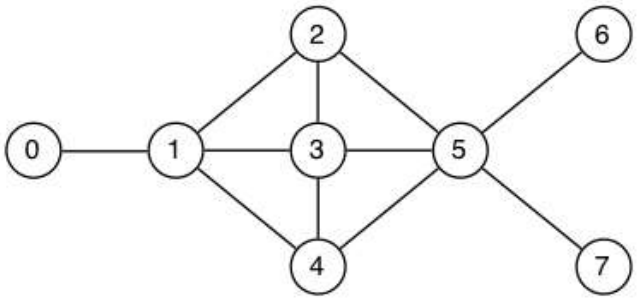
COMP9024 18x1 Data Structures and Algorithms

Week 04 Problem Set
Graph Traversal, Digraphs

1. (Graph traversal: DFS and BFS)

   Show the order in which the nodes of the graph depicted below are visited by

   a. DFS starting at node 0
   b. DFS starting at node 3
   c. BFS starting at node 0
   d. BFS starting at node 3



   Assume the use of a stack for depth-first search (DFS) and a queue for breadth-first search (BFS), respectively. Show the state of the stack or queue explicitly in each step. When choosing which neighbour to visit next, always choose the smallest unvisited neighbour.

   Answer:

   a. DFS starting at 0:

```
Current    Stack (top at left)
-          0
0          1
1          2 3 4
2          5 3 4
5          6 7 3 4
6          7 3 4
7          3 4
3          4
4          -
```

   b. DFS starting at 3:

```
Current    Stack (top at left)
-          3
3          1 2 4 5
1          0 2 4 5
0          2 4 5
2          4 5
4          5
5          6 7
6          7
7          -
```

   c. BFS starting at 0:

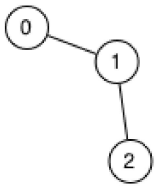
```
Current    Queue (front at left)
-          0
0          1
1          2 3 4
2          3 4 5
3          4 5
4          5
5          6 7
6          7
7          -
```

d. BFS starting at 3:

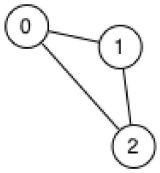
```
Current    Queue (front at left)
-          3
3          1 2 4 5
1          2 4 5 0
2          4 5 0
4          5 0
5          0 6 7
0          6 7
6          7
7          -
```

2. (Hamiltonian/Euler paths and circuits)

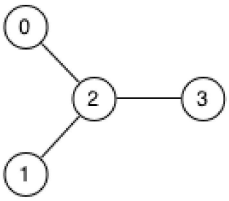
a. Identify any Hamiltonian/Euler paths/circuits in the following graphs:



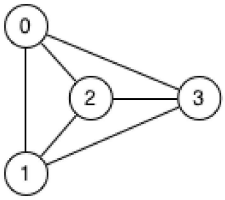
Graph 1



Graph 2

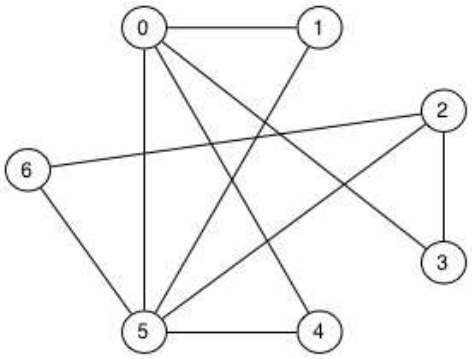


Graph 3



Graph 4

b. Find an Euler path and an Euler circuit (if they exist) in the following graph:



Answer:

a. Graph 1: has both Euler and Hamiltonian paths (e.g. 0-1-2), but cannot have circuits as there are no cycles.

Graph 2: has both Euler (e.g. 0-1-2-0) and Hamiltonian paths (e.g. 0-1-2); also has both Euler and Hamiltonian circuits (e.g. 0-1-2-0).

Graph 3: has neither Euler nor Hamiltonian paths, nor Euler nor Hamiltonian circuits.

Graph 4: has Hamiltonian paths (e.g. 0-1-2-3) and a Hamiltonian circuits (e.g. 0-1-2-3-0); it has neither an Euler path nor an Euler circuit.

b. An Euler path: 2-6-5-2-3-0-1-5-0-4-5

No Euler circuit since two vertices (2 and 5) have odd degree.

3. (Cycle check)

Design an algorithm to check for the existence of a cycle in a graph using depth-first search.

Answer:

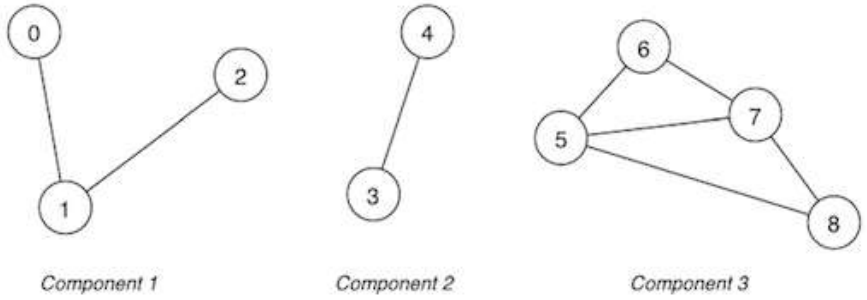
```
hasCycle(G):
|  Input  graph G
```

```
|  Output true if G has a cycle, false otherwise
|
|  mark all vertices as unvisited
|  for each vertex v∈G do          // make sure to check all connected components
|  |  if v has not been visited then
|  |     if dfsCycleCheck(G,v,v) then
|  |        return false
|  |     if dfsCycleCheck(G,v,v) then
|  |        return true
|  |     end if
|  |  end if
|  end for
|  return true
|  return false

dfsCycleCheck(G,v,u):      // look for a cycle that does not go back directly to u
|  mark v as visited
|  for each (v,w)∈edges(G) do
|  |  if w has not been visited then
|  |     return dfsCycleCheck(G,w,v)
|  |  else if w≠u then
|  |     return true
|  |  end if
|  end for
|  return false
```

4. (Connected components)

   a. Write a C program that computes the connected components in a graph. The graph should be built from user input in the same way as in exercise 5 last week (i.e., problem set week 6). Your program should use the Graph ADT (`Graph.h`, `Graph.c`) from the lecture. These files should not be changed.

   An example of the program executing is shown below for the following graph:




```
prompt$ ./components
Enter the number of vertices: 9
Enter an edge (from): 0
Enter an edge (to): 1
Enter an edge (from): 1
Enter an edge (to): 2
Enter an edge (from): 4
Enter an edge (to): 3
Enter an edge (from): 6
Enter an edge (to): 5
Enter an edge (from): 6
Enter an edge (to): 7
Enter an edge (from): 5
Enter an edge (to): 7
Enter an edge (from): 5
Enter an edge (to): 8
Enter an edge (from): 7
Enter an edge (to): 8
Enter an edge (from): done
Finished.
Number of components: 3
```

```
Component 1:
0
1
2
Component 2:
3
4
Component 3:
5
6
7
8
```

Note that:

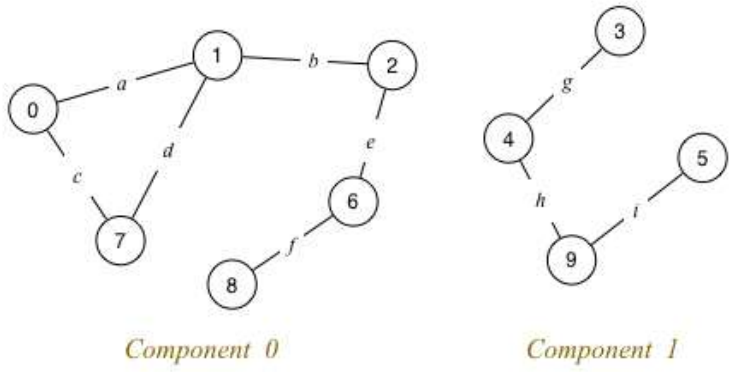
- the vertices within a component are printed in ascending order
- the components themselves are output in ascending order of their smallest node.

You may assume that a graph has a maximum of 1000 nodes.

We have created a script that can automatically test your program. To run this test you can execute the `dryrun` *program that corresponds to the problem set and week. It expects to find a program named* `components.c` *in the current directory. You can use dryrun as follows:*

```
prompt$ ~cs9024/bin/dryrun prob04
```

b. Consider the following graph with multiple components:



Assume a vertex-indexed connected components array `cc[0..nV-1]` that might form part of the `Graph` representation structure for this graph as suggested in the lecture:

```
cc[] = {0,0,0,1,1,1,0,0,0,1}
```

Show how the `cc[]` array would change if

1. edge *d* was removed
2. edge *b* was removed

c. Consider an adjacency matrix graph representation augmented by the two fields

- `nC` (number of connected components)
- `cc[]` (connected components array)

These fields are initialised as follows:

```
newGraph(V):
|  Input  number of nodes V
|  Output new empty graph
|
|  g.nV=V, g.nE=0, g.nC=V
|  allocate memory for g.edges[][]
|  for all i=0..V-1 do
|     g.cc[i]=i
```

```
|     for all j=0..V-1 do
|        g.edges[i][j]=0
|     end for
|  end for
|  return g
```

Modify the pseudo-code for edge insertion and edge removal from the lecture (week 6) to maintain the two new fields.

Answer:

a. The following two functions together implement the algorithm from the lecture that uses the following strategy:

- pick a not-yet-visited vertex
- find all vertices reachable from that vertex
- increment the count of connected components
- repeat above until all vertices have been visited

```
#define MAX_NODES 1000
int componentOf[MAX_NODES];

void dfsComponents(Graph g, int nV, int v, int id) {
   componentOf[v] = id;
   Vertex w;
   for (w = 0; w < nV; w++)
      if (adjacent(g, v, w) && componentOf[w] == -1)
         dfsComponents(g, nV, w, id);
}

// computes the connected component array
// and returns the number of connected components
int components(Graph g, int nV) {
   Vertex v;
   for (v = 0; v < nV; v++)
      componentOf[v] = -1;

   int compID = 0;
   for (v = 0; v < nV; v++) {
      if (componentOf[v] == -1) {
         dfsComponents(g, nV, v, compID);
         compID++;
      }
   }
   return compID;
}
```

Calling the function and printing the result:

```
int i, c = components(g, nV);
printf("Number of connected components: %d\n", c);
for (i = 0; i < c; i++) {
   printf("Component %d:\n", i+1);
   Vertex v;
   for (v = 0; v < nV; v++)
      if (componentOf[v] == i)
         printf("%d\n", v);
}
```

b. After removing *d*, `cc[] = {0,0,0,1,1,1,0,0,0,1}` (i.e. unchanged)
After removing *b*, `cc[] = {0,0,2,1,1,1,2,0,2,1}`

c. Inserting an edge may reduce the number of connected components:

```
insertEdge(g,(v,w)):
|  Input graph g, edge (v,w)
|
|  if g.edges[v][w]=0 then              // (v,w) not in graph
|  |  g.edges[v][w]=1, g.edges[w][v]=1   // set to true
```

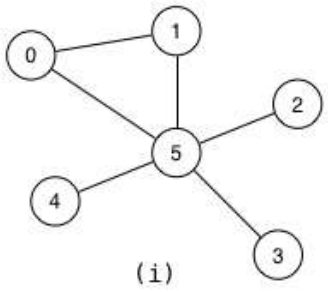
```
|  |  g.nE=g.nE+1
|  |  if g.cc[v]≠g.cc[w] then            // v,w in different components
|  |  |  c=min{g.cc[v],g.cc[w]}          // ⇒ merge components c and d
|  |  |  d=max{g.cc[v],g.cc[w]}
|  |  |  for all vertices v∈g do
|  |  |     if g.cc[v]=d then
|  |  |        g.cc[v]=c                 // move node from component d to c
|  |  |     else if g.cc[v]=g.nC-1 then
|  |  |        g.cc[v]=d                 // replace largest component ID by d
|  |  |     end if
|  |  |  end for
|  |  |  g.nC=g.nC-1
|  |  end if
|  end if
```

Removing an edge may increase the number of connected components:

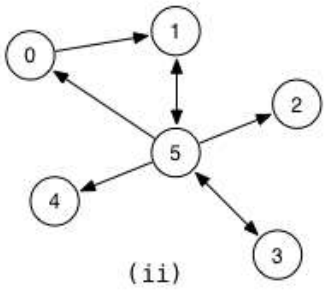
```
removeEdge(g,(v,w)):
|  Input graph g, edge (v,w)
|
|  if g.edges[v][w]≠0 then               // (v,w) in graph
|  |  g.edges[v][w]=0, g.edges[w][v]=0   // set to false
|  |  if ¬hasPath(g,v,w) then            // v,w no longer connected
|  |     dfsNewComponent(g,v,g.nC)       // ⇒ put v + connected vertices into new component
|  |     g.nC=g.nC+1
|  |  end if
|  end if

dfsNewComponent(g,v,componentID):
|  Input graph g, vertex v, new componentID for v and connected vertices
|
|  g.cc[v]=componentID
|  for all vertices w adjacent to v do
|     if g.cc[w]≠componentID then
|        dfsNewComponent(g,w,componentID)
|     end if
|  end if
```

5. (Digraphs)

   a. For each of the following graphs:



(i)



(ii)

   Show the concrete data structures if the graph was implemented via:

   - adjacency matrix representation (assume full $V \times V$ matrix)
   - adjacency list representation (if non-directional, include both *(v,w)* and *(w,v)*)

   b. Consider the following map of streets in the Sydney CBD:

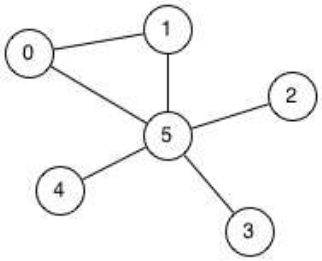
Represent this as a directed graph, where intersections are vertices and the connecting streets are edges. Ensure that the directions on the edges correctly reflect any one-way streets (this is a driving map, not a walking map). You only need to make a graph which includes the intersections marked with red letters Some things that don't show on the map: Castlereagh St is one-way heading south and Hunter St is one-way heading west.

For each of the following pairs of intersections, indicate whether there is a path from the first to the second. Show a path if there is one. If there is more than one path, show two different paths.

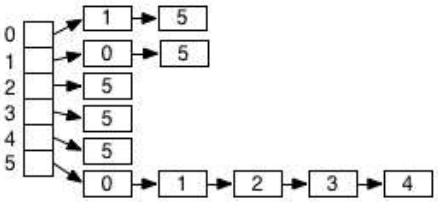
1. from intersection "D" on Margaret St to insersection "L" on Pitt St
2. from intersection "J" to the corner of Margaret St and York St (intersection "A")
3. from the intersection of Margaret St and York St ("A") to the intersection of Hunter St and Castlereagh St ("M")
4. from intersection "M" on Castlereagh St to intersection "H" on York St

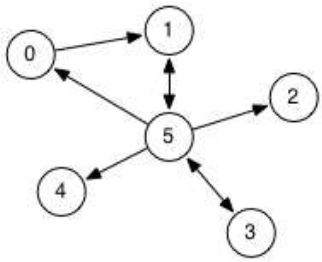
Answer:

a.

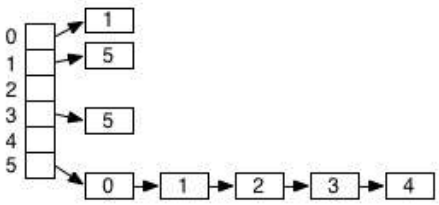


	0	1	2	3	4	5
0	0	1	0	0	0	1
1	1	0	0	0	0	1
2	0	0	0	0	0	1
3	0	0	0	0	0	1
4	0	0	0	0	0	1
5	1	1	1	1	1	0

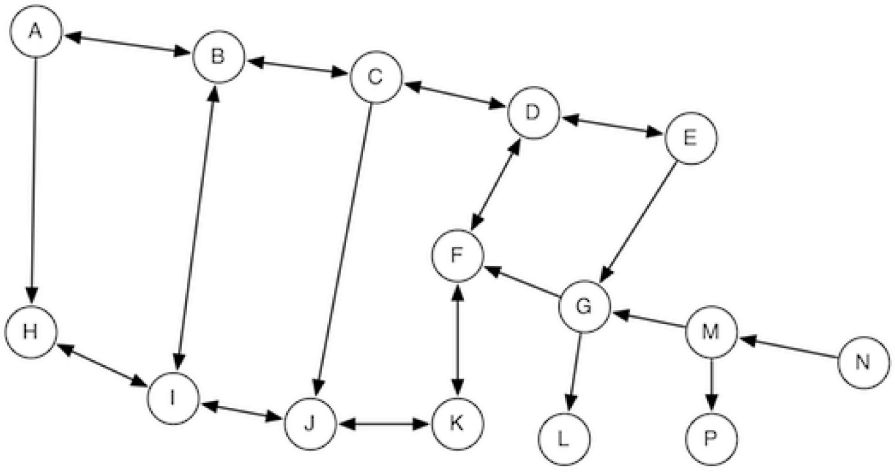




	0	1	2	3	4	5
0	0	1	0	0	0	0
1	0	0	0	0	0	1
2	0	0	0	0	0	0
3	0	0	0	0	0	1
4	0	0	0	0	0	0
5	1	1	1	1	1	0



b. The graph is as follows. Bi-directional edges are depicted as two-way arrows, rather than having two separate edges going in opposite directions.

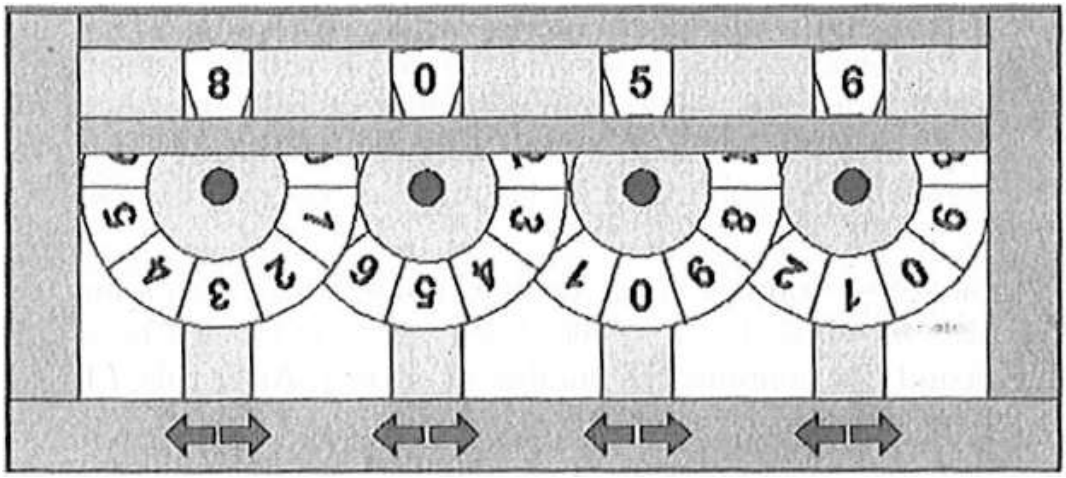


For the paths:

1. D → E → G → L and there are no other choices that don't involve loops through D
2. J → I → B → A or J → K → F → D → C → B → A
3. You can't reach M (or N or P) from A on this graph. Real-life is different, of course.
4. M → G → F → D → C → B → A → H or M → G → F → K → J → I → H

6. **Challenge Exercise**

Consider a machine with four wheels. Each wheel has the digits 0…9 printed clockwise on it. The current *state* of the machine is given by the four topmost digits `abcd`, e.g. 8056 in the picture below.



Each wheel can be controlled by two buttons: Pressing the button labelled with "←" turns the corresponding wheel clockwise one digit ahead, whereas presssing "→" turns it anticlockwise one digit back.

Write a C-program to determine the *minimum* number of button presses required to transform

- a given initial state, `abcd`
- to a given goal state, `efgh`
- without passing through any of $n \geq 0$ given "forbidden" states, `forbidden[]={`$w_1x_1y_1z_1,\ldots,w_nx_ny_nz_n$`}`.

For example, the state 8056 as depicted can be transformed into 0056 in 2 steps if `forbidden[]={}`, whereas a minimum of 4 steps is needed for the same task if `forbidden[]={9056}`. (Why?)

Use your program to compute the least number of button presses required to transform 8056 to 7012 if

a. there are no forbidden states
b. you are not permitted to pass through any state 7055–8055 (i.e., 7055, 7056, …, 8055 all are forbidden)
c. you are not permitted to pass through any state 0000–0999 or 7055–8055

Answer:

The problem can be solved by a breadth-first search on an undirected graph:

- nodes 0…9999 correspond to the possible configurations 0000 – 9999 of the machine;
- edges connect nodes v and w if, and only if, one button press takes the machine from v to w and vice versa.

The following code implements BFS on this graph with the help of the integer queue ADT from the lecture (`queue.h`, `queue.c`). Note that the graph is built *implicitly* : nodes are generated as they are encountered during the search.

```
#include "queue.h"

int shortestPath(int source, int target, int forbidden[], int n) {
   int visited[10000];

   int i;
   for (i = 0; i <= 9999; i++)
      visited[i] = -1;             // mark all nodes as unvisited
   for (i = 0; i < n; i++)
      visited[forbidden[i]] = -2;  // mark forbidden nodes as visited => they won't be selected
   visited[source] = source;

   queue Q = newQueue();
   QueueEnqueue(Q, source);
   bool found = (target == source);
   while (!found && !QueueIsEmpty(Q)) {
      int v = QueueDequeue(Q);
      int wheel, turn;
      for (wheel = 10; wheel <= 10000; wheel *= 10) {   // fancy way of generating the
         for (turn = 1; turn <= 9; turn += 8) {         // eight neighbour configurations of v
            int w = wheel * (v / wheel) + (v % wheel + (wheel/10) * turn) % wheel;
            if (visited[w] == -1) {
               visited[w] = v;
               if (w == target)
                  found = true;
               else
                  QueueEnqueue(Q, w);
            }
         }
      }
   }
   dropQueue(Q);
   if (found) {                    // unwind path to determine length
      int length = 0;
      while (target != source) {
         target = visited[target];  // move to predecessor on path
         length++;
      }
      return length;
   } else {
      return -1;                   // no solution
   }
}
```

a. 9
b. 17
c. ∞ (unreachable)