Explanation of Algorithm

- DFS numbers of all vertices initialized to 0.
  - It serves as vertices marked "unvisited"
  - Assigning DFS number to a vertes implies it is marked "visited"
- It then selects an arbitrary vertex $v = x$, builds a DFST, and computes LOW[v] in one single pass.
- An edge leading to a new vertex is a tree edge.
- The tree edge is pushed on to the stack before making the recursive call.

Explanation of Algorithm

- On return from call LOW[v] is computed.
- If v is an articulation point then its DFS number must be greater than or equal to LOW point of the child.
- At this point, all the edges up to and including v, w are output as a biconnected component.
- If w is a old vertex (DFS number is nonzero) then it is a back edge.
- Only back edge to a proper ancestor of the parent ($w \neq u$) would lower the LOW point.

Correctness of Algorithm

Lemma (*Correctness of LOW point computation*)

When Search(w) procedure completes, the edges in the stack above (v, w) are the edges in the same biconnected components as (v, w).

- To prove it, we use induction on the number of biconnected components, b.
- If there is just one biconnected components, i.e., $b = 1$, it is trivial.
  - In this case, there is no articulation point.
  - So all the edges of G will be on the stack.

Correctness of Algorithm

- Induction hypothesis: assume this to be true for all graphs having b biconnected components.
- Let G have $b+1$ biconnected components.
- Consider the first Bicon(w, v) call that ends with LOW[w] $\geq$ dfn[v] for a tree edge (v, w).
- No edge has been removed yet from the STACK.
- Since LOW[w] $\geq$ dfn[v], all the set of edges above (v, w) are incident on the descendants of w and first block B_1 has been detected.
- So the edges on the STACK above (v, w) are exactly the edges in the same biconnected component as (v, w).

Correctness of Algorithm

- Now after removing edges of B_1, the algorithm works on induced graph $G' = G - B_1$, in exactly the same way as it had worked on graph G.
- But G' has b biconnected components.
- By induction, algorithm should correctly obtain all b biconnected components of G'.

Notion of Connectedness in Directed Graphs

- For every pair of vertices u, v there exists a path.
- But path here means directed path.
- Connectivity in directed graph implies both $u \rightsquigarrow v$ and $v \rightsquigarrow u$.
- In general, it is possible that path $u \rightsquigarrow v$ may exist but $v \not\rightsquigarrow u$.
- Or even u and v are not reachable from each other.
- There is also a notion of weak connectivity: it possible to reach any vertex from any other vertex by traversing edges in some direction (ignoring direction).
- It essentially means every vertex has either indegree or outdegree of at least 1.

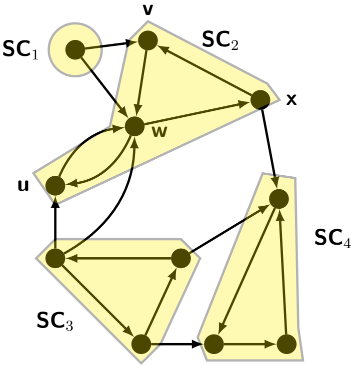
Strong Connected Components (SCC)

Definition (*SCC)*

Let $G = (V, E)$ be directed graph. G is strongly connected iff for every pair of vertices v, w, there is a directed path from v to w and also a directed path from w to v.

- SCC is an equivalence relation.
- If u, v are in same SCC then uRv and vRu where R: there exists a directed path.
- If uRv and vRw then obviously, uRw.
- Collapsing each SCC to a vertex we get a condensation graph which is a DAG.

Strong Connected Components

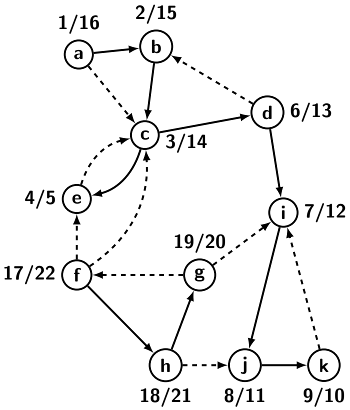


- There are four SCCs.
- One SCC has just one vertex.
- Consider $u, v \in SC_2$.
- There exists pair of paths: $u \to w \to x \to v$ and $v \to w \to u$.

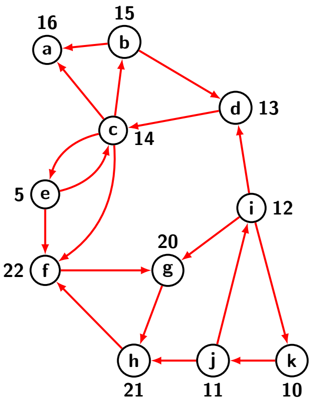
Strong Connected Components

- Compute Finish time f[u] for $u \in V$.
- Reverse each edge of G to obtain G^T.
- Call DFS on G^T but apply it decreasing order of finish time f[u] as computed in first DFS.
- Output the vertices of each tree in DFS forest of G^T as a separate SCC.

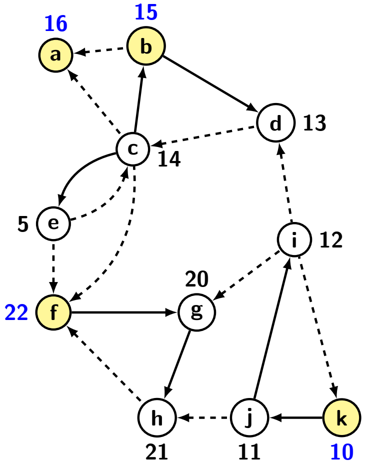
Original Graph G



Transformed Graph G^T (by reversing edge directions).



Strong Connected Components



- Start vertices in decreasing order of finish time: f, a, b, and k.
- DFS from from these vertices discover SCCs:
  - f: $\{f, g, h\}$
  - a: $\{a\}$
  - b: $\{b, c, d, e\}$
  - k: $\{i, j, k\}$

Correctness of SCC Algorithm

- Let $d[v]$: DFS discovery time, and
- Let $f[v]$: DFS finish time.
- Let S be subset of V.

$$d[S] = \min_{u \in S} d[u], \text{ and } f[S] = \max_{u \in S} f[u]$$

Lemma

Let S_1 and S_2 be two distinct SCCs. If $\exists$ an edge $(u, v) \in E$, where $u \in S_1$ and $v \in S_2$ then $f[S_1] > f[S_2]$

Correctness of SCC Algorithm

Proof.

- Case 1 ($d[S_1] < d[S_2]$): Let x be first vertex in S_1 to be discovered.
- At this time none of the vertices in S_1 and S_2 have been marked "visited"
- For any vertex $w \in S_2$, $\exists$ a path from x to w.
- So, all vertices in S_2 are descendants of x
- Therefore, $f[x] = f[S_2] < f[S_1]$

$\square$

Correctness of SCC Algorithm

Proof continues.

- Case 2 ($d[S_1] > d[S_2]$): Let y be the first vertex discovered in S_2.
- All vertices in S_2 are descendant of y.
- Therefore, by definition $f[y] = f[S_2]$.
- Since, S_1 and S_2 are distinct SCC there cannot be path from any vertex of S_2 to a vertex of S_1.
- So all vertices in S_1 are "unvisited" when DFS of S_2 is complete.
- Therefore, $f[S_2] < f[S_1]$.

□

Correctness of SCC Algorithm

A corollary of the above lemma is on G^T where the inequalities are reversed is as follows:

Lemma (*Corollary I*)

Let S_1 and S_2 be two distinct SCCs. If $\exists$ an edge $(u, v) \in E^T$ where $u \in S_1$ and $v \in S_2$ then $f[S_2] < f[S_1]$.

Proof.

- Since every edge is reversed in E^T, it implies $(v, u) \in E$.
- Therefore, $f[S_1] > f[S_2]$ from the previous lemma.

$\square$

Correctness of SCC Algorithm

Another corollary that follows from the above result is:

Lemma (*Corollary II*)

If $f[S_2] > f[S_1]$, then there cannot be any edge (u, v), where $u \in S_2$ to $v \in S_1$ in G^T.

- Use of decreasing finish time in exploring G^T is the intuition behind the algorithm's correctness.

Correctness of SCC Algorithm

- Consider a pair of vertices v and w.
- Let finish time of w is smaller than v: $f[w] < f[v]$.
- Now we consider three cases:
  - Case 1: $w \not\rightsquigarrow v$ (v is not reachable from w) then v and w are in different SCCs, and correctness of algorithm is not affected.
  - Case 2: $w \rightsquigarrow v$ (v is reachable from w) and w is not in DFS subtree of v in G^T.
  - Case 3: w is in DFS subtree of v.
- Therefore, we consider last two cases in more details.

Correctness of SCC Algorithm

- Consider case 2:
  - Since $d[w] < d[v]$ in DFS of G, w cannot be an ancestor of v.
  - Furthermore, there cannot be of a cross link on $w \rightsquigarrow v$ path, because the finish time of a cross link's source vertex is greater than the finish time of its end vertex.
  - So, case 2 is not possible, leaving case 3 as the only one possibility.

Correctness of SCC Algorithm

- Consider case 3:
  - When DFS enters a vertex v in G^T, all the vertices in SCCs with lower finish times remain unvisited.
  - And if there is a path $w \rightsquigarrow v$, it only means $w \in T_v$.
- So, proof is complete by noticing that for every two distinct components S_1 and S_2, if $f[S_1] > f[S_2]$ there cannot be an edge from S_1 to S_2 in G^T.
- The above fact follows directly from the Corollary II proved earlier.
- This implies w and v both belong to same SCC.

Summary of Graph Algorithms

- Graph terminology and representation were introduced.
- BFS, and DFS search were explained.
- We learnt about a number of applications of DFS, particularly in computing:
  - Connected components, biconnected components of a undirected graphs, and
  - Strong connected components of directed graph and topological sorting of a DAG

Summary of Graph Algorithms

- With reference to weighted graphs we talked about the problems of determining shortest paths, namely,
  - Dijkstra and Bellman-Ford algorithms using edge relaxation operation.
  - Prim and Kruskal's algorithms for finding MSTs.
- We also looked at almost linear time UNION and FIND operations on disjoint sets in the context of Kruskal's algorithm.
- However, a detailed analysis of UNION-FIND algorithms was not discussed recognizing involved mathematical complications.