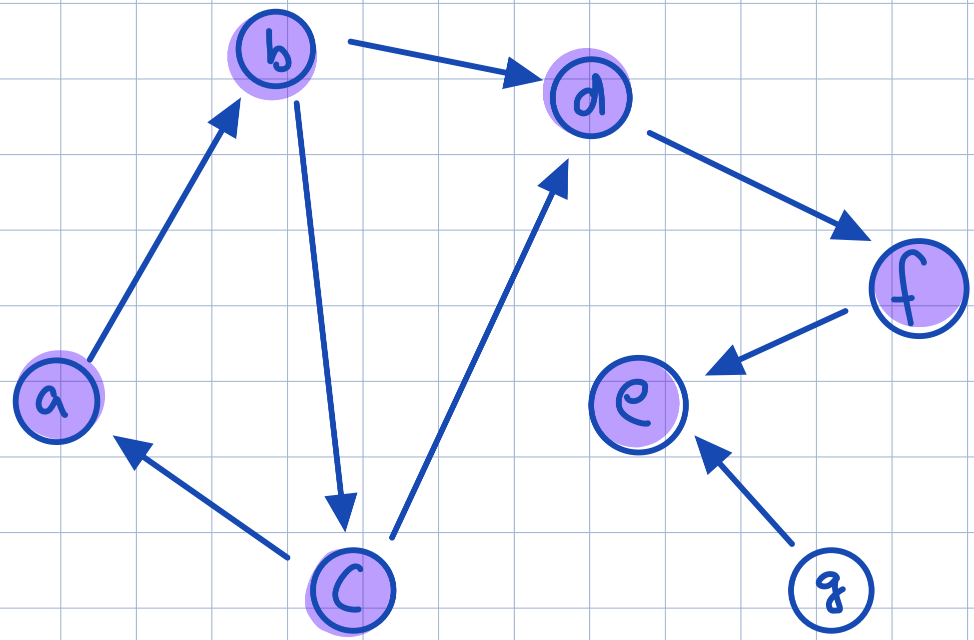
Q: Which nodes are reachable from a?

A Graph



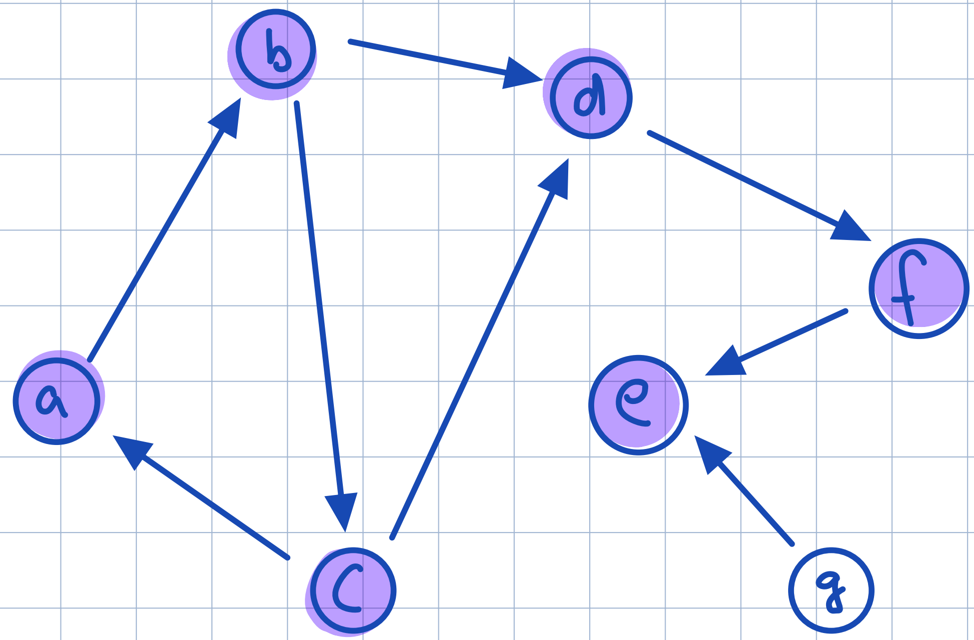
Q: Which nodes are reachable from a?

A Graph



Q: Which nodes are reachable from a?

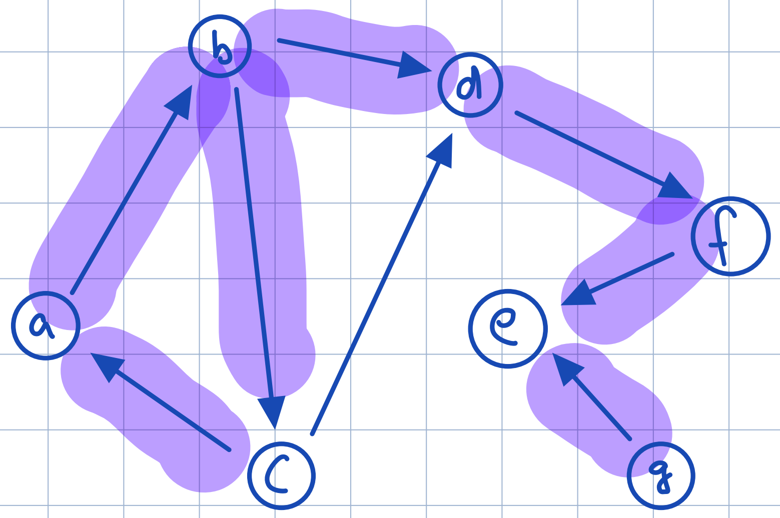
A Graph



Q: Which nodes are reachable from a?

A: {a, b, c, d, e, f} (but not g)

A Graph



edge (a, b)

edge (c, a)

edge (c, d)

edge (b, c)

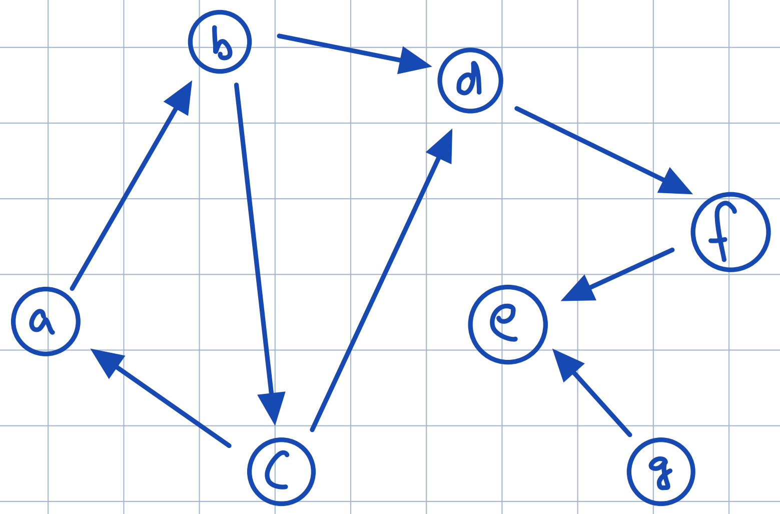
edge (b, d)

edge (d, f)

edge (f, e)

edge (g, e)

A Graph



$$\frac{}{\mathrm{edge}(a,b)} \qquad \frac{}{\mathrm{edge}(c,a)} \qquad \frac{}{\mathrm{edge}(c,d)}$$

$$\frac{}{\mathrm{edge}(b,c)} \qquad \frac{}{\mathrm{edge}(b,d)}$$

$$\frac{}{\mathrm{edge}(d,f)} \qquad \frac{}{\mathrm{edge}(f,e)} \qquad \frac{}{\mathrm{edge}(g,e)}$$

$$\frac{\mathrm{edge}(x,y)}{\mathrm{reach}(x,y)} \qquad \frac{\mathrm{edge}(x,y) \quad \mathrm{reach}(y,z)}{\mathrm{reach}(x,z)}$$

BNF Grammars

You have seen many BNF grammars:

$$A ::= 1 \mid A \times A \mid A \to A$$

$$\Gamma ::= \cdot \mid \Gamma, x : A$$

$$e ::= () \mid (e, e) \mid \lambda x. e \mid \pi_i(e) \mid e\,e \mid x$$

BNF Grammars

This is shorthand for:

$$A \Rightarrow 1$$
$$A \Rightarrow A \times A$$
$$A \Rightarrow A \to A$$

$$\Gamma \Rightarrow \cdot$$
$$\Gamma \Rightarrow \Gamma, x : A$$

Chomsky Normal Form

Any grammar can be rewritten so that every production is either

$A \longrightarrow a \qquad A \longrightarrow BC$

(Many new nonterminals will be created)

Example

$$A \longrightarrow 1$$

$$A \longrightarrow A \times A$$

Example

$A \rightarrow 1$

$A \rightarrow A \times A$

$T \rightarrow x$

$A \rightarrow 1$

$A \rightarrow A\, T\, A$

Example

$A \rightarrow 1$

$A \rightarrow A \times A$

$T \rightarrow x$

$A \rightarrow 1$

$A \rightarrow A\ T\ A$

$T \rightarrow x$

$A \rightarrow 1$

$A \rightarrow A\ K$

$K \rightarrow T\ A$

CYK Parsing

1. Suppose G is a grammar in Chomsky NF
2. Let w be a word of length n
3. w_i = ith symbol of w
4. Define:

$$\frac{\text{parse}(B, i, j) \quad \text{parse}(C, j, k)}{\text{parse}(A, i, k)} \quad \text{for each } A \to BC \text{ in } G$$

$$\frac{}{\text{parse}(A, i, i+1)} \quad \text{for each } A \to s \text{ s.t. } s = w_i \text{ in } G$$

CYK Example

$T \rightarrow x$

$A \rightarrow 1$

$A \rightarrow AK$

$K \rightarrow TA$

$w = 1x1$

CYK Example

$T \to x$

$A \to 1$

$A \to A\,K$

$K \to T\,A$

$w = 1 \times 1$

parse(A, 0, 1)

CYK Example

$T \to x$

$A \to 1$

$A \to A\,K$

$K \to T\,A$

$w = 1x1$

parse(A,0,1)	parse(T,1,2)

CYK Example

$T \to x$

$A \to 1$

$A \to A\ K$

$K \to T\ A$

$w = 1 \times 1$

parse(A, 0, 1)

parse(T, 1, 2)

parse(A, 2, 3)

CYK Example

$T \to x$

$A \to 1$

$A \to A\ K$

$K \to T\ A$

$w = 1 \times 1$

$$\frac{}{\text{parse}(A,0,1)} \qquad \frac{}{\text{parse}(T,1,2)} \qquad \frac{}{\text{parse}(A,2,3)}$$

$$\frac{\text{parse}(A,i,j) \quad \text{parse}(K,j,k)}{\text{parse}(A,i,k)}$$

CYK Example

$T \to x$

$A \to 1$

$A \to A\ K$

$K \to T\ A$

$w = \underline{1} \times \underline{1}$

$$\frac{}{\text{parse}(A,0,1)} \qquad \frac{}{\text{parse}(T,1,2)} \qquad \frac{}{\text{parse}(A,2,3)}$$

$$\frac{\text{parse}(A,i,j) \quad \text{parse}(K,j,k)}{\text{parse}(A,i,k)}$$

$$\frac{\text{parse}(T,i,j) \quad \text{parse}(A,j,k)}{\text{parse}(K,i,k)}$$

CYK Example

$$\frac{}{\text{parse}(A,0,1)} \qquad \frac{}{\text{parse}(T,1,2)} \qquad \frac{}{\text{parse}(A,2,3)}$$

$$\frac{\text{parse}(A,i,j) \quad \text{parse}(K,j,k)}{\text{parse}(A,i,k)}$$

$$\frac{\text{parse}(T,i,j) \quad \text{parse}(A,j,k)}{\text{parse}(K,i,k)}$$

CYK Example

$$\frac{}{parse(A,0,1)} \qquad \frac{}{parse(T,1,2)} \qquad \frac{}{parse(A,2,3)}$$

$$\frac{parse(A,i,j) \quad parse(K,j,k)}{parse(A,i,k)}$$

$$\frac{parse(T,i,j) \quad parse(A,j,k)}{parse(K,i,k)}$$

$$\frac{parse(T,1,2) \quad parse(A,2,3)}{parse(K,1,3)}$$

CYK Example

$$\overline{\text{parse}(A,0,1)} \qquad \overline{\text{parse}(T,1,2)} \qquad \overline{\text{parse}(A,2,3)}$$

$$\frac{\text{parse}(A,i,j) \quad \text{parse}(K,j,k)}{\text{parse}(A,i,k)}$$

$$\frac{\text{parse}(T,i,j) \quad \text{parse}(A,j,k)}{\text{parse}(K,i,k)}$$

$$\text{parse}(K,1,3)$$

CYK Example

$$\frac{}{parse(A,0,1)} \qquad \frac{}{parse(T,1,2)} \qquad \frac{}{parse(A,2,3)}$$

$$\frac{parse(A,i,j) \quad parse(K,j,k)}{parse(A,i,k)}$$

$$\frac{parse(T,i,j) \quad parse(A,j,k)}{parse(K,i,k)}$$

$$parse(K,1,3)$$

$$\frac{parse(A,0,1) \quad parse(K,1,3)}{parse(A,0,3)}$$

CYK Example

$$\frac{}{parse(A,0,1)} \qquad \frac{}{parse(T,1,2)} \qquad \frac{}{parse(A,2,3)}$$

$$\frac{parse(A,i,j) \quad parse(K,j,k)}{parse(A,i,k)}$$

$$\frac{parse(T,i,j) \quad parse(A,j,k)}{parse(K,i,k)}$$

$parse(K,1,3)$

$parse(A,0,3)$

CYK Example

$$\frac{}{\text{parse}(A, 0, 1)} \qquad \frac{}{\text{parse}(T, 1, 2)} \qquad \frac{}{\text{parse}(A, 2, 3)}$$

$$\frac{\text{parse}(A, i, j) \quad \text{parse}(K, j, k)}{\text{parse}(A, i, k)}$$

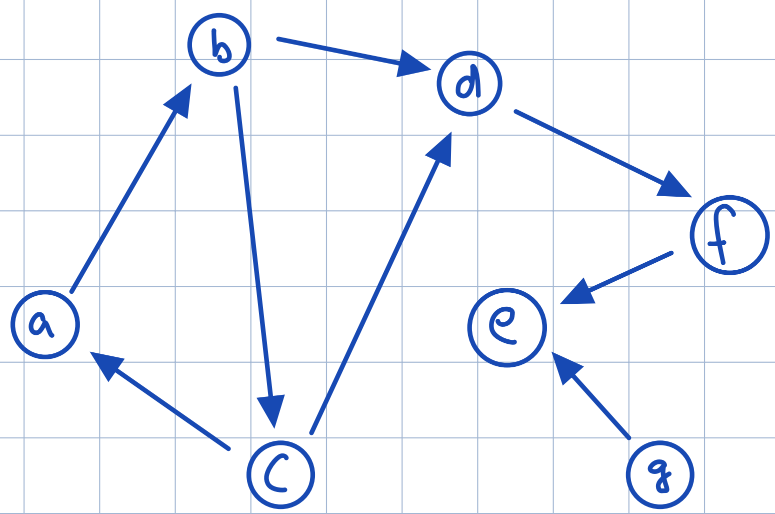
$$\frac{\text{parse}(T, i, j) \quad \text{parse}(A, j, k)}{\text{parse}(K, i, k)}$$

$\text{parse}(K, 1, 3)$

$\text{parse}(A, 0, 3)$

successful parse! → $\text{parse}(A, 0, 3)$

Relations, Mathematically



$$\frac{}{\mathit{edge}(a,b)}$$

$$\frac{}{\mathit{edge}(b,c)} \qquad \frac{}{\mathit{edge}(b,d)}$$

$$\frac{}{\mathit{edge}(d,f)} \qquad \frac{}{\mathit{edge}(f,e)} \qquad \frac{}{\mathit{edge}(g,e)}$$

$$\frac{\mathit{edge}(x,y)}{\mathit{reach}(x,y)} \qquad \frac{\mathit{edge}(x,y) \quad \mathit{reach}(y,z)}{\mathit{reach}(x,z)}$$

Relations, Mathematically

$$\frac{}{\mathit{edge}(a,b)}$$

$$\frac{}{\mathit{edge}(b,c)} \qquad \frac{}{\mathit{edge}(b,d)}$$

$$\frac{}{\mathit{edge}(d,f)} \qquad \frac{}{\mathit{edge}(f,e)} \qquad \frac{}{\mathit{edge}(g,e)}$$

$$\frac{\mathit{edge}(x,y)}{\mathit{reach}(x,y)}$$

$$\frac{\mathit{edge}(x,y) \quad \mathit{reach}(y,z)}{\mathit{reach}(x,z)}$$

Relations, Mathematically

$$\frac{}{edge(a,b)}$$

$$\frac{}{edge(b,c)} \qquad \frac{}{edge(b,d)}$$

$$\frac{}{edge(d,f)} \qquad \frac{}{edge(f,e)} \qquad \frac{}{edge(g,e)}$$

$Edge \subseteq Node \times Node$

$$\frac{edge(x,y)}{reach(x,y)}$$

$$\frac{edge(x,y) \quad reach(y,z)}{reach(x,z)}$$

Relations, Mathematically

$$\frac{}{edge(a,b)}$$

$$\frac{}{edge(b,c)} \qquad \frac{}{edge(b,d)}$$

$$\frac{}{edge(d,f)} \qquad \frac{}{edge(f,e)} \qquad \frac{}{edge(g,e)}$$

$Edge \subseteq Node \times Node$

$Edge = \{(a,b), (b,c), (b,d), (d,f), (f,e), (g,e)\}$

$$\frac{edge(x,y)}{reach(x,y)}$$

$$\frac{edge(x,y) \quad reach(y,z)}{reach(x,z)}$$

Relations, Mathematically

$$\frac{}{edge(a,b)}$$

$$\frac{}{edge(b,c)} \qquad \frac{}{edge(b,d)}$$

$$\frac{}{edge(d,f)} \qquad \frac{}{edge(f,e)} \qquad \frac{}{edge(g,e)}$$

$$Edge \subseteq Node \times Node$$

$$Edge = \{(a,b), (b,c), (b,d), (d,f), (f,e), (g,e)\}$$

$$\frac{edge(x,y)}{reach(x,y)}$$

$$Reach = Edge \cup \{(x,z) \mid (x,y) \in Edge, (y,z) \in Reach\}$$

$$\frac{edge(x,y) \quad reach(y,z)}{reach(x,z)}$$

Relations, Mathematically

$$\frac{}{edge(a,b)}$$

$$\frac{}{edge(b,c)} \qquad \frac{}{edge(b,d)}$$

$$\frac{}{edge(d,f)} \qquad \frac{}{edge(f,e)} \qquad \frac{}{edge(g,e)}$$

$$Edge \subseteq Node \times Node$$

$$Edge = \{(a,b),\ (b,c),\ (b,d),\ (d,f),\ (f,e),\ (g,e)\}$$

$$Reach = Edge \cup Edge ; Reach$$

$$\frac{edge(x,y)}{reach(x,y)}$$

$$\frac{edge(x,y) \quad reach(y,z)}{reach(x,z)}$$

A Recursive Definition

$$Edge = \{ (a, b), \ldots \}$$

$$Reach = Edge \cup Edge;Reach$$

A Recursive Definition

$\text{Edge} = \{ (a,b), \ldots \}$

$\text{Reach} = \text{Edge} \cup \text{Edge}; \text{Reach}$

Q: How do we know this definition makes sense?

An Intuitive Idea

Define $Reach_0 = \emptyset$

$Reach_1 = Edge \cup Edge ; Reach_0$

$Reach_2 = Edge \cup Edge ; Reach_1$

$\vdots$

$Reach_{n+1} = Edge \cup Edge ; Reach_n$

If $Reach_{n+1} = Reach_n$, then we have Reach!

An Intuitive Idea

Define $Reach_0 = \phi$

$Reach_1 = Edge \cup Edge ; Reach_0$

$Reach_2 = Edge \cup Edge ; Reach_1$

$\vdots$

$Reach_{n+1} = Edge \cup Edge ; Reach_n$

If $Reach_{n+1} = Reach_n$, then we have Reach!

Why Might This Work?

1. $Reach \subseteq Node \times Node$

2. If $|Node| = m$, then $|Reach| \leq m^2$

3. If $|Reach_{n+1}| > |Reach_n|$, then in $\leq m^2$ steps $Reach_{n+1}$ stabilizes

Monotonicity

Define $F(X) = Edge \cup Edge ; X$

<u>Lemma:</u> If $X \subseteq Y$ then $F(X) \subseteq F(Y)$

Monotonicity

Lemma: If $X \subseteq Y$ then $F(X) \subseteq F(Y)$

Proof:

1. Assume $X \subseteq Y$
2. Assume $(a,c) \in F(X) = Edge \cup Edge;X$
3. Case: $(a,c) \in Edge$
   Then $(a,c) \in F(Y) = Edge \cup Edge;Y$

   Case: $(a,c) \in Edge;X$
   $(a,b) \in Edge$ and $(b,c) \in X$
   $(b,c) \in Y$ since $X \subseteq Y$
   $(a,c) \in Edge;Y$
   $(a,c) \in Edge \cup Edge;Y$
   $(a,c) \in F(Y)$

Formalizing the Intuitive Idea

Suppose X is finite, and $F : \mathcal{P}(X) \to \mathcal{P}(X)$ monotone

Let $R_0 = \emptyset$ and $R_{n+1} = F(R_n)$

1. $\exists\, k$ s.t. $R_k = R_{k+1}$
2. This is the <u>smallest</u> fixed point of F

An Increasing Sequence

Lemma: $\forall n.\ R_n \subseteq R_{n+1}$

Proof: By induction on n

- Case $n = 0$

  $R_0 = \emptyset \qquad R_1 = F(\emptyset)$

  By definition $R_0 \subseteq R_1$

- Case $n = k+1$

  By induction, $R_k \subseteq R_{k+1}$

  By monotonicity, $F(R_k) \subseteq F(R_{k+1})$

  Hence $R_{k+1} \subseteq R_{k+2}$

  So $R_n \subseteq R_{n+1}$

A Fixed Point

We know $R_0 \subseteq R_1 \subseteq \dots \subseteq R_n \subseteq R_{n+1} \subseteq \dots$

Since X is finite, $\mathcal{P}(X)$ is also finite

Hence in at most $|X|$ steps $R_{|X|} = R_{|X|+1}$

A Least Fixed Point

If $F(S) = S$ then $\forall n.\ R_n \subseteq S$

Proof. Assume $S = F(S)$

Proceed by induction on n.

Case $n = 0$.

$R_0 = \emptyset \wedge \emptyset \subseteq S \implies R_0 \subseteq S$

Case $n = k+1$:

By induction, $R_k \subseteq S$

By monotonicity, $F(R_k) \subseteq F(S)$

$R_{k+1} \subseteq F(S)$

Since $S = F(S)$ $\quad R_{k+1} \subseteq S$

Datalog

1. If X finite and $F: \mathcal{P}(X) \to \mathcal{P}(X)$ monotone
then F has a least fixed point

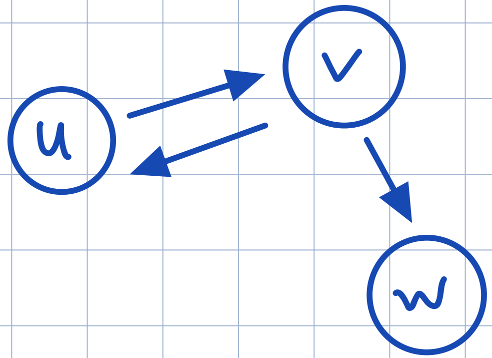
2. Every inductive relation defined
by inference rules over finite sets
has a least fixed point semantics

3. Defining sets by such relations is
the Datalog query language

Datalog

$$\frac{R_1(a, X) \quad \ldots \quad R_1(W, c)}{R(X, W)}$$

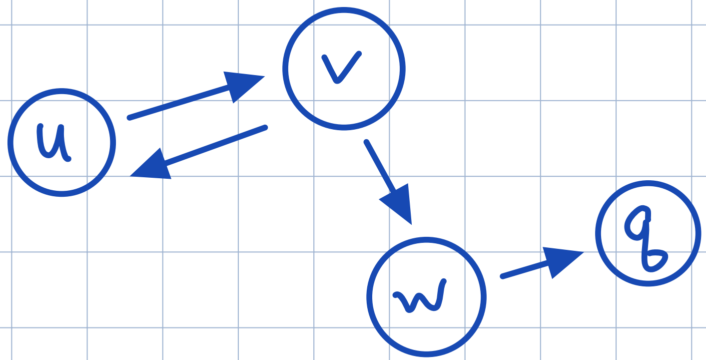
Limitations



$$\frac{}{edge_2(u,v)} \qquad \frac{}{edge_2(v,u)}$$

$$\frac{}{edge_2(v,w)}$$

Limitations



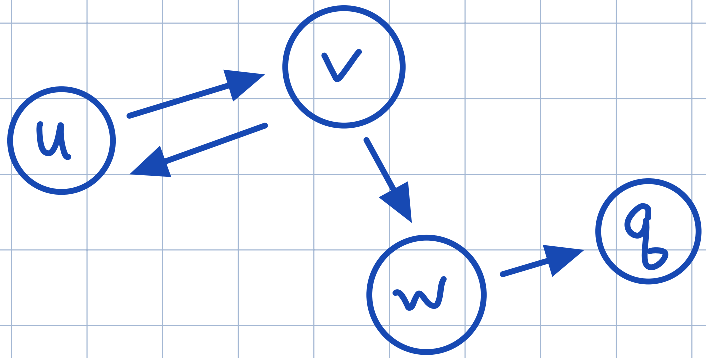
$$\frac{}{edge_2(u,v)} \qquad \frac{}{edge_2(v,u)}$$

$$\frac{}{edge_2(v,w)} \qquad \frac{}{edge(w,q)}$$

$$\frac{edge_2(x,y)}{reach_2(x,y)}$$

$$\frac{edge_2(x,y) \quad reach_2(y,z)}{reach_2(x,z)}$$

Limitations



$$\frac{}{edge_2(u,v)} \qquad \frac{}{edge_2(v,u)}$$

$$\frac{}{edge_2(v,w)} \qquad \frac{}{edge(w,q)}$$

$$\frac{edge_2(x,y)}{reach_2(x,y)}$$

$$\frac{edge_2(x,y) \quad reach_2(y,z)}{reach_2(x,z)}$$

No <u>generic</u> transitive closure

Design Question

Is there a version of Datalog which has better facilities for abstraction?