Graph Basics

- Graphs
- Properties of Graphs
- Graph Terminology

COMP2521 20T2 \Diamond Graph Basics [0/15]

>>

Graphs

Many applications require

- a collection of items (i.e. a set)
- relationships/connections between items

Examples:

- maps: items are cities, connections are roads
- web: items are pages, connections are hyperlinks

Collection types you're familiar with

- lists ... linear sequence of items (COMP1511)
- trees... branched hierarchy of items (Weeks 02/03)

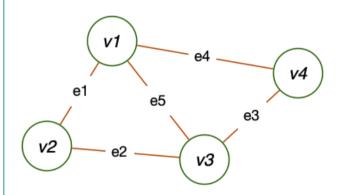
Graphs are more general ... allow arbitrary connections

COMP2521 20T2 \diamondsuit Graph Basics [1/15]

A graph G = (V,E)

- V is a set of vertices
- E is a set of edges (subset of $V \times V$)

Example:



$$V = \{ v1, v2, v3, v4 \}$$

$$E = \{ e1, e2, e3, e4, e5 \}$$

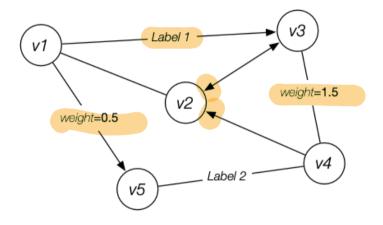
or

$$E = \{ (v1,v2), (v2,v3), \\ (v3,v4), (v1,v4), (v1,v3) \}$$

COMP2521 20T2 \diamondsuit Graph Basics [2/15]

Nodes are distinguished by a unique identifier

Edges may be (optionally) directed, labelled and/or weighted



COMP2521 20T2 \diamondsuit Graph Basics [3/15]

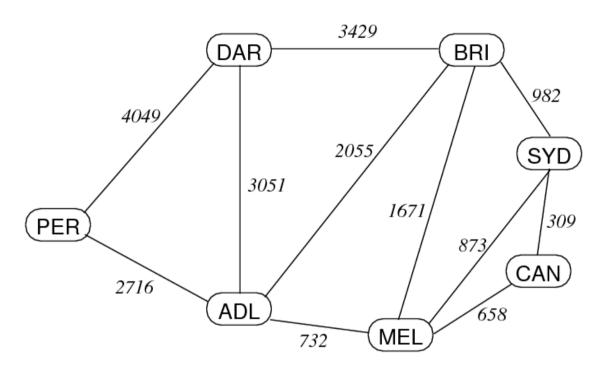
A real example: Australian road distances

Distance	Adelaide	Brisbane	Canberra	Darwin	Melbourne	Perth	Sydney
Adelaide	-	2055	1390	3051	732	2716	1605
Brisbane	2055	-	1291	3429	1671	4771	982
Canberra	1390	1291	-	4441	658	4106	309
Darwin	3051	3429	4441	-	3783	4049	4411
Melbourne	732	1671	658	3783	-	3448	873
Perth	2716	4771	4106	4049	3448	-	3972
Sydney	1605	982	309	4411	873	3972	-

Notes: vertices are cities, edges are distance between cities, symmetric

COMP2521 20T2 \diamondsuit Graph Basics [4/15]

Alternative representation of above:



COMP2521 20T2 \diamondsuit Graph Basics [5/15]

... Graphs

Questions we might ask about a graph:

- is there a way to get from item A to item B? W 维过其他总
- what is the best way to get from A to B?
- which items are directly connected (A B)?

Graph algorithms are generally more complex than tree/list ones:

- no implicit order of items
- graphs may contain cycles
- concrete representation is less obvious
- algorithm complexity depends on connection complexity

Properties of Graphs

Terminology: /V/ and /E/ (cardinality) normally written just as V and E.

A graph with V vertices has at most V(V-1)/2 edges.

50 /2

The ratio *E:V* can vary considerably. 每午点除愈初

- if E is closer to V^2 , the graph is dense $\frac{1}{2}$
- if E is closer to V, the graph is sparse 4 1/2
 - Example: web pages and hyperlinks

Knowing whether a graph is sparse or dense is important

- may affect choice of data structures to represent graph
- may affect choice of algorithms to process graph

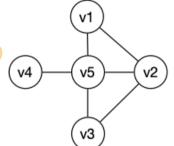
Graph Terminology

For an edge *e* that connects vertices *v* and *w*

- *v* and *w* are adjacent (neighbours)
- e is incident on both v and w

Degree of a vertex *v*

number of edges incident on



degree(v1) = 2degree(v2) = 3

1egree(vz) = 3 Nogroo(v2) = 2

degree(v3) = 2degree(v4) = 1

degree(v5) = 4

Synonyms:

- vertex = node
- edge = arc = link (Note: some people use arc for *directed* edges)

COMP2521 20T2 \diamondsuit Graph Basics [8/15]

... Graph Terminology

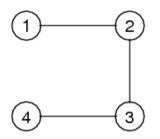
Path: a sequence of vertices where

• each vertex has an edge to its predecessor 複化

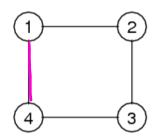
Cycle: a path where

• last vertex in path is same as first vertex in path

Length of path or cycle = #edges



Path: 1−2, 2−3, 3−4



Cycle: 1-2, 2-3, 3-4, 4-1

COMP2521 20T2 \diamondsuit Graph Basics [9/15]



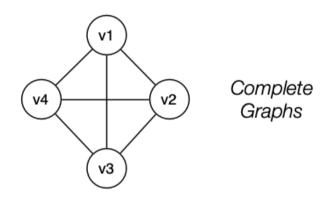
❖ ... Graph Terminology

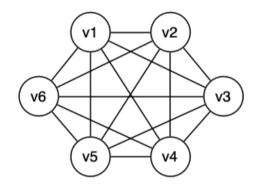
Connected graph

- 一个通所相
- there is a path from each vertex to every other vertex
- if a graph is not connected, it has ≥2 connected components

Complete graph K_V

- there is an edge rom each vertex to every other vertex
- in a complete graph, E = V(V-1)/2





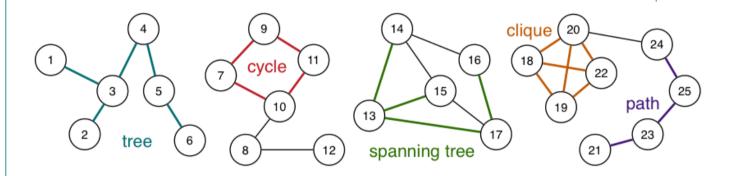
... Graph Terminology

Tree: connected (sub)graph with no cycles

Spanning tree: tree containing all vertices

Clique: complete subgraph

Consider the following single graph:



This graph has 25 vertices, 32 edges, and 4 connected components

Note: The entire graph has no spanning tree; what is shown in green is a spanning tree of the third connected component

... Graph Terminology

A spanning tree of connected graph G = (V,E)

- is a subgraph of G containing all of V
- and is a single tree (connected, no cycles)

A spanning forest of non-connected graph G = (V,E)

- is a subgraph of G containing all of V
- and is a set of trees (not connected, no cycles),
 - with one tree for each connected component

Can form spanning tree from graph by removing edges

e1 e2 e5 Graph (with cycles)

v2 e4 v3

A spanning tree of graph e1 (no cycles)

v2 v3

V1 e3 v3

v4

v1 e3 v4

v4

v1 e3 v4

v4

v4

Many possible spanning trees can be formed. Which is "best"?

❖ ... Graph Terminology

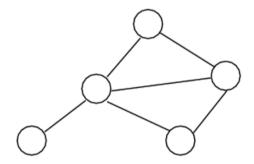
Undirected graph

• edge(u,v) = edge(v,u), no self-loops (i.e. no edge(v,v))

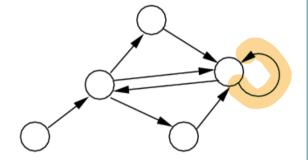
Directed graph

edge(u,v) ≠ edge(v,u), can have self-loops (i.e. edge(v,v))

Examples:



Undirected graph



Directed graph

COMP2521 20T2 \diamondsuit Graph Basics [14/15]

❖ ... Graph Terminology

Other types of graphs ...

Weighted graph

- each edge has an associated value (weight)
- e.g. road map (weights on edges are distances between cities)

Multi-graph

- allow multiple edges between two vertices
- e.g. function call graph (f() calls g() in several places)

Labelled graph

- edges have associated labels
- can be used to add semantic information

COMP2521 20T2 \Diamond Graph Basics [15/15]

Graph Representations

- Graph Representations
- Array-of-edges Representation
- Array-of-edges Cost Analysis
- Adjacency Matrix Representation
- Adjacency Matrix Cost Analysis
- Adjacency List Representation
- Adjacency List Cost Analysis
- Comparison of Graph Representations

COMP2521 20T2 \diamondsuit Graph Representations [0/24]

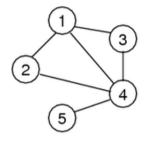
>

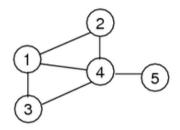
Graph Representations

Describing graphs:

- could describe via a diagram showing edges and vertices
- could describe by giving a list of edges
- assume we identify vertices by distinct integers

E.g. four representations of the same graph:





- 1-2 1-3 1-4
- 1–3 2–1 2–4

3-4

4-1 4-3

4–5

5-4

(a)

(b)

- (c)
- (d)

COMP2521 20T2 \diamondsuit Graph Representations [1/24]

... Graph Representations

We discuss three different graph data structures:

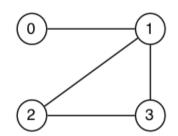
- 1. Array of edges
 - explicit representation of edges as (v,w) pairs
- 2. Adjacency matrix
 - edges defined by presence value in VxV matrix
- 3. Adjacency list
 - edges defined by entries in array of V lists

COMP2521 20T2 \diamondsuit Graph Representations [2/24]

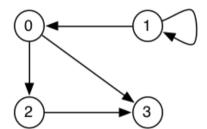
Array-of-edges Representation

Edges are represented as an array of **Edge** values (= pairs of vertices)

- space efficient representation
- adding and deleting edges is slightly complex
- undirected: order of vertices in an **Edge** doesn't matter
- directed: order of vertices in an **Edge** encodes direction



[(0,1), (1,2), (1,3), (2,3)]



[(1,0), (1,1), (0.2), (0,3), (2,3)]

For simplicity, we always assume vertices to be numbered **0..V-1**

... Array-of-edges Representation

Graph initialisation

```
Assumes = struct Graph { int nV; int nE; Edge edges[]; }
```

COMP2521 20T2 \diamondsuit Graph Representations [4/24]

<< ^

... Array-of-edges Representation

Edge insertion

We "normalise" edges so that e.g (v < w) in all (v,w)

COMP2521 20T2 \Diamond Graph Representations [5/24]

... Array-of-edges Representation

Edge removal

COMP2521 20T2 \Diamond Graph Representations [6/24]

... Array-of-edges Representation

Print a list of edges

COMP2521 20T2 \diamondsuit Graph Representations [7/24]

Array-of-edges Cost Analysis

Storage cost: O(E)

Cost of operations:

- initialisation: *O*(1)
- insert edge: *O(E)* (need to check for edge in array)
- delete edge: O(E) (need to find edge in edge array)

If array is full on insert

 allocate space for a bigger array, copy edges across ⇒ still O(E)

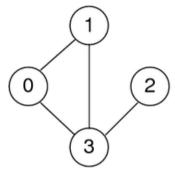
If we maintain edges in order

use binary search to find edge ⇒ O(log E)

COMP2521 20T2 \Diamond Graph Representations [8/24]

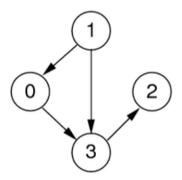
❖ Adjacency Matrix Representation

Edges represented by a $V \times V$ matrix



Undirected graph

A	0	1	2	3
0	0	1	0	1
1	1	0	0	1
2	0	0	0	1
3	1	1	1	0



Directed graph

A -	> 0	1	2	3
0	0	0	0	1
1 -	> 1	0	0	1
2	0	0	0	0
3	0	0	1	0

COMP2521 20T2 \diamondsuit Graph Representations [9/24]

... Adjacency Matrix Representation

Advantages

- easily implemented as 2-dimensional array
- can represent graphs, digraphs and weighted graphs
 - o graphs: symmetric boolean matrix
 - digraphs: non-symmetric boolean matrix
 - weighted: non-symmetric matrix of weight values

Disadvantages:

• if few edges (sparse) \Rightarrow memory-inefficient ($O(V^2)$ space)

内存与用声多

COMP2521 20T2 \Diamond Graph Representations [10/24]

... Adjacency Matrix Representation

Graph initialisation

COMP2521 20T2 \Diamond Graph Representations [11/24]

... Adjacency Matrix Representation

Edge insertion

```
insertEdge(g,(v,w)):
    Input graph g, edge (v,w)
    Output graph g containing (v,w)

if g.edges[v][w] = 0 then // (v,w) not in graph
    g.edges[v][w]=1 // set to true
    g.edges[w][v]=1
    g.nE=g.nE+1
    end if
```

COMP2521 20T2 \Diamond Graph Representations [12/24]

... Adjacency Matrix Representation

Edge removal

```
removeEdge(g,(v,w)):
    Input graph g, edge (v,w)
    Output graph g without (v,w)

if g.edges[v][w] ≠ 0 then // (v,w) in graph
    g.edges[v][w]=0 // set to false
    g.edges[w][v]=0
    g.nE=g.nE-1
    end if
```

COMP2521 20T2 \Diamond Graph Representations [13/24]

... Adjacency Matrix Representation

Print a list of edges

```
showEdges(g):
    Input graph g

for all i=0 to g.nV-1 do
    for all j=i+1 to g.nV-1 do
        if g.edges[i][j] ≠ 0 then
            print i"-"j
        end if
    end for
end for
```

COMP2521 20T2 \Diamond Graph Representations [14/24]

Storage cost: $O(V^2)$

3×3

If the graph is sparse, most storage is wasted.

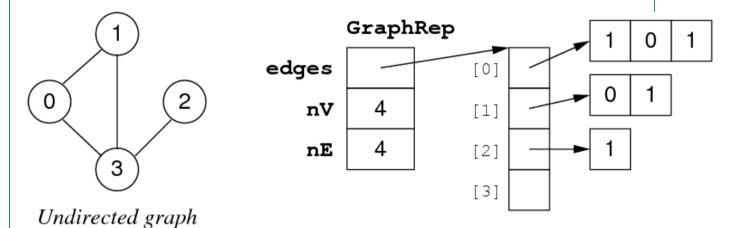
Cost of operations:

- initialisation: $O(V^2)$ (initialise $V \times V$ matrix)
- insert edge: O(1) (set two cells in matrix)
- delete edge: O(1) (unset two cells in matrix)

COMP2521 20T2 \diamondsuit Graph Representations [15/24]

... Adjacency Matrix Cost Analysis

A storage optimisation: store only top-right part of matrix.



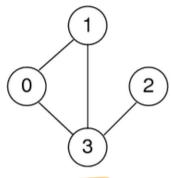
New storage cost: V-1 int ptrs + V(V+1)/2 ints (but still $O(V^2)$)

Requires us to always use edges (v,w) such that v < w.

COMP2521 20T2 \Diamond Graph Representations [16/24]

❖ Adjacency List Representation

For each vertex, store linked list of adjacent vertices:



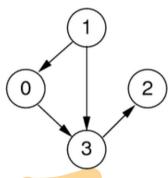
Undirected graph



$$A[1] = <0, 3>$$

$$A[2] = <3>$$

$$A[3] = <0, 1, 2>$$



Directed graph

$$A[1] = <0, 3>$$

$$A[2] = <>$$

$$A[3] = <2>$$

... Adjacency List Representation

Advantages

- relatively easy to implement in languages like C
- can represent graphs and digraphs
- memory efficient if E:V relatively small

Disadvantages:

one graph has many possible representations
 (unless lists are ordered by same criterion e.g. ascending)

COMP2521 20T2 \Diamond Graph Representations [18/24]

... Adjacency List Representation

Graph initialisation

... Adjacency List Representation

Edge insertion:

```
insertEdge(g,(v,w)):
    Input graph g, edge (v,w)
    Output graph g containing (v,w)

if not ListMember(g.edges[v],w) then
    // (v,w) not in graph
    ListInsert(g.edges[v],w)
    ListInsert(g.edges[w],v)
    g.nE=g.nE+1
    end if
```

<< ^ >>

... Adjacency List Representation

Edge removal:

```
removeEdge(g,(v,w)):
    Input graph g, edge (v,w)
    Output graph g without (v,w)

if ListMember(g.edges[v],w) then
    // (v,w) in graph
    ListDelete(g.edges[v],w)
    ListDelete(g.edges[w],v)
    g.nE=g.nE-1
    end if
```

... Adjacency List Representation

Print a list of edges

```
showEdges(g):
    Input graph g

for all i=0 to g.nV-1 do
    for all v in g.edges[i] do
        if i < v then
            print i"-"v
        end if
    end for
end for</pre>
```

COMP2521 20T2 \Diamond Graph Representations [22/24]

Adjacency List Cost Analysis

Storage cost: O(V+E)

Cost of operations:

- initialisation: O(V) (initialise V lists)
- insert edge: *O(E)* (need to check if vertex in list)
- delete edge: *O(E)* (need to find vertex in list)

Could sort vertex lists, but no benefit (although no extra cost)

COMP2521 20T2 \diamondsuit Graph Representations [23/24]

Comparison of Graph Representations

Summary of operations above:

House	array of edges		adjacency list
space usage	E	V^2	V+E
initialise	1	V^2	V
insert edge	E	1	Ε
remove edge	E	1	E

Other operations:

	array of edges	adjacency matrix	adjacency list
disconnected(v)?	E	V	1
isPath(x,y)?	E·log V	V^2	V+E
copy graph	Ε	V^2	V+E
destroy graph	1	V	V+E

Graph ADT

- Graph ADT
- Graph ADT (Array of Edges)
- Graph ADT (Adjacency Matrix)
- Graph ADT (Adjacency Lists)
- Example: Graph ADT Client

COMP2521 20T2 \Diamond Graph ADT [0/24]

Graph ADT

Data: set of edges, set of vertices

Operations:

- building: create graph, add edge
- deleting: remove edge, drop whole graph
- scanning: check if graph contains a given edge

Things to note:

- set of vertices is fixed when graph initialised
- we treat vertices as ints, but could be arbitrary Items

Will use this ADT as a basis for building more complex operations later.

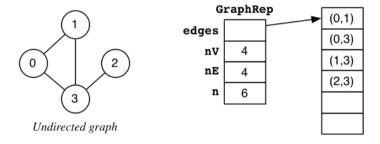
... Graph ADT

Graph ADT interface Graph.h

COMP2521 20T2 \Diamond Graph ADT [2/24]

Implementation of **GraphRep** (array-of-edges representation)

```
typedef struct GraphRep {
   Edge *edges; // array of edges
   int nV; // #vertices (numbered 0..nV-1)
   int nE; // #edges
   int n; // size of edge array
} GraphRep;
```



Implementation of graph initialisation (array-of-edges)

```
Graph newGraph(int V) {
   assert(V >= 0);
   Graph g = malloc(sizeof(GraphRep));
   assert(g != NULL);
   g->nV = V; g->nE = 0;
   // allocate enough memory for edges
   g->n = Enough;
   g->edges = malloc(g->n*sizeof(Edge));
   assert(g->edges != NULL);
   return g;
}
```

How much is enough? ... No more than V(V-1)/2 ... Much less in practice (sparse graph)

からとメンショクスショウ

COMP2521 20T2 \Diamond Graph ADT [4/24]

Some useful utility functions:

Implementation of edge insertion (array-of-edges)

Implementation of edge removal (array-of-edges)

```
void removeEdge(Graph g, Edge e) {
    // ensure that g exists
    assert(g != NULL && validE(g,e));
    int i = 0;
    while (i < g->nE && !eq(e,g->edges[i]))
        i++;
    if (i < g->nE) // edge e found
        g->edges[i] = g->edges[--g->nE];
}
```

Implementation of edge check (array-of-edges)

```
bool adjacent(Graph g, Vertex x, Vertex y) {
   assert(g != NULL && validV(g,x) && validV(g,y));
   Edge e;
   e.v = x; e.w = y;
   for (int i = 0; i < g->nE; i++) {
      if (eq(e,g->edges[i])) // edge found
        return true;
   }
   return false; // edge not found
}
```

Re-implementation of edge insertion (array-of-edges)

Implementation of graph removal (array-of-edges)

```
void freeGraph(Graph g) {
   assert(g != NULL);
   free(g->edges); // free array of edges
   free(g); // remove Graph object
}
```

COMP2521 20T2 \Diamond Graph ADT [10/24]

Graph ADT (Adjacency Matrix)

Implementation of **GraphRep** (adjacency-matrix representation)

```
typedef struct GraphRep {
   int
         **edges; // adjacency matrix
   int
            nV; // #vertices
           nE;
   int
                  // #edges
} GraphRep;
                              GraphRep
                                                      1
                                                         0
                        edges
                                        [0]
                                                      0
                                                         0
       0
                                 4
                           nV
                                        [1]
                                 4
                                                      0
                                                         0
                                                            1
                           nΕ
                                        [2]
                                        [3]
                                                         1
                                                            0
      Undirected graph
```

COMP2521 20T2 \Diamond Graph ADT [11/24]

<< ^ >>

... Graph ADT (Adjacency Matrix)

Implementation of graph initialisation (adjacency-matrix)

```
Graph newGraph(int V) {
   assert(V >= 0);
   Graph g = malloc(sizeof(GraphRep));
   assert(g != NULL);
   g->nV = V;   g->nE = 0;
   // allocate array of pointers to rows
   g->edges = malloc(V * sizeof(int *));
   assert(g->edges != NULL);
   // allocate memory for each column and initialise with 0
   for (int i = 0; i < V; i++) {
      g->edges[i] = calloc(V, sizeof(int));
      assert(g->edges[i] != NULL);
   }
   return g;
}
```

Standard library function calloc(size_t nelems, size_t nbytes)

- allocates a memory block of size **nelems*nbytes**
- and sets all bytes in that block to zero

COMP2521 20T2 \diamondsuit Graph ADT [12/24]

... Graph ADT (Adjacency Matrix)

Implementation of edge insertion (adjacency-matrix)

```
void insertEdge(Graph g, Edge e) {
  assert(g != NULL && validE(g,e));

if (!g->edges[e.v][e.w]) { // edge e not in graph
    g->edges[e.v][e.w] = 1;
    g->edges[e.w][e.v] = 1;
    g->nE++;
  }
}
```

COMP2521 20T2 \Diamond Graph ADT [13/24]

... Graph ADT (Adjacency Matrix)

Implementation of edge removal (adjacency-matrix)

```
void removeEdge(Graph g, Edge e) {
  assert(g != NULL && validE(g,e));

if (g->edges[e.v][e.w]) { // edge e in graph
    g->edges[e.v][e.w] = 0;
    g->edges[e.w][e.v] = 0;
    g->nE--;
}
```

COMP2521 20T2 \Diamond Graph ADT [14/24]

<< ^ >>

... Graph ADT (Adjacency Matrix)

Implementation of edge check (adjacency matrix)

```
bool adjacent(Graph g, Vertex x, Vertex y) {
   assert(g != NULL && validV(g,x) && validV(g,y));

return (g->edges[x][y] != 0);
}
```

Note: all operations, except creation, are *O(1)*

COMP2521 20T2 \Diamond Graph ADT [15/24]

... Graph ADT (Adjacency Matrix)

Implementation of graph removal (adjacency matrix)

COMP2521 20T2 \Diamond Graph ADT [16/24]

Implementation of **GraphRep** (adjacency-lists representation)

```
typedef struct GraphRep {
   Node **edges; // array of lists
   int
            nV;
                      // #vertices
   int
            nE;
                     // #edges
} GraphRep;
                                               ist
                           GraphRep
                     edges
                                    [0]
                2
      0
                             4
                        nV
                                    [1]
                        nΕ
                             4
                                    [2]
                                    [3]
      Undirected graph
```

COMP2521 20T2 \Diamond Graph ADT [17/24]

Assume that we have a linked list implementation

```
typedef struct Node {
    Vertex v;
    struct Node *next;
} Node;
```

with operations like inLL, insertLL, deleteLL, freeLL, e.g.

```
bool inLL(Node *L, Vertex v) {
   while (L != NULL) {
     if (L->v == v) return true;
     L = L->next;
   }
   return false;
}
```

COMP2521 20T2 \Diamond Graph ADT [18/24]

<< ^ >>

... Graph ADT (Adjacency Lists)

Implementation of graph initialisation (adjacency lists)

COMP2521 20T2 \Diamond Graph ADT [19/24]

Implementation of edge insertion/removal (adjacency lists)

```
void insertEdge(Graph g, Edge e) {
   assert(g != NULL && validE(g,e));
   if (!inLL(g->edges[e.v], e.w)) { // edge e not in graph
      g->edges[e.v] = insertLL(g->edges[e.v], e.w);
      g->nE++)
   }
}
void removeEdge(Graph g, Edge e) {
   assert(g != NULL && validE(g,e));
   if (inLL(g->edges[e.v], e.w)) { // edge e in graph
      g->edges[e.v] = deleteLL(g->edges[e.v], e.w);
      g->edges[e.w] = deleteLL(g->edges[e.w], e.v);
      g->nE--;
   }
}
```

COMP2521 20T2 \Diamond Graph ADT [20/24]

Implementation of edge check (adjacency lists)

```
bool adjacent(Graph g, Vertex x, Vertex y) {
   assert(g != NULL && validV(g,x) && validV(g,y));

return inLL(g->edges[x], y);
}
```

Note: all operations, except creation, are O(E)

COMP2521 20T2 \Diamond Graph ADT [21/24]

Implementation of graph removal (adjacency lists)

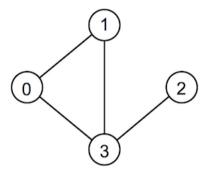
```
void freeGraph(Graph g) {
   assert(g != NULL);
   for (int i = 0; i < g->nV; i++)
        freeLL(g->edges[i]); // free one list
   free(g->edges); // free array of list pointers
   free(g); // remove Graph object
}
```

COMP2521 20T2 \Diamond Graph ADT [22/24]

❖ Example: Graph ADT Client

A program that uses the graph ADT to

- build the graph depicted below
- print all the nodes that are incident to vertex 1 in ascending order



COMP2521 20T2 \Diamond Graph ADT [23/24]

... Example: Graph ADT Client

```
#include <stdio.h>
#include "Graph.h"

#define NODES 4
#define NODE_OF_INTEREST 1

int main(void) {
    Graph g = newGraph(NODES);
    Edge e;

    while (scanf("%d %d", &(e.v), &(e.w)) == 2)
        insertEdge(g,e);

    for (Vertex v = 0; v < NODES; v++) {
        if (adjacent(g, v, NODE_OF_INTEREST))
            printf("%d\n", v);
    }

    freeGraph(g);
    return 0;
}</pre>
```

COMP2521 20T2 \diamondsuit Graph ADT [24/24]