CS 303, Paul Cao, Lecture 31

Plan: transitive closure, Johnson's algorithm

Review

- Floyd-Marshall's algorithm for all pair shortest path- $O(n^3)$ using matrix operations

Warshall's algorithm for transitive closure

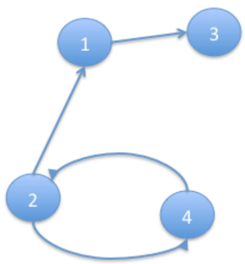
Definition:

The transitive closure of a directed graph with n vertices can be defined as the n-by-n matrix $T=\{t_{ij}\}$, in which the element in the i^{th} row and j^{th} column is 1 if there exists a non-trivial directed path (i.e. a directed path of a positive length) from the i^{th} vertex to the j^{th} vertex. Otherwise, t_{ij} is zero.

It basically tells us the existence of all nontrivial paths in a digraph.

Its format is kind of similar to adjacency matrix.

e.g.



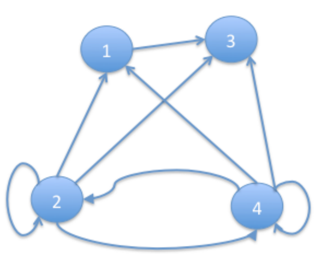
$$\begin{array}{c|c} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \hline \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

Adjacency Matrix

$$\begin{array}{c|c} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \hline \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

Transitive Closure

e.g.



$$\begin{array}{c|c} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \hline \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

Adjacency Matrix

$$\begin{array}{c|c} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \hline \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

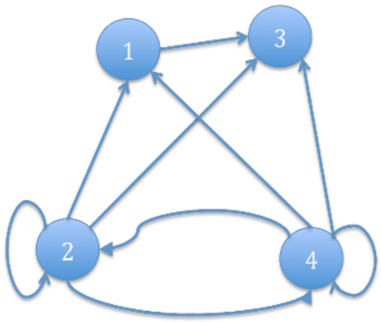
Transitive Closure

Q:/ How to find the transitive closure of a graph?

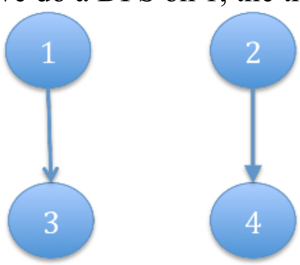
A:/ Solution1:

- Use DFS or BFS for each vertex, i.e. vertex i.
- Use the search tree to indicate which vertices are reachable from vertex i

e.g.

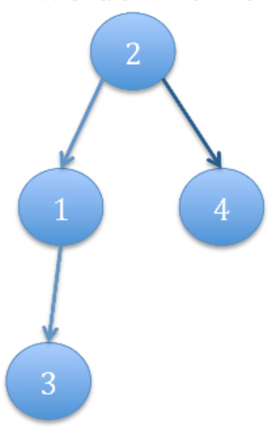


if we do a DFS on 1, the tree we have is



Thus 1 can only go to 3

if we do another DFS on 2, the tree we have is



Thus 2 can go to 1, 3, and 4

Complexity: $|V|*\Theta(|V|^2)=\Theta(|V|^3)$

Solution 2: Warshall's algorithm based on DP

Idea:

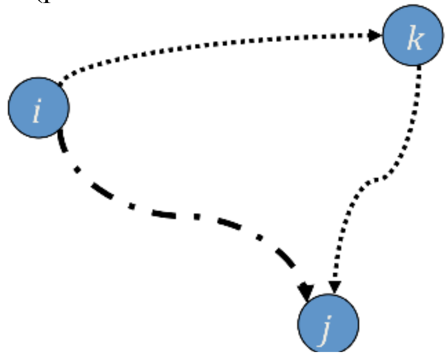
Important observation: If there is a path from a to z via s then there must be a path from a to s and from s to z
Let $R^{(k)}$ be the optimal answer when we only allow the first k nodes to be intermediate nodes in paths. We can compute the optimal solution for k + 1 nodes $R^{(k+1)}$ efficiently. In here, k is the number in the node.

So we can constructs transitive closure T as the last matrix in the sequence of n-by-n matrices $R^{(0)}, \ldots, R^{(k)}, \ldots, R^{(n)}$ where

$R^{(k)}[i,j] = 1$ iff there is nontrivial path from i to j with only first k vertices allowed as intermediate
Q:/ What is $R^{(0)}$?
A:/ It is the adjacency matrix

Q:/ What is $R^{(n)}$?
A:/ It is the transitive closure

As k goes to k+1, the recurrence can be set up as the following
On the k-th iteration, the algorithm determines for every pair of vertices i, j if a path exists from i and j with just vertices 1,…,k allowed as intermediate
$R^{(k)}[i,j]$
1. $= R^{(k-1)}[i,j]$ (path using just 1 ,…,k-1)
   or
2. $= R^{(k-1)}[i,k]$ and $R^{(k-1)}[k,j]$ (path from i to k and from k to i using just 1 ,…,k-1)



where the dash represent the first scenario and the dotted line represent the second scenario.

With this idea, recurrence relating elements $R^{(k)}$ to elements of $R^{(k-1)}$ is: $R^{(k)}[i,j] = R^{(k-1)}[i,j]$ or ($R^{(k-1)}[i,k]$ and $R^{(k-1)}[k,j]$)

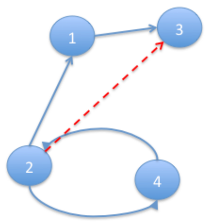
It implies the rules above for generating $R^{(k)}$ from $R^{(k-1)}$:
- Rule 1 If an element in row i and column j is 1 in $R^{(k-1)}$, it remains 1 in $R^{(k)}$
- Rule 2 If an element in row i and column j is 0 in $R^{(k-1)}$, it has to be changed to 1 in $R^{(k)}$ if and only if the element in its row i and column k and the element in its column j and row k are both 1's in $R^{(k-1)}$

e.g.
Find the transitive closure of the following graph

Iteration 1

$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 0 & 1 & 0 & 0 \end{array}$$

$R^{(0)}$ **(Adjacency Matrix)**



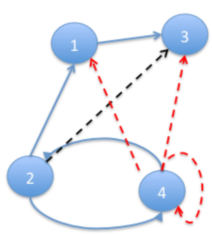
$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 0 & 1 & 0 & 0 \end{array}$$

$R^{(1)}$ **(with intermediate vertices numbered no higher than 1, i.e. just vertex 1)**

Iteration 2

$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 0 & 1 & 0 & 0 \end{array}$$

$R^{(1)}$ **(with intermediate vertices numbered no higher than 1, i.e. just vertex 1)**



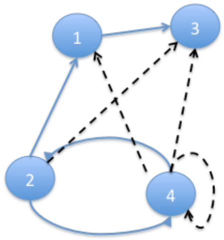
$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 1 & 1 & 1 & 1 \end{array}$$

$R^{(2)}$ **(with intermediate vertices numbered no higher than 2, i.e. vertices 1 and 2)**

Iteration 3

$$\begin{array}{cc} & \begin{array}{cccc}1&2&3&4\end{array}\\ \begin{array}{c}1\\2\\3\\4\end{array} & \begin{bmatrix}0&0&1&0\\1&0&1&1\\0&0&0&0\\1&1&1&1\end{bmatrix}\end{array}$$

$R^{(2)}$ (with intermediate vertices numbered no higher than 2, i.e. vertices 1 and 2)



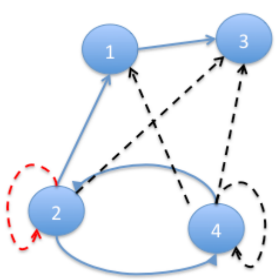
$$\begin{array}{cc} & \begin{array}{cccc}1&2&3&4\end{array}\\ \begin{array}{c}1\\2\\3\\4\end{array} & \begin{bmatrix}0&0&1&0\\1&0&1&1\\0&0&0&0\\1&1&1&1\end{bmatrix}\end{array}$$

$R^{(3)}$ (with intermediate vertices numbered no higher than 3, i.e. vertices 1,2 and 3)

Iteration 4

$$\begin{array}{cc} & \begin{array}{cccc}1&2&3&4\end{array}\\ \begin{array}{c}1\\2\\3\\4\end{array} & \begin{bmatrix}0&0&1&0\\1&0&1&1\\0&0&0&0\\1&1&1&1\end{bmatrix}\end{array}$$

$R^{(3)}$ (with intermediate vertices numbered no higher than 3, i.e. vertices 1,2 and 3)



$$\begin{array}{cc} & \begin{array}{cccc}1&2&3&4\end{array}\\ \begin{array}{c}1\\2\\3\\4\end{array} & \begin{bmatrix}0&0&1&0\\1&1&1&1\\0&0&0&0\\1&1&1&1\end{bmatrix}\end{array}$$

$R^{(4)}$ (with intermediate vertices numbered no higher than 3, i.e. vertices 1,2,3,4)
It is the transitive closure

The overall algorithm is

ALGORITHM *Warshall*$(A[1..n, 1..n])$

```
//Implements Warshall's algorithm for computing the transitive closure
//Input: The adjacency matrix A of a digraph with n vertices
//Output: The transitive closure of the digraph
R^(0) ← A
for k ← 1 to n do
    for i ← 1 to n do
        for j ← 1 to n do
            R^(k)[i, j] ← R^(k-1)[i, j] or (R^(k-1)[i, k] and R^(k-1)[k, j])
return R^(n)
```

Efficiency:

Time efficiency: $\Theta(n^3) \rightarrow$ same as DFS traversals but very very succinct

Space efficiency: Matrices can be written over their predecessors

Topic 2: Johnson's algorithm

Johnson's algorithm

Idea: graph re-weighting and try to use Dijkstra $|V|$ times

Graph reweighting:

Given function $h: V \to \mathbb{R}$, reweight each edge $(u,v) \in E$ by

$$w_h(u,v) = w(u,v) + h(u) - h(v)$$

then for any $u, v \in V$, all paths from u to v are reweighted by the same amount.

Proof: Let P be $\overset{(u)}{v_1} \to v_2 \to v_3 \to \cdots \to \overset{(v)}{v_k}$ a path from u to v

$$\begin{aligned} w_h(p) &= \sum_{i=1}^{k-1} w_h(v_i, v_{i+1}) \\ &= \sum_{i=1}^{k-1} \left(w(v_i, v_{i+1}) + h(v_i) - h(v_{i+1}) \right) \\ &= \sum_{i=1}^{k-1} w(v_i, v_{i+1}) + h(u) - h(v) \\ &= w(u,v) + h(u) - h(v) \qquad \# \end{aligned}$$

Corollary: $\delta_h(u,v) = \delta(u,v) + h(u) - h(v)$

Thus we'll try to find $h: V \to \mathbb{R}$ s.t.

$$w_h(u,v) \geq 0 \text{ for all } (u,v) \in E$$

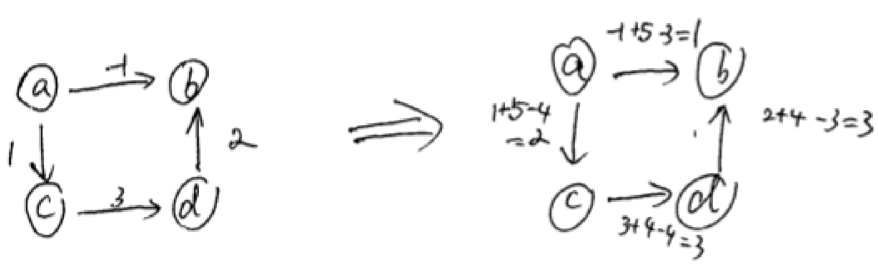
Then run Dijkstra to get $\delta_h(u,v)$, then get $\delta(u,v)$

i.e. $w(u,v) + h(u) - h(v) \geq 0$

Solve a system of linear equations $\begin{cases} h(v) - h(u) \leq w(u,v) \\ \vdots \end{cases}$

then we are done.

e.g.



$$\begin{cases} h(b) - h(a) \leq -1 \\ h(a) - h(c) \leq 1 \\ h(d) - h(c) \leq 3 \\ h(b) - h(d) \leq 2 \end{cases} \xRightarrow{\text{one solution}} \begin{cases} h(a) = 5 \\ h(b) = 3 \\ h(c) = 4 \\ h(d) = 4 \end{cases}$$

If we run Dijkstra on a on the new graph

$a \to d = \underset{w_h(a,d)}{\underline{5}} \longrightarrow$ distance on the old graph is

$$w_h(a,d) - h(a) + h(d)$$
$$= 5 - 5 + 4 = 4$$