Warshall's algorithm

This is a more efficient method for determining the reachability of all node pairs in a graph (directed or undirected).

For a graph $G$ with $n$ nodes $1, \ldots, n$, the method constructs a sequence of $n$ graphs (or adjacency matrices) $P_1, \ldots, P_n$. Each one uses the same node set. We'll describe the method for simple undirected graphs. (There is no point to considering loops because the presence of loops does not alter whether or not different nodes are connected by a path.)

To start, set $P_0 = G$.

Suppose $P_k$ has been defined. $P_{k+1}$ contains all the edges of $P_k$, plus any additional edges needed to make sure that

- in the undirected case: every pair of nodes joined by an edge of $P_k$ to node $k+1$ are joined to each other by an arc of $P_{k+1}$
- in the directed case: for every path $a \rightarrow k+1 \rightarrow b$, $(a,b)$ is an edge of $P_{k+1}$.

The method terminates after $n$ iterations: $P_n$ contains all adjacency relationships represented as edges.

Proof that the algorithm works: By induction, we will show that for each $k$, $P_k$ has this property (induction hypothesis):

$P_k$ contains all the edges of the original graph, and

- For any nodes $u$, $v$, if there exists a path of the original graph from $u$ to $v$ using only nodes in the set $\{1, \ldots, k\}$, then $(u,v)$ is an edge of $P_k$.

The property is clear when $k=1$.

Suppose the property has been established for $k$, and consider $P_{k+1}$. Let $u,v$ be nodes that are joined by a path using only nodes in the set $\{1, \ldots, k+1\}$. If this path does not use $k+1$, then $(u,v)$ is an edge of $P_k$ (by the induction hypothesis) and hence is also an edge of $P_{k+1}$. Suppose this path does use $k+1$ and contains no duplicate vertices.

Write the path as

$$u = n_1, \ldots, n_r, k+1, n_s, \ldots, n_t = v$$

$u$ is connected to $k+1$ by a path that uses nodes only in $1, \ldots, k$, so $(u,k+1)$ is an edge of $P_k$.

$v$ is connected to $k+1$ by a path that uses nodes only in $1, \ldots, k$, so $(v,k+1)$ is an edge of $P_k$.

Therefore $(u,v)$ is an edge of $P_{k+1}$.

Since the property holds for all $k$, it holds in particular for $k=n$, so in the final graph $P_n$ there is an edge joining every pair of nodes that are connected by a path using any nodes.

homework: p224- (section 3.3) 1-8, 25, 26