Graph Traversals

Lecture 27
By Marina Barsky
Graph Traversal

A graph traversal algorithm explores every vertex (and edge) of a graph.

It is one of the most important algorithms used as an archetype for solving many interesting problems:

- Courses and prerequisites (topological sort)
- Sub-networks (connected components)
- “Weak link” (articulation points)
- “Degree of separation” (shortest paths)
- ...

Example: Modeling Mazes with Graphs

A maze can be viewed as a grid graph.

The vertices are pairs of \((x,y)\) coordinates, and two neighboring cells are connected by an edge if there is no wall between them.
Perfect mazes

A paper-and-pencil maze that has no loops and no inaccessible areas is called a "perfect maze".

Perfect maze

Imperfect maze
Solving perfect mazes

Perfect mazes can be solved using the right-hand rule.
(The left-hand rule also works and explores the maze in the opposite order.)

"Always keep your right hand on the wall." — Right-Hand Rule

If you don't take the exit after you found it, then you will traverse the entire maze graph before returning to the entrance.
Solving perfect mazes

Perfect mazes are easier to traverse than arbitrary graphs since the underlying graph is *acyclic*.

The right-hand rule will not work for imperfect graphs or for arbitrary graphs.

"Always keep your right hand on the wall." — Right-Hand Rule
When solving paper mazes humans "see" the entire maze all at once and can use their intuition.

Computer programs work one step at a time and only "see" one thing.

In graph traversal algorithms we need to use a first-person perspective.
Types of Graph Traversal

Recall tree traversal algorithms:

- Depth First:
  - In-order
  - Pre-order
  - Post-order
- Breadth First

Traversing general graphs is similar:

- Breadth-First Search
- Depth-First Search

When traversing general graphs (as opposed to trees) we need to be more careful.

We may return to the same node and loop around it endlessly.
We may end up looping around this cycle.
Endless Looping

Super 3D Noah's Ark

We try to avoid doing this!
Labeling Vertices

To avoid the looping problem we will keep track of some data. Each vertex will be marked as being in one of three different states.

1. A vertex is **undiscovered** if the algorithm has not seen the vertex yet.
2. A vertex is **discovered** if the algorithm has seen the vertex but we have not fully explored all of its incident edges.
3. A vertex is **processed** if the algorithm has seen the vertex and followed all of its incident edges.
We would also need: Queues and Stacks

Queue

Stack
To-Do List

In a special To-Do list we will keep track of vertices that have been discovered but not yet processed.

Whenever we encounter an undiscovered vertex we add it to a *to-do list*.

**Illustration:**

- Initially the to-do list contains the single vertex where we start.
- For consistency of the discussion, when processing we examine a vertex's neighbors clockwise from 12 o'clock.
Generic Algorithm for Graph Traversal

The following algorithm will correctly and efficiently explore all the vertices.

Algorithm `traverse` (graph G, vertex start)

1. for each u in vertices of G:
   - u.state := “undiscovered”
2. start.state := “discovered”
3. todo := new to_do_list()
4. todo.add(start)
5. while todo is not empty:
   - current_vertex := todo.remove()
   - for each u in neighbors(current_vertex):
     - if u.state = “undiscovered”:
       - u.state := “discovered”
       - todo.add(u)
7. current_vertex.state := “processed”
Breadth-First Search

In breadth-first search (BFS) the to-do list is a (FIFO) queue.

Algorithm **BFS** (graph G, vertex start)

1. for each u in vertices of G:
   
   u.state := “undiscovered”

2. start.state := “discovered”

3. todo := new Queue()

4. todo.enqueue(start)

5. while todo is not empty:

6.   current_vertex := todo.dequeue()

7.   for each u in neighbors(current_vertex):

8.     if u.state = “undiscovered”:

9.       u.state := “discovered”

10.      todo.enqueue(u)

11.  current_vertex.state := “processed”
Example: BFS

Breadth-first search starting from vertex g.
Example: BFS

Breadth-first search starting from vertex g.

current_vertex: g

<table>
<thead>
<tr>
<th>color</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td>discovered</td>
<td></td>
</tr>
<tr>
<td>discovered</td>
<td></td>
</tr>
<tr>
<td>processed</td>
<td></td>
</tr>
</tbody>
</table>

Graph:

- Vertices: a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q
- Edges:
  - a to b
  - b to e
  - e to g
  - g to b
  - b to c
  - c to e
  - e to g
  - g to f
  - f to c
  - c to d
  - d to h
  - h to i
  - i to n
  - n to o
  - o to p
  - p to q
  - q to g
  - g to j
  - j to l
  - l to m
  - m to o
  - o to n

State:
- g: discovered
- all other vertices: undiscovered
Example: BFS

Breadth-first search starting from vertex g.
Example: BFS

Breadth-first search starting from vertex g.

<table>
<thead>
<tr>
<th>color</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>undiscovered</td>
</tr>
<tr>
<td></td>
<td>discovered</td>
</tr>
<tr>
<td></td>
<td>processed</td>
</tr>
</tbody>
</table>

current_vertex  todo  

<table>
<thead>
<tr>
<th></th>
<th>e</th>
<th>h</th>
<th>l</th>
<th>j</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example: BFS

Breadth-first search starting from vertex g.

color | state
--- | ---
undiscovered
discovered
processed

current_vertex  e  todo h l j
Example: BFS

Breadth-first search starting from vertex g.

color | state
undiscovered
discovered
processed

current_vertex | e |
--- | ---
todo | h | l | j | a
Example: BFS

Breadth-first search starting from vertex g.

<table>
<thead>
<tr>
<th>color</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>undiscovered</td>
</tr>
<tr>
<td></td>
<td>discovered</td>
</tr>
<tr>
<td></td>
<td>processed</td>
</tr>
</tbody>
</table>

current_vertex: h

todo: h l j a
Example: BFS

Breadth-first search starting from vertex g.

```
color    state
undiscovered
discovered
processed
```

current_vertex | h | todo | l | j | a |
Example: BFS

Breadth-first search starting from vertex g.
Example: BFS

Breadth-first search starting from vertex g.

<table>
<thead>
<tr>
<th>color</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td>discovered</td>
<td></td>
</tr>
<tr>
<td>discovered</td>
<td></td>
</tr>
<tr>
<td>discovered</td>
<td></td>
</tr>
<tr>
<td>processed</td>
<td></td>
</tr>
</tbody>
</table>
Example: BFS

Breadth-first search starting from vertex g.
Example: BFS

Breadth-first search starting from vertex g.

current_vertex

<table>
<thead>
<tr>
<th>color</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>discovered</td>
</tr>
<tr>
<td></td>
<td>processed</td>
</tr>
</tbody>
</table>
Example: BFS

Breadth-first search starting from vertex g.

<table>
<thead>
<tr>
<th>color</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>discovered</td>
</tr>
<tr>
<td></td>
<td>processed</td>
</tr>
</tbody>
</table>

current_vertex | todo | j | a | c | i | p | f | m
Example: BFS

Breadth-first search starting from vertex g.

color | state
--- | ---
undiscovered
discovered
processed

current_vertex j
todo a c i p f m
Example: BFS

Breadth-first search starting from vertex g.

current_vertex: j

todo: a, c, i, p, f, m, o, ...

color state
undiscovered
discovered
processed
Example: BFS

Breadth-first search starting from vertex g.
Example: BFS

Breadth-first search starting from vertex g.

We have processed every vertex that is one edge away from g.

Discovered vertices in the queue are distance two from g.
BFS

1. Breadth-first search will always visit closer nodes first. In other words, a node at distance \( d \) edges from the start will be visited before any node at distance \( d+1 \).

2. The nodes are discovered in the same order that they are processed. Both of these orders can be obtained from the todo list by crossing out the values instead of erasing them.

**Note:** On the exam you may be asked to specify the order that the nodes are discovered / processed. For the previous example the answer would be:

```
g e h l j a c i p f m o n k b d q
```
Depth-First Search

In depth-first search (DFS) we process the to-do list as a stack.

Algorithm DFS (graph G, start)

```plaintext
for each u in vertices of G:
    u.state := "undiscovered"

start.state := "discovered"
todo := new stack()
todo.push(start)

while todo is not empty:
    current_vertex := todo.pop()
    for each u in neighbors(current_vertex):
        if u.state = "undiscovered":
            u.state := "discovered"
            todo.push(u)

    current_vertex.state := "processed"
```
Example: DFS

Depth-first search starting from vertex g.

<table>
<thead>
<tr>
<th>color</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td>discovered</td>
<td></td>
</tr>
<tr>
<td>processed</td>
<td></td>
</tr>
</tbody>
</table>

Current vertex: g

Todo: g
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.

color | state
---|---
undiscovered
discovered
processed

current_vertex: g
todo: e h l j
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.

color
undiscovered
discovered
processed

current_vertex  todo  e  h  l  j

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.

color | state
--- | ---
undiscovered | discovered | processed

current_vertex | todo | e | h | l | a | o | n | k
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.

current_vertex: n

todo: e h l a o

color state
undiscovered
discovered
processed
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.

color | state
------|------
       | undiscovered
       | discovered
       | processed

current_vertex | o  todo | e | h | l | a

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.

Current vertex: o

<table>
<thead>
<tr>
<th>color</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>undiscovered</td>
</tr>
<tr>
<td></td>
<td>discovered</td>
</tr>
<tr>
<td></td>
<td>processed</td>
</tr>
</tbody>
</table>
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.

color  state
undiscovered
discovered
processed

current_vertex  p  todo  e  h  l  a  m  i  q
Example: DFS

Depth-first search starting from vertex g.
Example: DFS

Depth-first search starting from vertex g.

Some vertices of distance **one** have only been discovered.

Some vertices of distance **two** from g have been processed.

<table>
<thead>
<tr>
<th>color</th>
<th>state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>discovered</td>
</tr>
<tr>
<td></td>
<td>processed</td>
</tr>
</tbody>
</table>
**Note:** on the exam you could be asked for the order that the nodes are discovered / processed. In the previous example the answer would be the following.

```
g e h l j a o n k m p i q d c b f
```

discovered

```
g j k n o p q i c b d m a l h f e
```

processed
Breadcrumb (DFS) algorithm for getting out of the woods

Image of BFS is from wikipedia
Solving mazes with graph traversals

- DFS (as well as BFS) can be used to solve mazes.
- Both algorithms traverse the entire graph - so eventually they will find a path from start to exit, if such path exists.

Warm memories of Lab 5...
We can modify the **DFS** to mark the vertex as processed *only after all vertices reachable from it have been discovered/processed*.

### Algorithm DFS (graph G, start)

<table>
<thead>
<tr>
<th>Line</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>for each</strong> u in vertices of G:</td>
</tr>
<tr>
<td>2</td>
<td>u.state := “undiscovered”</td>
</tr>
<tr>
<td>3</td>
<td>start.state := “discovered”</td>
</tr>
<tr>
<td>4</td>
<td>todo := new <strong>stack</strong>()</td>
</tr>
<tr>
<td>5</td>
<td>todo.push(start)</td>
</tr>
<tr>
<td>6</td>
<td><strong>while</strong> todo is not empty:</td>
</tr>
<tr>
<td>7</td>
<td>current_vertex := todo.top()  # don’t pop yet: just read</td>
</tr>
<tr>
<td>8</td>
<td><strong>for each</strong> u in neighbors(current_vertex):</td>
</tr>
<tr>
<td>9</td>
<td>count := 0</td>
</tr>
<tr>
<td>10</td>
<td>if u.state = “undiscovered”:</td>
</tr>
<tr>
<td>11</td>
<td>u.state := “discovered”</td>
</tr>
<tr>
<td>12</td>
<td>todo.push(u)</td>
</tr>
<tr>
<td>13</td>
<td>count := count + 1</td>
</tr>
<tr>
<td>14</td>
<td>if count = 0:  # nowhere to go</td>
</tr>
<tr>
<td>15</td>
<td>current_vertex.state = “processed”</td>
</tr>
<tr>
<td>16</td>
<td>todo.pop()</td>
</tr>
</tbody>
</table>
Depth-First Search 2 (Recursive)

This later variation of the Depth-first search can also be implemented recursively. This implementation implicitly replaces the _todo stack_ with the _call stack_.

---

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>DFS(G, current)</th>
</tr>
</thead>
<tbody>
<tr>
<td>current.state:= “discovered”</td>
<td></td>
</tr>
<tr>
<td>for each u in neighbors(current)</td>
<td></td>
</tr>
<tr>
<td>if u.state = “undiscovered” then</td>
<td></td>
</tr>
<tr>
<td>DFS(G, u)</td>
<td></td>
</tr>
<tr>
<td>current.state:= “processed”</td>
<td></td>
</tr>
</tbody>
</table>

| |
| for each u in vertices of G |
| u.state:= “undiscovered” |
| DFS(G, start)  // start is a vertex in G |

Recursive pseudocode for DFS
**BFS and DFS: side-by-side**

**Algorithm BFS**\((G, \text{start})\)

1. For each \(u\) in vertices of \(G\)
   - \(u\).\text{state} := “un”

2. \(\text{start}\).\text{state} := “d”

3. \(\text{todo} := \text{new queue}()\)
4. \(\text{todo}.\text{enqueue}(\text{start})\)

5. While \(\text{todo is not empty}\):
   - \(v := \text{todo}.\text{dequeue}()\)
   - For each \(u\) in neighbors\((v)\)
     - If \(u\).\text{state} = “un”:
       - \(u\).\text{state} := “d”
       - \(\text{todo}.\text{enqueue}(u)\)
   - \(v\).\text{state} := “p”

**Algorithm DFS**\((G, \text{start})\)

1. For each \(u\) in vertices of \(G\)
   - \(u\).\text{state} := “un”

2. \(\text{start}\).\text{state} := “d”

3. \(\text{todo} := \text{new stack}()\)
4. \(\text{todo}.\text{push}(\text{start})\)

5. While \(\text{todo is not empty}\):
   - \(v := \text{todo}.\text{pop}()\)
   - For each \(u\) in neighbors\((v)\)
     - If \(u\).\text{state} = “un”:
       - \(u\).\text{state} := “d”
       - \(\text{todo}.\text{push}(u)\)
   - \(v\).\text{state} := “p”

Note that the stack and queue operations must guarantee \(O(1)\)-time.
Runtime of BFS and DFS

• Eventually, both algorithms traverse all the edges of G

• The running time is therefore $O(m)$, where $m$ is the number of edges

• There are at most $O(n^2)$ edges, so this is $O(n^2)$-time

$O(n^2)$ for **dense graphs**

$O(n)$ for **sparse graphs**
Graph traversals in real life: hiking trails
Graph traversals in real life: food
Exploration strategies

Some people fully explore areas around them before moving on, and others continue moving further and further from their home, and return to explore the surrounding areas later.

These strategies roughly correspond to *breadth-first* and *depth-first* respectively.
Group work

You are given the following Graph:

Show the order in which the vertices of this graph are discovered and processed for 3 types of traversals.
BREADTH-FIRST TRAVERSAL starting with node C

• Specify the order in which nodes are DISCOVERED:

• Specify the order in which nodes are PROCESSED:
BREADTH-FIRST TRAVERSAL starting with node C

Solution

• Specify the order in which nodes are DISCOVERED: cdbhafge

• Specify the order in which nodes are PROCESSED: cdbhafge
DEPTH-FIRST TRAVERSAL (NON-RECURSIVE)
starting with node C

- Specify the order in which nodes are DISCOVERED:

- Specify the order in which nodes are PROCESSED:
DEPTH-FIRST TRAVERSAL (NON-RECURSIVE) starting with node C

Solution

- Specify the order in which nodes are DISCOVERED: cdbhafge

- Specify the order in which nodes are PROCESSED: cafgehbd
DEPTH-FIRST TRAVERSAL (RECURSIVE!) starting with node C

- Specify the order in which nodes are DISCOVERED:

- Specify the order in which nodes are PROCESSED:
DEPTH-FIRST TRAVERSAL (RECURSIVE!) starting with node C

Solution

• Specify the order in which nodes are DISCOVERED: cdfgebha

• Specify the order in which nodes are PROCESSED: ahbegfdc