Shortest Paths

COMP4128 Programming Challenges

School of Computer Science and Engineering UNSW Australia

Single Source Shortest PathsDijkstra's AlgorithmBellman-Ford Algorithm

Single Source Shortest Paths

- 2 All Pairs Shortest Paths
- 3 Implicit Graphs

- Given a weighted directed graph G with two specific vertices s and t, we want to find the shortest path that goes between s and t on the graph.
- Note that the unweighted case is solved by BFS.
- Generally, algorithms which solve the shortest path problem also solve the single source shortest path problem, which computes shortest paths from a single source vertex to every other vertex.
- You can represent all the shortest paths from the same source as a tree.

 It's very important to distinguish between graphs where all edges are positive and graphs with negative weight edges! Why?

Single Source Shortest Paths

- Imagine a graph with a cycle whose total weight is negative.
- Even if there are no negative cycles, this may still cause problems depending on your algorithm choice!
- If the graph is acyclic, negative weight edges generally don't cause problems, but care should be taken regardless.

- point, we keep what we think is the shortest path to each Most single source shortest paths algorithms rely on the basic idea of building shortest paths iteratively. At any vertex and we update this by "relaxing" edges.
 - Relax(u, v): if the currently found shortest path from our source s to a vertex v could be improved by using the
- distance. But with negative weights, this becomes trickier. with only relaxing vertices for which we know the optimal For graphs with non-negative weights, we can get away edge (u, v), update it.

- If we keep track for each ν of its most recently relaxed incoming edge, we can find the actual path from the source to ν . How?
- For each ν , we know the vertex we would've come from if we followed the shortest path from the source.
- We can work backwards from v to the source to find the shortest path from the source to v.

• If we keep relaxing our edges until we can't anymore, then we will have a shortest path.

Single Source Shortest Paths

How do we choose which edges to relax?

- For now, suppose there are no negative edges.
- Visit each vertex *u* in turn, starting from the source *s*. Whenever we visit the vertex *u*, we relax all of the edges coming out of *u*.

Dijkstra's Algorithm

- How do we decide the order in which to visit each vertex?
- We can do something similar to breadth-first search.
- The next vertex we process is always the unprocessed vertex with the smallest distance from the source.
- This ensures that we only need to process each vertex once: by the time we process a vertex, we have definitely found the shortest path to it. Prove this inductively!

- To decide which vertex we want to visit next, we can either just loop over all of them, or use a priority queue keyed on each vertex's current shortest known distance from the source.
- Since we know that we have a complete shortest path to every vertex by the time we visit it in Dijkstra's algorithm, we know we only visit every vertex once.
- Complexity Dijkstra's Algorithm is $O(E \log V)$ using a binary heap as a priority queue, or $O(V^2)$ with a loop.

- The above only holds for graphs without negative edges!
- With negative edges, we may need to visit each vertex more than once, and it turns out this makes the runtime exponential in the worst case (and it's even worse with negative cycles).
- In short: don't use Dijkstra's if there's any negative edges! (But most graphs you see won't have them).

- How do we handle negative edges in a more efficient way?
- How do we handle negative cycles?
- **Key Observation:** If a graph has no negative cycle then all shortest paths from u have length $\leq |V| 1$. Conversely, a negative cycle implies there is a shortest path of length |V| better than any path of length |V| 1.
- So we should instead build up shortest paths by number of edges, not just from u outwards.
- Bellman-Ford involves trying to relax every edge of the graph (a *global relaxation*) |V|-1 times and update our tentative shortest paths each time.
- Because every shortest path has at most |V|-1 edges, if after all of these global relaxations, relaxations can still be made, then there exists a negative cycle.

- The time complexity of Bellman-Ford is O(VE).
- However, if we have some way of knowing that the last global relaxation did not affect the tentative shortest path to some vertex ν , we know that we don't need to consider edges coming out of ν in our next global relaxation.
- This heuristic doesn't change the overall time complexity of the algorithm, but makes it run very fast in practice on random graphs.
- Sometimes called Shortest Path Faster Algorithm (SPFA) for some reason...

Bellman-Ford Algorithm

Implementation (continued)

```
// Puts distances from source (n-1) in dist
// Returns true if there is a negative cycle, false otherwise.
// NOTE: You can't trust the dist array if this function returns True.
rector(int) find_dists_and_check_neg_cycle() {
    fill(dist.begin(), dist.end(), INF);
    dist(n-1) = 0;
    // VI-1 global relaxations
    for (int i = 0; i < n - 1; i++) relax();
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ), // Otherwise, no negative cycle, that condition is actually a iff return false;
                                                                                                                                                                                                                                                                                                                       // If any edge can be relaxed further, there is a negative cycle for (auto e = edges begin(); e != edges.end(); ++e) {
   if (dist[e->u] != IMF &&
        return true;
```

- Slight technical note: Due to how we coded relax, after n-1 iterations, dist[i] may be the distance of a path of length > n-1. But this does not matter, everything still holds.
- If there is a negative cycle, you can't trust the distances computed. Call a vertex v ruined if its shortest distance from u is actually $-\infty$.
- For every negative cycle, in every relaxation round at least one of its vertices will be updated.
- Hence, to find all ruined vertices, DFS out of all vertices who were relaxed in the *V*-th round.
- To find a specific negative cycle, backtrack from any 'ruined' vertex.

3 Implicit Graphs

• The all pairs shortest path problem involves finding the shortest path between every pair of vertices in the graph.

All Pairs Shortest Paths

• Surprisingly, this can be found in $O(V^3)$ time and $O(V^2)$ memory, by dynamic programming.

- Let f(u, v, i) be the length of the shortest path between u and v using only the first i vertices (i.e. the vertices with the i smallest labels) as intermediate vertices.
- The key is to build this up for increasing values of *i*.

• Base Case: Then f(u, u, 0) = 0 for all vertices u, and $f(u, v, 0) = w_e$ if there is an edge e from u to v, and infinity otherwise.

$$f(u, v, i) = \min(f(u, v, i-1), f(u, i, i-1) + f(i, v, i-1)).$$

- The solution we already had, f(u, v, i-1), definitely doesn't use i as an intermediate vertex.
- If i is the only new intermediate vertex we can use, the only new path that could be better is the shortest path u → i concatenated with the shortest path i → v.

- $f(u, v, i) = \min(f(u, v, i-1), f(u, i, i-1) + f(i, v, i-1))$
- Thus, f(u, v, n) will give the length of the shortest path from u to v.
- Noting that to calculate the table for the next i, we only need the previous table, we see that we can simply overwrite the previous table at each iteration, so we only need $O(V^2)$ space.

- But what if f(u, i, i 1) or f(i, v, i 1) is overwritten in the table before we get to use it?
 - the table before we get to use it?Allowing the use of i as an intermediate vertex on a path to or from i is not going to improve the path, unless we have a negative-weight cycle.

What if there is a negative cycle?

- If there is a negative-weight cycle, our invariant is instead $f(u,v,i) \leq$ shortest simple path from $u \rightarrow v$ only using the first i vertices as intermediaries.
- Hence f(u, u, n) will be negative for vertices in negative cycles. Also you can't trust the calculated distances, same as Bellman-Ford.

- Note that if there are negative-weight edges, but no negative cycles, Floyd-Warshall will correctly find all distances.
- Every *undirected* graph with a negative-weight edge contains a negative cycle.

- How can we find the actual shortest path?
- As well as keeping track of a dist table, any time the improved path (via i) is used, note that the next thing on the path from u to v is going to be the next thing on the path from u to i, which we should already know because we were keeping track of it!

- When initialising the table with the edges in the graph, don't forget to set ν as next on the path from u to ν for each edge $u \to \nu$.
- Implementing this functionality is left as an exercise.

- Single Source Shortest PathsDijkstra's AlgorithmBellman-Ford Algorithm
- 2 All Pairs Shortest Paths

Implicit Graphs

Implicit Graphs

- Although some graph interpretations are obvious (e.g. cities and highways), it's often the case that the graph you must run your algorithm on is non-obvious.
- Often this doesn't admit a clean implementation using something like an explicit adjacency list.
- In many cases like this, it may be a better idea to compute the adjacencies on the fly.
- Also, sometimes modifying the graph you consider is helpful!

• **Problem Statement** You have found a strange device that has a red button, a blue button, and a display showing a single integer, initially *n*. Pressing the red button multiplies the number by two; pressing the blue button subtracts one from the number. If the number stops being positive, the device breaks. You want the display to show the number *m*. What is the minimum number of button presses to make this happen?

Implicit Graphs

- Input Two space-separated integers n and m $(1 \le n, m \le 10^7)$.
- **Output** A single number, the smallest number of button presses required to get from *n* to *m*.

- In this example, we should think of our button presses as transitions.
- Hence our graph has numbers as its vertices and edges representing which numbers can reach each other through button presses.

Implicit Graphs • The graph is unweighted, so we just need to do a simple BFS to find the answer.

```
#include <br/>
using namespace std;<br/>
const int MAXVAL=20*1000*1000+5;<br/>
const int INF = 1000*1000*1000+7;<br/>
int n, w v[MAXVAL];<br/>
queue<int> Y;<br/>
int main () {<br/>
cin >> n >> m;<br/>
fill(v, v + MAXVAL, 1e9);<br/>
q.push(n);<br/>
v[n] = 0;<br/>
while (!q.empty()) {<br/>
int i = q.front(); q.pop();<br/>
if (i-1 > 0 && v[i] + 1 < v[i-1]) {<br/>
v[i-1] = v[i] + 1;<br/>
q.push(i - 1);<br/>
j (i.x2 < MAXVAL && v[i] + 1 < v[i+2]) {<br/>
v[i*2] = v[i] + 1;<br/>
q.push(i * 2);<br/>
if (x < v[m] << v\n";<br/>
} cout << v[m] << v\n";<br/>
}
```

Implicit Graphs

- **Problem Statement** You are a rock climber trying to climb a wall. On this wall, there are N rock climbing holds for you to use. Whenever you are on the wall, you must be holding on to exactly three holds, each of which can be at most D distance from the other two. To move on the wall, you can only disengage from one of the holds and move it to another hold that is within D distance of the two holds that you are still holding onto. You can move from hold to hold at a rate of 1m/s. How can you reach the highest hold in the shortest amount of time, starting from some position that includes the bottom hold?
- Input A set of up to N $(1 \le N \le 50)$ points on a 2D plane, and some integer D $(1 \le D \le 1000)$. Each point's coordinates will have absolute value less than 1,000,000.
- **Output** A single number, the least amount of time needed to move from the bottom to the top.

- If there was no restriction that required you to always be using three holds, then this would just be a standard shortest path problem that is solvable using Dijkstra's algorithm.
- We would just need to take the points as the vertices and the distance between points as the edge weights.

Implicit Graphs

- However, we need to account for the fact that we must be using three holds clustered together at any time.
- But there is a natural interpretation of the hold restriction in terms of a graph: when we move from some position that uses holds $\{a,b,c\}$ to some position where we use holds $\{a,b,d\}$, we can say that we are moving from some vertex labelled $\{a,b,c\}$ to some vertex labelled $\{a,b,d\}$.
- It can be determined whether or not such a move is allowed, i.e. if there is an edge between these vertices, in constant time.

- Now, we have a graph where we have $O(N^3)$ vertices and $O(N^4)$ edges.
 - Running our shortest path algorithm on this graph directly will give us the answer we want, by definition.

Implicit Graphs

• So we can solve this problem in $O(E\log V) = O(N^4\log N^3) = O(N^4\log N) \text{ time.}$

```
struct state {
    int pid[3];
    int dist;
};

bool operator< (const state &a, const state &b) {
        return a.dist > b.dist;
}

priority_queue<state> pq;
pq.push(begin);
bool running = true;
while (ipq.empty() && running) {
        state cur = pq.top();
        pq.pop();
        /check &if done
        inf (cur.pid[j] = n) {
            running = false;
            break;
        }

// to be continued
```

Implicit Graphs

Implementation (continued)

- **Problem Statement** There are *N* cities, each in one of *C* countries. There are two modes of travel.
- There are A unidirectional direct flight routes connecting two cities, *u*₁, *v*₁ with weight *w*₁.
- There are *B* unidirectional inter-country flights, connecting two countries, a_i, b_i with a weight w_i . These flights can be boarded from any city in the source country and disembarked from in any city in the destination country.

Find the shortest distance from city 1 to city N.

- Input First line, 4 integers N, C, A, B.
- $1 \le N, C, A, B \le 100, 000$. Next line, N integers, c_i , denoting the country the i-th city is in. Next A lines each with 3 integers $u_i, v_i, w_i, 1 \le u_i, v_i \le N, 1 \le w_i \le 10^9$. Next B lines each with 3 integers
 - $a_i, b_i, w_i, 1 \le a_i, b_i \le C, 1 \le w_i \le 10^9.$

Sample Input:

4 3 2 2 1 1 4 100 2 4 2 0 1 1 2 5 0 1 1 3 80

• Sample Output: 70

• **Explanation**: Fly with an intercountry flight of cost 50 to city 2. Then take a direct flight with cost 20 to city 4.

- How to view this as a graph? Without inter-country flights, this is routine.
- With inter-country flights, we could just generate all edges between cities in country a_i and b_i .
 - But this is too many edges.

- **Observation:** We should only consider inter-country flights originating from A the first time we reach a city in A. Similarly, we should only consider inter-country flights to B once.
- This sounds just like how we treat cities in Dijkstra's.
- To encode this, we should treat countries just like cities in our graph, they should have nodes. Be careful, we need 2 nodes per country, one to encode outgoing flights and one to encode incoming flights. What are the edges though?

- The natural ones.
- Cities go to the "outgoing" node for their country.
- The "incoming" node for a country goes to all cities in that country.
- "Outgoing" country nodes connect by inter-country flights to "incoming" country nodes.
- O(N+C) nodes, O(N+A+B) edges. Okay!

```
#include <br/>
using namespace std;<br/>
const int MAXW = 100005;<br/>
const int MAXXW = 100005;<br/>
int W. C. 4. B;<br/>
// (dest, dist)<br/>
// (dest, dist)<br/>
// outgoing country nodes are at MAXW + (country id)<br/>
// intgoing country nodes are at MAXW + MAXC + (country id)<br/>
// intgoing country nodes are at MAXW + MAXC + (country id)<br/>
// intgoing country nodes are at MAXW + AXXC + (country id)<br/>
// int main() {<br/>
cin >> W >> C >> A >> B;<br/>
for (int i = 1; i <= W; i++) {<br/>
    int c; cin >> c; n >> d; i++) {<br/>
    int a, b; long long w;<br/>
    cin >> a >> b >> w;<br/>
    for (int i = 0; i < A; i++) {<br/>
        int a, b; long long w;<br/>
    cin >> a >> b >> w;<br/>
    for (int i = 0; i < B; i++) {<br/>
        int a, b; long long w;<br/>
    cin >> a >> b >> w;<br/>
        for (int i = 0; i < B; i++) {<br/>
        int a, b; long long w;<br/>
        cin >> a >> b >> w;<br/>
        allE[MAXW+all omplace_back(MAXW+MAXC+b);<br/>
    }
}<br/>
// Run your favorite shortest dist algo!
```

• Problem Statement You are at some position on a grid and wish to reach your safe house at some other location on the grid. However, also on certain cells on the grid are enemy safe houses, which you do not want to go near. What is the maximum possible distance you can stay away from every enemy safe house, and still be able to reach your own safe house? When there are multiple paths that keep the same distance from the enemy safe houses, print the shortest one. Distance in this problem is measured by Manhattan distance.

- Input An $N \times M$ grid $(1 \le N, M, \le 1000)$, and the location of your starting point, your safe house, and all the enemy safe houses. There are up to 10,000 enemy safe houses.
- Output Two integers, the maximum distance that you can stay away from every enemy safe house and still be able to reach your safe house, and the shortest length of such a path.

- If there was no restriction stating that you must stay as far away from the enemy safe houses as possible, this would be a simple shortest path problem on a grid.
 - What if we already knew how far we need to stay away from each enemy safe house?

- \bullet Call the distance that we know we need to stay away from the enemy safe houses X .
- We just need to BFS out from every enemy safe house to a distance of X squares, marking all of those squares as unusable. Just marking them as seen will suffice.
- Then we can find the answer with a simple BFS from the starting point. It will ignore the squares that are too close to enemy safe houses because we've marked them as seen.

- How do we view our original optimisation problem in terms of this decision problem?
- Our simpler problem is a decision problem because we answer whether or not it's possible to get from the starting point to the safe house with distance X.
 - The original problem is an optimisation problem because it
 - requires a 'best' answer.

 Observe that if we can stay X distance away from the enemy safe houses, then any smaller distance is also feasible, and if we cannot stay X distance away, then any larger distance is also infeasible.

- This monotonicity allows us to binary search for the largest X such that we can still reach our safe house from our starting point, which we check using the BFS procedure outlined earlier.
- Complexity Each check takes O(NM) time, and we need to perform $\log X_{MAX} = \log(N+M)$ of these checks in our binary search, so this algorithm takes $O(NM\log(N+M))$ total

Implementation

```
const int di[4] = { -1, 1, 0, 0 };
const int dj[4] = { 0, 0, -1, 1 };

vector<pair<int,int>> enemies;

// search from all enemy safe houses to find
// each square's minimum distance to an enemy safe house
queue<pair<int,int>> q;
for (auto it = enemies.begin(); it != enemies.end(); ++it) {
q.push(*it);
}

while (iq.empty()) {
    pair<int,int> enemy = q.front(); q.pop();
    int i = enemy.first, j = enemy.second;
    // try all neighbours
    for (int d = 0; d < 4; ++d) {
        int ni = i + di[d];
        // if off board, *gnore
    if (dist_to_enemy[ni][nj] = -1) continue;
    if (dist_to_enemy[ni][nj] = -1) continue;
    if (dist_to_enemy[ni][nj] = dist_to_enemy[i][j] + 1;
    }
}</pre>
```

Implementation (continued)

```
// binary search
int lo = -1, hi = min(dist_to_enemy[il][jl], dist_to_enemy[i2][j2]),
sol = -1;
while (lo != hi) {
   int X = (lo + hi + 1) / 2;
   // BFS, since the edges are unit weight
   vector<vector(int) > d2(il) | jl);
   q.push(make_pair(il, jl));
   q.push(make_pair(il, jl));
   q.push(make_pair(il, jl));
   int i = q.front() ifirst, j = q.front().second; q.pop();
   for (int d = 0; d < 4; ++d) {
      int i = q.front() ifirst, j = dif[d];
      int (dist[ni][nj] = n) = h | ln nj > h) continue;
   if (dist[ni][nj] = dist[li][j] + 1;
   dist[ni][nj] = dist[li][j] + 1;
   d.push(make_pair(ni, nj));
}

if (dist[li2][j2] == -1) hi = X - 1;
else lo = X, sol = dist[li2];
}
```

- Given M inequalities of the form $x_i x_j \le c_k$, are there real numbers x_1, x_2, \ldots, x_N that satisfy those inequalities?
 - What does this have to do with shortest paths??
- ullet This can actually be solved in O(MN) using Bellman-Ford!

- We need to transform the problem into a graph problem, by creating a graph from the difference constraints.
- Create a vertex for every variable, and for every constraint $x_i-x_j \leq c_k$, create an edge from vertex j to vertex i with weight c_k .

Implicit Graphs • Then, create a source s that has a zero weight edge to every other vertex.

- If we run Bellman-Ford from this source vertex and examine the shortest paths from the source, we obtain a solution.
- For every edge from j to i, from its definition we know that the length of the shortest path to i, say d_i , and the length of the shortest path to j, d_j , satisfy the inequality $d_i \le d_j + c_k$.

Implicit Graphs • We can rewrite $d_i \le d_j + c_k$ as $d_i - d_j \le c_k$, which means that the distances from the source satisfy the difference constraints.

- What happens when there is a negative weight cycle?
- What happens when there is no solution to our system of difference constraints?

• When we have a negative weight cycle, we have a set of edge inequalities with the variables in that cycle:

$$x_1 - x_2 \le c_1$$
$$x_2 - x_3 \le c_2$$

$$x_3-x_1\leq c_3$$

• If we add all these inequalities together, because they form a cycle, all the variables cancel out, giving

$$0 \le c_1 + c_2 + \ldots + c_k,$$

where c_1, c_2, \ldots, c_k are the edge weights on our cycle.

Implicit Graphs • Since the cycle is negative, the RHS adds up to a negative number. Thus we have a contradiction, so there is no solution.

- **Problem Statement** Given a weighted directed graph with non-negative edge weights, what is the length of the second shortest path on the graph from s to t? The second shortest path is the walk with the shortest length that is distinct from the shortest path; in particular, it is the same length as the shortest path if and only if there are multiple shortest paths.
- Input A weighted directed graph with N vertices and M edges (1 \leq N, M \leq 100,000).
- Output A single integer, the length of the second shortest

- Observe that the second shortest walk will differ from a given shortest path by at least one edge.
- Then after finding the shortest path, we can iterate over all edges on the shortest path, and see what happens if we go along a different edge instead.

- Let the shortest path from u to v have length d_{uv} . If some edge (u,v) is blocked, and we take some other edge (u,w) instead, with weight e_{uw} we get a path of length $d_{su} + e_{uw} + d_{wt}$.
 - Since we're iterating over all smallest possible differences to the shortest path, one of these distances will be our answer.
- So we just try them all (i.e. all choices of (u, w)) and take the one of minimum length.
- \bullet How do we efficiently compute all the $d_{wt}?$

Implementation

Implementation (continued)

```
set<edge> s;
s.insert(edge(0, n));
vctor&cool> seen(n+1);
prev2[n] = -2;
for (s.insert(edge(0, n)); is.empty(); s.erase(s.begin())) {
    dge cur = *s.begin();
    if (seen[cur.y] = cur.x;
        seen[cur.y] = true;
    foreach(v, adj[cur.y]) {
        if (issen[v-y])
        if (issen[v-y])
        if (issen[v-y])
        if (dist2[v-y]) = -1 || dist2[v-y] > cur.x+v->x) {
            prev2[v-y])
        is insert(edge(cur.x+v->x, v-y));
}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           int res = 1<<30;
for (int v = n; v != -2; v = prev1[v])
foreach(e, adj[v])
res = min(res, dist1[v] + e->x + dist2[e->y]);
// dist2[v] = shortest distance from v to n vector(int) dist2(n+1, -1);

vector(int) prev2(n+1, -1);
```