

OME: Optimisation Methods

1 Important Notes About These Notes

These notes were written by me, Mark Ormesher, during my revision for Computer Science exams at King's College London. I am publishing them here to help my classmates and other students in the years below me. To the best of my knowledge they are correct.

These notes are **not endorsed** by King's College London, the module lecturers, or any member of College staff. These notes are **not checked** by any qualified individual, and I offer no guarantee as to how accurate or complete these notes are. They are offered for free, as-is, and I accept no responsibility for how you choose to use them.

These notes were relevant for my year (2016/17) but **the content for your course may be different**. Check your lecture slides, syllabi, tutorial guides, etc.

These notes were produced for my own personal use (it's part of how I study and revise). That means that some annotations may be irrelevant to you, and **some topics might be skipped** entirely.

Feel free to **share** these notes, however please only share a link to the repo (see the link below), not individual files.

Notes are originally from **<https://github.com/markormesher/CS-Notes>**. All original work is and shall remain the copyright-protected work of Mark Ormesher. Any excerpts of other works, if present, are considered to be protected under a policy of fair use.

Contents

1	Important Notes About These Notes	1
2	Single-Source Shortest Paths	4
2.1	Notation	4
2.2	Useful Facts	4
2.3	Negative Weights	5
2.3.1	Exchange Rate Example	5
2.3.2	Negative Cycles	5
2.4	Shortest-Path Trees	5
2.5	Relaxation Technique	6
2.5.1	Initialisation	6
2.5.2	Relaxation	7
2.5.3	Relaxation Properties	7
2.5.4	Checking for Cycles	8
2.6	Bellman-Ford Algorithm	8
2.6.1	Correctness	9
2.6.2	Running Time	9
2.6.3	Improving the Running Time	10
2.6.4	Optimisations: Early Termination and Queue of Active Vertices	10
2.7	Dijkstra's Algorithm	11
2.7.1	Running Time	12
2.7.2	Correctness - Invariants	12
2.7.3	Correctness - Induction on Invariants	13
2.8	Directed Acyclic Graphs (DAGs)	14
2.8.1	Topological Order	14
2.9	Geographic Networks	15
2.9.1	Re-Weighting Edges	15
3	All-Pairs Shortest Paths	16
3.1	Existing Algorithms	16
3.2	Floyd-Warshall Algorithm	16
3.3	Johnson's Algorithm	16
3.3.1	Re-Weighting Edges	16
3.3.2	Calculating h -values	17
3.3.3	Full Algorithm	18
3.3.4	Running Time	18
3.3.5	Correctness	18
4	Network Flow Problems	19
4.1	Notation	19
4.2	Types of Flow Problem	19
4.3	Flow Formalisation	19
4.4	From Flows to Paths	20
5	Max-Flow Problems	21
5.1	Residual Capacity	21
5.1.1	Adding Flow Using the Residual Network	21
5.2	General Approach for Finding Maximum Flow	22

5.3	Side Note: Cuts	22
5.3.1	Theorem One: Max-Flow Min-Cut Theorem	23
5.3.2	Theorem Two	23
5.4	Ford-Fulkerson Algorithm	24
5.5	Edmonds-Karp Algorithm	24
6	Flow Feasibility Problems	26
6.1	Reduction to Maximum Flow Problem	26
7	Minimum Cost Flow Problems	28
7.1	Notation	28
7.2	Sending Flow Along Cheapest Paths	28
7.3	Residual Network	29
7.4	Successive Shortest Path Algorithm	30
7.4.1	Correctness and Runtime	31
8	Multi-Commodity Flow Problems	32
8.1	Notation	32
8.2	Example	32
8.3	Flows of Commodities	33
8.4	Types of Multi-Commodity Flow Problems	33
8.4.1	Feasibility	33
8.4.2	Minimum Cost	33
8.4.3	Minimum Congestion	33
8.5	Computational Approaches	34
9	Maximum Bipartite Matching	35
9.1	Computing a Maximum Matching	35
10	Linear Programming	36
11	Optimisations for NP-Hard Problems	37

2 Single-Source Shortest Paths

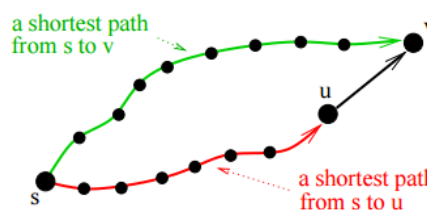
Objective: given a weighted directed graph, find the **minimum-weight paths** from a single vertex s to **all other vertices**. Algorithms that solve this problem can also be adapted to solve **point-to-point** problems.

2.1 Notation

- $G = (V, E)$ is a **directed graph** G with vertices V and edges E .
 - Vertices may also be referred to as nodes or sites.
 - $|V| = n$, $|E| = m$.
- $w(u, v)$ is the **weight of the edge** (u, v) .
- $s \in E$ is the **source** vertex.
- $p = \langle v_1, v_2, v_3, \dots, v_k \rangle$ is a **path** from vertex v_1 to v_k via v_2, v_3 , etc.
- $w(p) = w(v_1, v_2) + w(v_2, v_3) + \dots + w(v_{k-1}, v_k)$ is the **weight of a path**, equal to the sum of the weights of all edges in the path.
- $\delta(u, v)$ is the **minimum of** $w(p)$ for all p from u to v , or $+\infty$ if no such p exists.
 - Multiple minimum-weight paths may exist.
 - A shortest path from u to v is any path $p = \langle u, \dots, v \rangle$ such that $w(p) = \delta(u, v)$.

2.2 Useful Facts

- A sub-path of a shortest path is also a shortest path.
 - If a shortest path from a to e is $\langle a, b, c, d, e \rangle$, a shortest path from b to d is $\langle b, c, d \rangle$.
- For each edge $(u, v) \in E$, it holds that $\delta(s, v) \leq \delta(s, u) + w(u, v)$.
 - This states that the weight of the shortest path between s and v is **at most** the weight from s to u , plus the weight of the edge from u to v .
 - This is an example of the **triangle inequality** and is shown in the following diagram:



2.3 Negative Weights

Negative weights can be useful/essential in solving a problem - the most obvious example is that of finding the **maximum-weight** paths. Rather than create new algorithms, it is possible to negate every edge and then find the minimum-weight paths as normal.

The **minimum-weight paths for negated edges** are equivalent to the **maximum-weight paths for non-negated edges**.

2.3.1 Exchange Rate Example

Taking the example of a graph showing exchange rates between currencies, the most profitable path **maximises the product** of its edge weights. This is unsuitable, because the algorithms we study **minimise the sum** of edges along a path.

This can be resolved by taking the negative logarithm of every edge. If $\gamma(u, v)$ is the exchange rate from u to v , $w(u, v) = -\ln(\gamma(u, v))$.¹ This works because the sum of logarithms of a set of numbers is equal to the logarithm of their products².

However, we **cannot avoid negative weights** now: if $\gamma(u, v) > 1$ then $w(u, v) < 0$.

2.3.2 Negative Cycles

If a negative cycle exists on a path from s to v , repeated journeys around this cycle will continually reduce the weight of the path. Therefore, if a negative cycle exists on a path from s to v , $w(s, v) = -\infty$.

We consider only shortest-paths algorithms that:

- Compute the shortest paths for graphs where **no negative cycles are reachable** from the source.
- **Correctly detect when negative cycles are reachable** from the source and are not required to compute anything else.

Why not look for the **shortest simple paths** (i.e. avoid cycles). It's a valid question, but it's NP-hard and we do not cover it in this section. [See more: Optimisations for NP-Hard Problems, page 37.](#)

2.4 Shortest-Path Trees

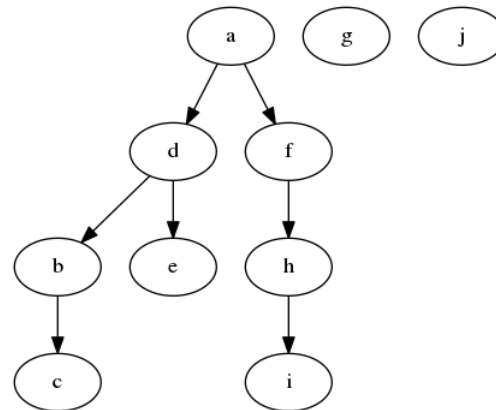
For each vertex v reachable from the source s , we want to find any one of the minimum-weight paths that exist. If there are no negative cycles, these paths can be easily represented with a **shortest-paths tree**. Such a tree contains exactly one shortest path from s to each v reachable from s .

¹Any fixed logarithm base greater than 1 would work.

² $\ln(a) + \ln(b) + \ln(c) = \ln(abc)$

The tree T can be represented by an array $parent[v], v \in V$ in $\Theta(n)$ space. $parent[v]$ is the parent of the vertex v in the tree T . An optional array $d[]$ can store the weight of the shortest paths to each vertex. For example, from the source vertex a :

v	a	b	c	d	e	f	g	h	i	j
$parent[v]$	nil	d	b	a	d	a	nil	f	h	nil
$d[v]$	0	6	9	2	4	3	∞	5	6	∞



A disadvantage of this representation is that only one path for each reachable vertex can be stored. Listing every path explicitly leads to $O(n^2)$ space requirements.

2.5 Relaxation Technique

This is the most common approach for solving single-source shortest paths problems. It uses an **initialisation** step followed by a sequence of **relaxation** steps. Algorithms following this pattern differ in their **termination condition(s)** and the **order that edges are relaxed**.

At each vertex v , the following is maintained:

- $d[v]$ is the shortest-path **estimate** for reaching v from the source s . It is an upper-bound.
- $parent[v]$ is the current predecessor of v .

2.5.1 Initialisation

This stage sets the upper-bound for all non-source vertices at infinity and creates no parent relationships.

```

1 fun INITIALISE(G, s):
2     d[s] = 0
3     parent[s] = nil
4     for each vertex v in V - {s}, do:
5         d[v] = inf
6         parent[v] = nil
  
```

Initialisation is a $\Theta(n)$ operation. For large graphs this can be too slow, so an alternative is to only initialise s at the start of the algorithm and then **initialise other vertices as they are discovered**.

2.5.2 Relaxation

Applied to an edge (u, v) , this stage checks whether a path to v via u would have a lower weight than the current estimate of reaching v , $d[v]$. If so, $d[v]$ is updated with the weight of the path via u and the parent relationship is updated so that $parent[v] = u$. This process stems from the triangle inequality ([see more: Useful Facts, page 4](#)).

```

1 fun RELAX(u, v, w):
2     if (d[v] > d[u] + w(u, v)):
3         d[v] = d[u] + w(u, v)
4         parent[v] = u

```

Relaxation is a $\Theta(1)$ operation.

2.5.3 Relaxation Properties

- For every vertex v , throughout the computation...
 - $d[v]$ will only **decrease**.
 - $d[v]$ is either ∞ or the weight of some path from s to v .
 - $d[v] \geq \delta(s, v)$.
- An **effective relaxation operation** is defined as an operation that decreases $d[v]$.
- During computation, there is an edge $(u, v) \in E$ such that $d[v] > d[u] + w(u, v)$ (i.e. there is an effective relaxation operation) if and only if there is a vertex x such that $d[x] > \delta(s, x)$.
 - From this, it follows that if no effective relaxation operations remain then all shortest paths have been found.
- If there **is no negative cycle** reachable from s ...
 - There will be a **finite number of effective relaxation operations**.
 - The *parent* pointers will form a tree rooted at s .
 - When no effective relaxation operations remain, then for each vertex v , $d[v] = \delta(s, v)$ and *parent* $[v]$ points to v 's predecessor on a shortest path from s .
- If there **is a negative cycle** reachable from s ...
 - There will always be an effective relaxation operation that can be applied.
 - The *parent* pointers will eventually form a cycle. This suggests that negative cycles in the graph can be detected by periodically checking *parent* for cycles.

2.5.4 Checking for Cycles

If the *parent* array contains a cycle then the graph contains a negative cycle, so computation should indicate this and terminate.

A cycle can be detected by following parent pointers from all vertices, marking vertices as visited during traversal, and reducing work by avoiding re-visiting any vertices. On each traversal, if a vertex that was already seen is encountered then a cycle exists.

```
1 fun CHECK-CYCLE(parent):  
2   visited = [false, false, ...]  
3   for each v in V, do:  
4     if (!visited[v]):  
5       seen = { }  
6       while (true):  
7         visited[v] = true  
8         seen.add(v)  
9  
10        next = parent[v]  
11        if (next == nil):  
12          break  
13  
14        if (seen.contains(next)):  
15          return CYCLE  
16  
17        if (visited[next]):  
18          break  
19  
20        v = next  
21  
22   return NO_CYCLE
```

2.6 Bellman-Ford Algorithm

The Bellman-Ford algorithm is a **relaxation-based algorithm** for finding shortest single-sources paths of a graph. It **permits negative-weight edges**.


```

1 fun BELLMAN-FORD(G, w, s):
2
3     // run initialisation and relaxation:
4     INITIALISE(G, s)
5     for i from 1 to n - 1, do:
6         for each edge (u, v) in G, do:
7             RELAX(u, v, w)
8
9     // determine outcome:
10    for each edge (u, v) in G, do:
11        if (d[v] > d[u] + w(u, v)):
12            return FALSE // negative cycle reachable from s
13        else:
14            return TRUE // no negative cycle, d[] now holds shortest path weights

```

2.6.1 Correctness

If a **negative cycle is reachable** from *s* then the algorithm will always **correctly** return **FALSE** because an effective relaxation will always be possible.

If **no negative cycles are reachable** from *s* then the algorithm will always **correctly** return **TRUE** because at the end of the first section no more effective relaxations will be possible. This is guaranteed, because of two lemmas:

- If the shortest simple path from *s* to *v* is of length *k*, then after *k* iterations of the first loop $d[v] = \delta(s, v)$.
 - Assume the *k*-length path is $\langle s, x_1, x_2, \dots, v \rangle$.
 - After the first iteration $d[x_1] = \delta(s, x_1)$.
 - After the second iteration $d[x_2] = \delta(s, x_2)$.
 - ...
 - After *k* iterations $d[k] = \delta(s, k)$.
- The shortest path from *s* to any vertex contains at most $n - 1$ edges, because the longest non-cyclic path will visit every vertex once.

2.6.2 Running Time

This algorithm **runs the $\Theta(1)$ relaxation step exactly $n - 1$ times** for each edge in the first section, then in the second section it checks to see if any further relaxations are possible. If more effective relaxations are possible, a negative cycle must exist and **FALSE** is returned; if no effective relaxations are possible then the shortest path weights have been found and **TRUE** is returned.

The running time is $\Theta(nm)$. The worst case for **any** single-source shortest paths problem with negative weights is $\Omega(nm)$.

2.6.3 Improving the Running Time

The $\Theta(nm)$ running time comes from the first section of the algorithm, so we focus our efforts there:

- We could try to **decrease the number of iterations** of the outer loop by terminating it early in certain conditions:
 - Not every graph will require $n - 1$ iterations to find the solution. If **no effective relaxation operations took place** in a given iteration the loop can **terminate early** because the shortest paths were already computed.
 - Some iteration may create a cycle in the *parent* array. It can be checked periodically (after every iteration?); if a *parent* **cycle is found** then a negative loop is reachable from s and the entire algorithm can **terminate early** with `FALSE`.
- We could try to **reduce the work done in each iteration** of the outer loop by considering only edges that will give effective relaxations.
 - A vertex u is **active** if its outgoing edges have not been relaxed since the last time $d[u]$ was decreased.
 - We perform relaxations only on the edges out of active vertices.
 - We store active vertices in a **FIFO** queue.

2.6.4 Optimisations: Early Termination and Queue of Active Vertices

This optimisation keeps track of active vertices and **only considers relaxations of their outgoing edges**, as described above.

```

1 fun BELLMAN-FORD-FIFO(G, w, s):
2
3     // initialisation
4     INITIALISE(G, s)
5     Q = empty queue // active vertices
6     Q.enqueue(s)
7
8     // relaxation of active vertices
9     while (Q is not empty), do:
10         u = Q.dequeue()
11         for each v in adj[u], do:
12             RELAX(u, v, w)
13
14         if (CHECK-CYCLE(parent)):
15             return FALSE
16
17     return TRUE

```

The relaxation method is augmented to **enqueue the vertex** v if $d[v]$ is updated when considering the edge (u, v) (i.e. if it is 'active').

```

1 fun RELAX(u, v, w):
2     if (d[v] > d[u] + w(u, v)):
3         d[v] = d[u] + w(u, v)
4         parent[v] = u
5
6         // enqueue v if it was updated
7         if (v is not in Q):
8             Q.enqueue(v)

```

Note that a mechanism for detecting negative cycles has been added, because Q will never be empty if a negative cycle is reachable from s . This could be in the form of checking for cycles in *parent* after each set of relaxations.

The running time is now $O(nm)$, but still $\Theta(nm)$ in the worst case.

2.7 Dijkstra's Algorithm

- Critical assumption: **all weights are non-negative.**
 - As a consequence, there are also no negative cycles.
- We assume that all vertices are reachable from s .

A **set** S is maintained of all vertices for which $d[v] = \delta(s, v)$ (i.e. vertices that are 'finished'). A **min-priority queue** Q is maintained of all non-finished vertices and their current path weight estimate.

```

1 fun DIJKSTRA(G, w, s):
2     INITIALISE(G, s)
3     S = []
4     Q = V
5
6     while (Q is not empty), do:
7         u = Q.removeMin()
8         S.add(u)
9         for each v in adj[u], do:
10             RELAX(u, v)

```

The relaxation method is augmented to **update the min-priority queue** value of the vertex v if $d[v]$ is updated when considering the edge (u, v) .

```

1 fun RELAX(u, v, w):
2     if (d[v] > d[u] + w(u, v)):
3         d[v] = d[u] + w(u, v)
4         parent[v] = u
5         Q.decreaseKey(v, d[v])

```

2.7.1 Running Time

- The data structures will be instantiated $\Theta(1)$ times.
- The main **while** loop will run $\Theta(n)$ times, because all vertices start in the queue, each vertex is removed, and no vertex is replaced.
- A total of $\Theta(n)$ calls to `Q.removeMin()` will be made.
- A total of $\Theta(m)$ relaxations will take place, each of which may include a call to `Q.decreaseKey()`.

It is clear that the priority queue implementation will be the main factor in the algorithm's running time. There are three implementations to consider, each with different running times:

- **Unsorted array:**
 - Instantiation is $\Theta(n)$.
 - `removeMin()` requires an $O(n)$ search and will be done $\Theta(n)$ times.
 - `decreaseKey()` requires a $\Theta(1)$ update and will be done $O(m)$ times.
 - The total running time is $\Theta(n) + O(n^2) + \Theta(m) = O(n^2)$.
- **Sorted array:**
 - Instantiation is $\Theta(n)$.
 - `removeMin()` requires a $\Theta(1)$ removal and will be done $\Theta(n)$ times (this assumes that values are not shifted to fill the gap created).
 - `decreaseKey()` requires an $O(n)$ re-ordering and will be done $O(m)$ times.
 - The total running time is $\Theta(n) + \Theta(1) + O(mn) = O(mn)$.
 - * Note: $n - 1 \leq m \leq n(n - 1)$, so the total is $O(n^3)$ for a dense graph.
- **Heap:**
 - Instantiation is $\Theta(n)$.
 - `removeMin()` requires an $O(\log_2(n))$ removal and will be done $\Theta(n)$ times.
 - `decreaseKey()` requires an $O(\log_2(n))$ re-ordering and will be done $O(m)$ times.
 - The total running time is $\Theta(n) + O(n \cdot \log_2(n)) + O(m \cdot \log_2(n)) = O(m \cdot \log_2(n))$.

If the input graph is **dense**, such that $m = \Omega(n^2/\log_2(n))$, then the **unsorted array** implementation of the priority queue gives a better worst-case running time. If the graph is **not dense**, the **heap implementation** gives a better worst-case running time.

For most applications the input graphs are **not dense**, so Dijkstra's algorithm is **assumed to use a heap-based priority queue**.

2.7.2 Correctness - Invariants

We establish **invariants** for the algorithm, based on each vertex's membership in S or Q .

- At the end of the algorithm's execution $S = V$, so only the first set of invariants apply for each vertex. These state that $d[v] = \delta(s, v)$ for all vertices v and that the current three is the shortest path three from s to all other vertices. Therefore, if these invariants hold, the algorithm is correct.

Basis step: prove that the invariants hold for the first iteration ($i = 0$).

13

- Assume that the invariants hold at the end of some iteration i (possibly the first iteration, described above).
- Let u denote the vertex selected in the iteration $i + 1$.
- Show that the invariants hold at the end of iteration $i + 1$ when u has moved from Q to S .

Proving Invariant 1b

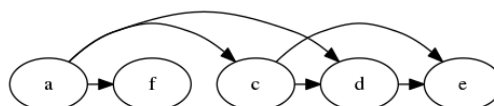
- We have selected vertex u to add to S , so we must prove that at the end of the iteration $d[u] = \delta(s, u)$.
- We already know that $d[u] \geq \delta(s, u)$ because it is a basic property of the relaxation operation, so we must prove $d[u] \leq \delta(s, u)$.
- Select any path R from s to u and show that $w(R) \geq d[u]$.
 - Split R into $R1$ and $R2$, such that $R1$ ends at the first vertex y outside of S .
 - $w(R) = w(R1) + w(R2) \geq d[y] + w(R2) \geq d[u] + w(R2) \geq d[u]$
 - Therefore, no path from s to u has a smaller weight than $d[u]$, so $d[y] \leq \delta(s, u)$.
- This works because u has the lowest d -value, therefore $d[y] \geq d[u]$, and all edge weights are non-negative, therefore $d[y] + w(R2) \geq d[u]$.

2.8 Directed Acyclic Graphs (DAGs)

The objective is the same as before: find the **minimum-weight paths** from a single vertex s to **all other vertices**. Dijkstra's algorithm won't work because negative-edges are possible; the Bellman-Ford algorithm could work in $O(mn)$, but the restricted graph format (DAG) allows for a $\Theta(n + m)$ algorithm.

2.8.1 Topological Order

The algorithm relies on creating a **topological order** for the vertices in the graph. A topological order rearranges vertices (without changing edges) so that all edges point in the same direction, as pictured:



Once the topological order has been created, the algorithm is simple:

```

1 fun DAG-SHORTEST-PATHS (G, w, s):
2     TOPO-SORT (G)
3     INITIALISE (G)
4     for each vertex u in topological order, do:
5         for each vertex v in adj[u], do:
6             RELAX(u, v, w)

```

The running time is made up of one $\Theta(n + m)$ topological order generation, one $\Theta(n)$ initialisation and $\Theta(m)$ relaxations, giving $\Theta(n + m)$.

The algorithm works because when a vertex is considered all of its incoming edges have already been relaxed, therefore $d[u] = \delta(s, u)$ when u is considered.

2.9 Geographic Networks

When looking for a **single destination** d , one approach is to run Dijkstra's algorithm from the source and stop when d is considered (i.e. removed from Q). This may require checking the entire network, wasting a lot of effort.

Another approach is to **run two Dijkstra computations**, one searching 'forwards' from the source and another searching 'backwards' from the destination. When a shortest path from s to d is found both computations can stop. This **bi-directional** search can result in a smaller part of the networking being examined, but is very tricky to implement with an optimal stopping condition.

When we have a geographic network we know the **coordinates** of each vertex. With this extra information we can compute the **straight-line distance between vertices**, which is a lower-bound on the shortest paths between them. Straight-line distances to the destination can be used to **re-weight** edges.

2.9.1 Re-Weighting Edges

Graph edges can be re-weighted so that edges leading geographically *towards* d become cheaper and edges leading *away* from d get more expensive. In a similar fashion to Johnson's Algorithm ([see more: Johnson's Algorithm, page 16](#)), edges can re-weighted as follows:

$$\hat{w}(u, v) = w(u, v) - \text{straight_line}(u, d) + \text{straight_line}(v, d)$$

Values in \hat{w} are always ≥ 0 , satisfying the requirements of Dijkstra's algorithm.

Note that this is a **heuristic** that will improve on the average-case running time, but **will not improve the worst-case** scenario.

3 All-Pairs Shortest Paths

Objective: given a weighted directed graph, find the **minimum-weight path** from **all vertices to all other vertices**. This will create $n(n - 1)$ paths and weights.

The output is two $n * n$ matrices:

- D , such that $D[i, j] = \delta(i, j)$.
- P , such that $P[i, j]$ holds the parent of j in a path from i to j .

For simplicity, we will only consider computation of the first matrix.

3.1 Existing Algorithms

- If all edge weights are **non-negative** we can run **Dijkstra's algorithm** from each vertex.
 - $n \cdot O(\min\{m \cdot \log_2(n), n^2\}) = O(\min\{mn \cdot \log_2(n), n^3\})$.
 - This is the best worst-case running time that is known.
- In the general case where negative edges are allowed we can run the **Bellman-Ford** algorithm from each vertex.
 - $n \cdot O(mn) = O(mn^2)$, up to $O(n^4)$.

3.2 Floyd-Warshall Algorithm

This algorithm runs in $\Theta(n^3)$ and is not covered in this course.

3.3 Johnson's Algorithm

This algorithm runs in $O(\min\{mn \cdot \log_2(n), n^3\})$, making it **suitable for sparse graphs** but **unsuitable for dense graphs**.

It works by **re-weighting edges** to remove negatives, then applying **Dijkstra's algorithm** at every vertex.

3.3.1 Re-Weighting Edges

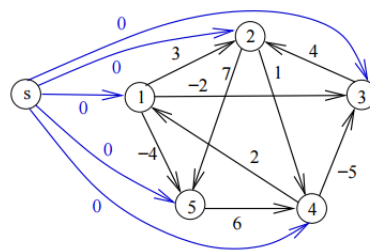
We must change the weight of edges to **remove negatives without changing shortest paths**. Adding some large value to each edge won't work because the shortest paths may change (because the path length becomes a major component in its weight).

Johnson's algorithm re-weights edges by first assigning a value $h(v)$ to each vertex v and then creating an updated set of weights \hat{w} such that $\hat{w}(i, j) = w(i, j) + h(i) - h(j)$. That is, each edge weight is increased by its source vertex h -value and decreased by its destination vertex h -value.

For any path the h -values of all intermediate nodes **cancel each other out**. For a path $P = \langle v_1, v_2, \dots, v_k \rangle$ its new weight is calculated as $\hat{w}(P) = w(P) + h(v_1) - h(v_k)$. All paths between v_1 and v_k are offset by the same amount regardless of their length, so shortest paths are not affected.

3.3.2 Calculating h -values

The **Bellman-Ford** algorithm can be used to calculate the h -value for each vertex. An 'imaginary vertex' s is created with a 0-weight edge to all vertices of the graph, and then the shortest path weights are found from s . Once complete, $h(v) = d[v] = \delta(s, v)$ for all vertices in the 'real' graph.



3.3.3 Full Algorithm

```

1 fun JOHNSON(G, w):
2     D = output matrix
3
4     // add imaginary vertex
5     V' = V + { s }
6     E' = E + { (s, v) for all v in V }
7     w(s, v) = 0 for all v in V
8     G' = (V', E')
9
10    // detect negative cycle or find h-values
11    if (BELLMAN-FORD(G', w, s) == false):
12        return false // negative cycle
13    else:
14        // re-weight edges
15        for each vertex v in V, do:
16            h(v) = d[v]
17        for each edge (u, v) in E, do:
18            w'(u, v) = w(u, v) + h(u) - h(v)
19
20    // run Dijkstra and record (adjusted) result in D
21    for each vertex u in V, do:
22        d' = DIJKSTRA(G, w', u)
23        for each vertex v in d', do:
24            D[u, v] = d'[v] - h(u) + h(v)
25
26    return D

```

3.3.4 Running Time

The running time is made up of one iteration of the Bellman-Ford algorithm and n iterations of Dijkstra's algorithm.

$$O(mn) + n \cdot O(\min\{m \cdot \log_2(n), n^2\}) = O(\min\{mn \cdot \log_2(n), n^3\})$$

3.3.5 Correctness

- **Detecting negative cycles:** we know that the Bellman-Ford algorithm can do this correctly and that the 'imaginary vertex' s cannot create a cycle (because it has no incoming edges), so therefore this algorithm can correctly detect negative cycles.
- **Computing shortest path weights:** we know that Dijkstra's algorithm can do this correctly and we apply it at every vertex, so therefore this algorithm can correctly compute the shortest path weights.

4 Network Flow Problems

4.1 Notation

- $G = (V, E, c, s, t)$ is a **flow network** G with vertices V and edges E .
 - Vertices may also be referred to as nodes or sites.
 - $|V| = n, |E| = m$.
- $c(u, v)$ is the **capacity of the edge** $(u, v) \in E$.
- $s \in E$ is the **source** vertex.
- $t \in E$ is the **sink/destination** vertex.
- $s \neq t$.

4.2 Types of Flow Problem

- **Max-Flow**: find a **maximum flow** from the **source** to the **sink** of a flow network. A maximum flow sends the **maximum amount of the underlying commodity** from the source to the sink without exceeding the capacity of any edge.
- **Flow Feasibility** (Transshipment Problem): find a flow which satisfies edge capacities *and* the specified supply/demand values at various vertices.
 - This can be reduced to the maximum flow problem.
- **Minimum Cost Flow**: find a **maximum flow** or a **supply/demand satisfying flow** that minimises the cost incurred by using each edge.
- **Multi-Commodity Flow**: find a **maximum flow**.

4.3 Flow Formalisation

We assume that if $(u, v) \in E$ then $(v, u) \in E$. Edges with zero capacity are added when required, but usually not shown on diagrams.

A **flow is a function** $f : E \mapsto \mathbb{R}$ where $f(u, v) \geq 0$ is the flow on the edge (u, v) with the following properties:

- **Capacity constraints**: for each edge $(u, v) \in E$, $0 \leq f(u, v) \leq c(u, v)$.
- **Flow conservation**: for each vertex $v \in V - \{s, t\}$, the total flow in to v is equal to the total flow out of v (i.e. the net flow through v is zero).
- **Single-direction flow**: for each edge $(u, v) \in E$, if $f(u, v) > 0$ then $f(v, u) = 0$.

The **value of a flow** is the net flow into the sink (which is the same as the net flow out of the source, because of flow conservation):

$$|f| = \sum_{(x,t) \in E} f(x,t) - \sum_{(t,z) \in E} f(t,z)$$

$$|f| = \sum_{(s,x) \in E} f(s,x) - \sum_{(z,s) \in E} f(z,s)$$

For a given flow f , if $f(u,v) = c(u,v)$ then f is said to **saturate** the edge (u,v) and the edge (u,v) is said to be a **saturated edge**.

4.4 From Flows to Paths

Given a network $G = (V, E, c, s, t)$ and a flow $f : E \mapsto \mathbb{R}$, the flow f can be **decomposed** into at most m paths from s to t , where $m = |E|$.

Algorithm: **select and remove maximal $\langle s, \dots, t \rangle$ flow paths** from f until f is empty. Each path removes all remaining flow from at least one edge, so at most m paths are created.

Each path can be found in $O(n)$ and there are $O(m)$ paths, so a flow can be decomposed into paths in $O(mn)$. This is less than the time needed to find a maximum flow, so this is okay.

Note that this algorithm does not work when **flow cycles** exist, but can be modified to accommodate them.

5 Max-Flow Problems

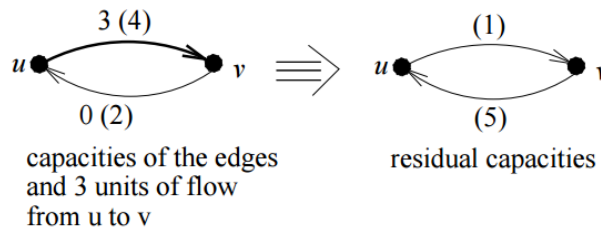
Maximum flow problems are solved by continually adding flow to a network. The amount that can be added (i.e. the network's remaining capacity) is the **residual capacity**.

5.1 Residual Capacity

If f is a flow defined in the network $G = (V, E, c, s, t)$, then residual edges c_f are defined as follows:

- $c_f(u, v) = c(u, v) - f(u, v)$ if $f(u, v) \geq 0$ and $f(v, u) = 0$.
 - When there is flow from u to v , the residual capacity of (u, v) decreases by the amount of that flow.
- $c_f(u, v) = c(u, v) + f(u, v)$ if $f(u, v) = 0$ and $f(v, u) \geq 0$.
 - Where there is flow from v to u , the residual capacity of (u, v) increases by the amount of that flow.

In the example below, $c(u, v) = 4$, $c(v, u) = 2$ and $f(u, v) = 3$. The residual edges $c_f(u, v) = 1$ and $c_f(v, u) = 5$ are created.



This shows that we can send 1 more unit of flow from u to v , or we can send 5 unit from v to u by 'sending back' the 3 that are already being sent in the other direction, plus saturating the edge (v, u) .

If $c_f(u, v) > 0$ then (u, v) is a **residual edge**.

The **residual network** of G induced by the flow f is the flow network $G_f(V, E_f, c_f, s, t)$ where c_f are the residual capacities and E_f is the set of residual edges.

5.1.1 Adding Flow Using the Residual Network

If f is a flow in the network G and f' is a flow in the residual network G_f , then f and f' can be combined to generate a new flow h in G .

The new flow $h = f \oplus f'$ in G is defined as follows:

For each (u, v) such that $f(u, v) \geq 0$ and $f(v, u) = 0$:

- If $f'(u, v) \geq 0$ and $f'(v, u) = 0$, then:
 - $h(u, v) = f(u, v) + f'(u, v)$
 - $h(v, u) = 0$
 - i.e. when the flows go in the **same direction** just add them together and leave the opposite direction as zero.
- If $f'(u, v) = 0$ and $f'(v, u) > 0$, then:
 - $h(u, v) = \max\{f(u, v) - f'(v, u), 0\}$
 - $h(v, u) = \max\{f'(v, u) - f(u, v), 0\}$
 - i.e. when the flows go in **opposite directions**, cancel out the smaller flow and add whatever is left.

The value of h is defined as $|h| = |f| + |f'|$.

5.2 General Approach for Finding Maximum Flow

- Start with f as a zero flow (i.e. $f(u, v) = 0$ for all $(u, v) \in E$).
- Loop forever:
 - Construct the residual network G_f .
 - Find a non-zero flow f' in G_f .
 - * If no such flow exists, exit loop.
 - $f = f \oplus f'$ (combine the flows)
- Return f as the maximum flow.

Most maximum flow algorithms use this approach.

5.3 Side Note: Cuts

A cut (S, T) in a network divides the nodes into **two disjoint groups** S and T such that $S \subseteq V$ and $T = V - S$. The source and sink nodes are in different groups, such that $s \in S$ and $t \in T$.

The **capacity** of a cut is the sum of capacities of all edges from S to T :

$$c(S, T) = \sum_{(u,v) \in E: u \in S, v \in T} c(u, v)$$

The **net flow** across a cut is the sum of net flows leaving S minus the sum of net flows coming in to S :

$$f(S, T) = \sum_{(u,v) \in E: u \in S, v \in T} f(u, v) - \sum_{(x,y) \in E: x \in T, y \in S} f(x, y)$$

By the conservation of flow property, for any cut (S, T) on a network with some flow f , it holds that $f(S, T) = |f|$.

For any cut (S, T) on a network with some flow f , it also holds that $|f| = f(S, T) \leq c(S, T)$. This is intuitive: the net flow across a cut **cannot exceed the capacity** of that cut.

Therefore, the **maximum flow in a network is not greater than the minimum capacity of a cut**:

$$\max\{|f| : f \text{ is a flow in } G\} \leq \min\{c(S, T) : (S, T) \text{ is a cut in } G\}$$

5.3.1 Theorem One: Max-Flow Min-Cut Theorem

The maximum value of a flow is equal to the minimum value of a cut.

$$\max\{|f| : f \text{ is a flow in } G\} = \min\{c(S, T) : (S, T) \text{ is a cut in } G\}$$

5.3.2 Theorem Two

For a flow f in G , the following conditions are equivalent (*i.e. all are true, or all are false*).

- (a) f is a maximum flow in G .
- (b) There is no augmenting path in the residual network G_f .
- (c) There exists a cut (S', T') in G such that $f(S', T') = c(S', T')$.

Theorem two can be shown by proving that $(a) \implies (b) \implies (c) \implies (a)$.

- $(a) \implies (c)$: if f is a maximum flow, there is a cut (S', T') such that $f(S', T') = c(S', T')$.
 - We know that $|f| = f(S', T')$.
 - We know that $f(S', T') = c(S', T')$ from (c).
 - We know that $c(S', T') \geq \min\{c(S, T) : (S, T) \text{ is a cut in } G\}$, because it cannot be smaller than the smallest.
 - We know that $c(S', T') \not\leq \min\{c(S, T) : (S, T) \text{ is a cut in } G\}$, because no flow can be greater than the minimum cut.

- Therefore, $|f| = f(S', T') = c(S', T') = \min\{c(S, T) : (S, T) \text{ is a cut in } G\}$.
- (a) \implies (b): if f is a maximum flow, there is no augmenting path in G_f .
 - If there is an augmenting path then f cannot be a maximum flow, because the path could be used to add more flow.
- (b) \implies (c): if there is no augmenting path in G_f , a cut (S', T') exists such that $f(S', T') = c(S', T')$.
 - Consider a set S' , constructed so that no augmenting path exists.
 - All edges from S' to T' are full saturated (because no path exists).
 - All edges from T' to S' have zero flow (because no path exists).
 - Therefore, $|f| = f(S', T') = c(S', T')$.
- (c) \implies (a): if a cut (S', T') exists such that $f(S', T') = c(S', T')$, then f is a maximum flow.
 - $f(S', T') = c(S', T') = |f|$ (from the statement).
 - There is no flow greater than $c(S', T')$, so f is a maximum flow.

5.4 Ford-Fulkerson Algorithm

The Ford-Fulkerson algorithm **follows the general pattern** above - start with a zero flow and a residual network, find an augmenting path in the residual network, apply the flow from that path, then repeat until no more augmenting paths can be found.

Correctness: the algorithm terminates when there is no augmenting path in the residual network (second cut theorem, condition b), which means that the flow found must be a maximum (second cut theorem, condition a).

The **running time** is $O(m)$ for one iteration: constructing G_f takes $\Theta(m)$ (or (n) incrementally), searching for a path takes $O(m)$, and updating the flow takes $O(n)$.

The **number of iterations** depends on the selection strategy used to find augmenting paths. Assuming all edge capacities are integral:

- Capacities never become non-integral, so the current total flow in G increases by some integer (i.e. by at least 1).
- Therefore, if f^* denotes a maximum flow, the number of iterations is $O(|f^*|)$.
- The **running time of the whole algorithm** is therefore $O(|f^*| * |m|)$.

5.5 Edmonds-Karp Algorithm

This algorithm is the Ford-Fulkerson algorithm with the following **selection strategy**: in each iteration, select the **shortest augmenting path** (counting by number of edges, not capacities).

The **running time** can be explored as follows:

- Again, the running time is $O(m)$ for one iteration.
- The number of iteration is $O(nm)$:
 - Let q be the number of iterations, and let k_1, k_2, \dots, k_q be the augmenting path lengths selected in iterations $1, 2, \dots, q$.
 - We know that $1 \leq k_1 \leq k_2 \leq \dots \leq k_q \leq n - 1$.
 - A path of the same length appears at most m times in $\langle k_1, k_2, \dots, k_q \rangle$.
 - Therefore, $q \leq nm$.
- The overall running time is therefore $O(nm^2)$, **independent** of the maximum flow value.

6 Flow Feasibility Problems

This variation adds **supply and demand** at various vertices, such that:

- $G = (V, E, c, d)$ is a **flow network** G with vertices V and edges E .
- $c(u, v)$ is the **capacity of the edge** $(u, v) \in E$.
- $d(v)$ is the initial supply/demand at the vertex v .
 - $d(v) > 0$ indicates a supply of $d(v)$ units at v .
 - $d(v) < 0$ indicates a demand for $d(v)$ units at v .
 - $d(v) = 0$ indicates a 'transitional' vertex.
- We assume that supply matches demand, i.e. $\sum_{v \in V} d(v) = 0$.

The objective is to find a flow f that 'moves' all supply to meet all demand. This means that each vertex should have a net flow of zero, i.e. for each vertex v ,

$$\sum_{(v,x) \in E} f(v,x) - \sum_{(z,v) \in E} f(z,v) = d(v)$$

6.1 Reduction to Maximum Flow Problem

For a given $G = (V, E, c, d)$ for a flow feasibility problem, $G' = (V', E', c', s, t)$ can be constructed to solve it as a maximum flow problem:

- Create an **artificial source vertex** s with edges **to** all 'supply' vertices v , each with capacity equal to the supply, $d(v)$.
- Create an **artificial sink vertex** t with edges **from** all 'demand' vertices u , each with capacity equal to the negative of the demand, $-d(u)$.

Formally:

- $V' = V + \{s, t\}$
- $E' = E + \{(s, v) : v \in V, d(v) > 0\} + \{(u, t) : u \in V, d(u) < 0\}$
- $c'(u, v) = c(u, v)$
- $c'(s, v) = d(v)$
- $c'(u, t) = -d(u)$

A maximum flow in G' saturates all outgoing edges from s (and therefore saturates all incoming edges to t) if and only if there is a feasible flow in G .

- If a maximum flow f' in G' saturates all edges outgoing from s , then remove the artificial vertices s and t to get a feasible flow for G .
- If f is a feasible flow in G , then saturate all edges outgoing from s and all edges incoming to t to get a maximum flow in G' .

7 Minimum Cost Flow Problems

Minimum cost flow problems add **costs** to each edge and attempt to find a flow that **satisfies supply/demand** or **maximises flow** (depending on the question type) whilst **minimising cost**.

7.1 Notation

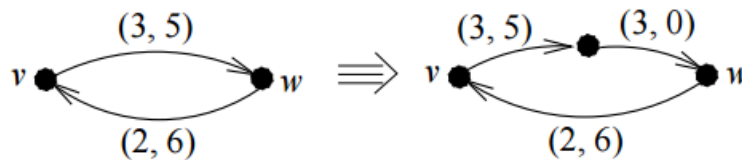
- $G = (V, E, u, c, d)$ is a **flow network** G with vertices V and edges E .
 - $|V| = n, |E| = m$.
- $u(v, w)$ is the **capacity of the edge** $(v, w) \in E$.
- $c(v, w)$ is the **cost of the edge** $(v, w) \in E$ per unit of flow.
- $d(v)$ is the initial supply/demand at the vertex v .
 - $d(v) > 0$ indicates a supply of $d(v)$ units at v .
 - $d(v) < 0$ indicates a demand for $d(v)$ units at v .
 - $d(v) = 0$ indicates a 'transitional' vertex.
- Edges are typically drawn with labels in the form $(\text{capacity}, \text{cost})$.
- We assume that supply matches demand, i.e. $\sum_{v \in V} d(v) = 0$.

The **cost of a flow** f is the sum of the costs of all used edges:

$$\text{cost of } f = \sum_{(v,w) \in E} c(v,w)f(v,w)$$

Usual capacity constraints, flow conservation constraints and net flow definitions apply.

For convenience, the following **assumption** can be made: if $(w, v) \in E$ then $(v, w) \in E$, but at least one of $u(w, v)$ and $u(v, w)$ is 0. Where this does not hold, one edge can be split as followed without changing the result:

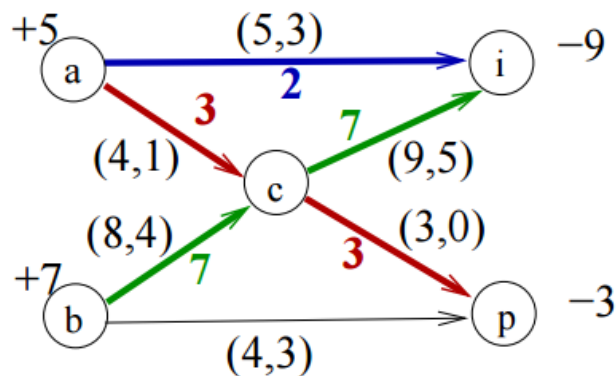


7.2 Sending Flow Along Cheapest Paths

A naïve approach is to find the cheapest paths between supply and demand. Formally:

- Take a node v with remaining supply.
- Find a cheapest path to a node w with remaining demand.
- Send as much flow as possible from v to w along this path.

For example:

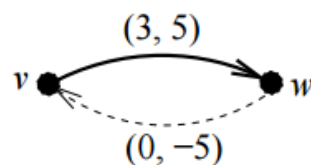


The paths identified are $\langle a, c, p \rangle$, $\langle a, i \rangle$ and $\langle b, c, i \rangle$

The cost of this flow is $(2 * 3) + (3 * 1) + (3 * 0) + (7 * 5) + (7 * 4) = 72$. This is **not optimal** - sometimes we need to **re-route** flow to get the optimal cost.

7.3 Residual Network

- Edges exist in both directions between vertices, but $u(v, w) = 0$ and/or $u(w, v) = 0$.
- If $u(v, w) > 0$ then $c(v, w) \geq 0$ and $u(w, v) = 0$.
- $c(w, v) = -c(v, w)$



Let f be a flow in a network $G = (V, E, u, c, d)$.

The **residual network** of G with respect to f is $G_f = (V, E_f, u_f, c, d_f)$, where:

- **Capacity:** for each edge (v, w) where $u(v, w) > 0$ and $u(w, v) = 0$:
 - $u_f(v, w) = u(v, w) - f(v, w)$ (i.e. the remaining flow)
 - $u_f(w, v) = f(v, w)$ (i.e. the flow that can be reversed)

- **Supply/Demand:** for each vertex $v \in V$:
 - $d_f(v) = d(v) - \sum_{(v,x) \in E} f(v,w) + \sum_{(z,v) \in E} f(z,v)$
 - (i.e. the original supply/demand, minus flow going out, plus flow going in)
- **Edges:** E_f is the set of edges with positive, non-zero residual capacities.

7.4 Successive Shortest Path Algorithm

The correct, optimal solution to this style of problem is to send flow along the **cheapest paths in the residual network** that is created after each action, rather than using just the original network (as in the earlier approach).

```

1  fun SUCCESSIVE_SP(G)
2      f = zero flow
3
4      // initial residual network and nodes with supply
5      G_f = residual network of G
6      S = {v in V, where d_f(v) > 0}
7
8      while (S is not empty):
9          G_f = residual network of G
10
11         // find a vertex with supply
12         v = select any vertex from S
13
14         // find a vertex with demand
15         T = shortest-paths tree in G_f from v
16         if (T contains a vertex w with d_f(w) < 0):
17
18             P = path in T from v to w
19
20             // find the path's flow bottleneck
21             q = MIN( d_f(v), -d_f(w), min edge-capacity in P )
22
23             // compute new flow and combine it with existing flow
24             f_P = q units of flow along path P
25             f = f + f_P
26
27             // remove v from S if all residual supply was used
28             if (d_f(v) = 0):
29                 S.remove(v)
30
31         else:
32             // remove v from S if no supply can be sent
33             S.remove(v)

```

7.4.1 Correctness and Runtime

- At the end of each iteration, f has the minimum cost of all flows which satisfy the same supply/demand quantities satisfied by f . This can be proved by induction, but such a proof is not covered here.
- The cost of residual edges can be negative, but there can be no negative cycles.
- **Dijkstra's algorithm** can be used to compute the shortest paths, even with negative residual edges: appropriate $h(v)$ values can be maintained in each iteration to allow the re-weighting system from Johnson's algorithm to create non-negative edges.
- The running time of **each iteration** is bounded by Dijkstra's algorithm, $O(\min\{m \cdot \log_2(n), n^2\})$.
- If all input edge capacities, supplies and demands are integral, the **number of iterations** is at most $\sum\{d(v) : v \in V, d(v) > 0\}$ (i.e. the sum of supply).
- The number of iterations **can be exponential** in the number of nodes, but more complex algorithms exist which run in **polynomial** time. The best known is $O(m^2 \cdot \log_2(n)^2)$.

8 Multi-Commodity Flow Problems

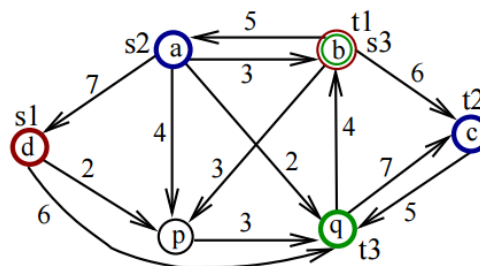
This variation adds multiple commodities to the flow network. The objective is to find a **simultaneous flow of all commodities** that satisfies some specified **global objective**.

8.1 Notation

- $G = (V, E, u)$ is a **flow network** G with vertices V and edges E .
 - $|V| = n, |E| = m$.
 - $u(v, w)$ is the **capacity of the edge** $(v, w) \in E$.
- $C = \langle (s_1, t_1, d_1), \dots, (s_k, t_k, d_k) \rangle$ specifies k commodities where, for $q = 1, 2, \dots, k$:
 - $s_q \in V$ and $t_q \in V$ are the source and target vertices of commodity q .
 - $d_q (> 0)$ is the demand of this commodity.

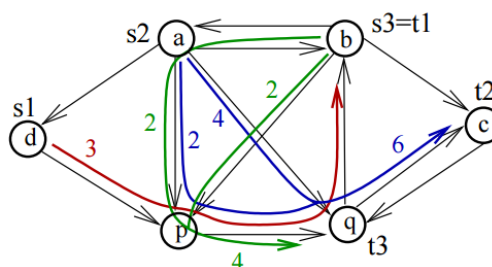
For some multi-commodity flow problems, the flows selected do not need to satisfy capacity constraints.

8.2 Example



Commodity (k)	Source (s_k)	Target (t_k)	Demand (d_k)
1	d	b	3
2	a	c	6
3	b	q	4

The following is one flow that satisfies these demands. Note that it does not satisfy the capacity constraints given on the previous diagram.



8.3 Flows of Commodities

- A simultaneous flow of the specified commodities consists of flows of individual commodities $\langle f_1, \dots, f_k \rangle$.
- A flow f_q of commodity q is a function $f_q : E \mapsto \mathbb{R}$ where $f_q(x, y)$ is the flow of commodity q along the edge $(x, y) \in E$.
 - f_q must satisfy **flow conservation constraints**.
 - **Capacity constraints** may not be present.
- Flow conservation for commodity q :
 - The net flow of q outgoing from the source s_q is d_q .
 - The net flow of q incoming to the target t_q is d_q .
 - The net flow outgoing from all other vertices $v \in V - \{s_q, t_q\}$ is 0.

8.4 Types of Multi-Commodity Flow Problems

8.4.1 Feasibility

Total flow is defined as $f(v, w) = \sum_{q=1}^k f_q(v, w)$.

The objective is to design a simultaneous flow $f = \langle f_1, \dots, f_k \rangle$ of all commodities that satisfies **capacity constraints**.

Formally, for each edge $(v, w) \in E$, the total flow $f(v, w) \leq u(v, w)$.

8.4.2 Minimum Cost

Edge costs are introduced to the flow network, such that $c(v, w)$ is the cost of sending one unit of flow along the edge $(v, w) \in E$. The total cost of a flow $f = \langle f_1, \dots, f_k \rangle$ is defined as the sum of the cost of all flows along all edges.

Total cost: $\sum_{(v,w) \in E} c(v, w) \cdot [f_1(v, w) + \dots + f_k(v, w)]$.

The objective is to design a simultaneous flow that satisfies **capacity constraints** and **minimises total cost**.

8.4.3 Minimum Congestion

The congestion of a flow on an edge $(v, w) \in E$ is defined as $f(v, w)/u(v, w)$ (i.e. total flow divided by capacity).

The objective is to design a simultaneous flow of all commodities that **minimises the maximum congestion**.

Note: the optimal (minimum) congestion may be greater than 100%.

8.5 Computational Approaches

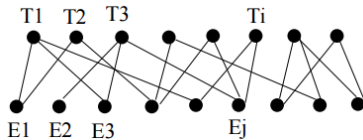
- **Specialised algorithms:** start with some initial simultaneous flow, and iteratively keep improving it by re-routing some of the flow of each commodity via better paths.
- **Mathematical programming:** express the problem as a linear programming problem (or other suitable model) and use appropriate general solving algorithms.

9 Maximum Bipartite Matching

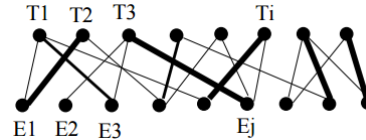
A **bipartite graph** is a graph with nodes that can be divided into two **disjoint sets**, such that all edges have **one end in each set**.

A **bipartite matching** is a subset of edges M such that every node belongs to at most one edge. A **maximum bipartite matching** maximises $|M|$.

Bipartite graph G :



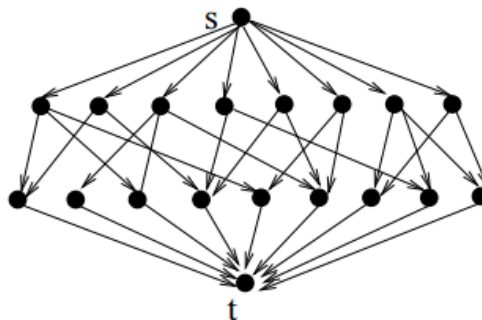
Bipartite matching M :



This approach is useful when considering the **allocation of resources**.

9.1 Computing a Maximum Matching

- Label the two disjoint sets of nodes as A and B .
- Make the edges between A and B directed, from A to B .
- Create pseudo-nodes s and t .
- Create edges from s to all nodes in A .
- Create edges from all nodes in B to t .
- Set the capacity of all edges to 1.
- Find the maximum flow from s to t .
 - Edges selected for the flow are part of the matching.
 - The size of the maximum flow is the size of the matching.
 - [See more: Max-Flow Problems, page 21.](#)



The runtime of such an approach is $O(nm)$, because the maximum value of the flow is $n/2$ and cost per iteration is $O(m)$.

10 Linear Programming

TODO: Lecture 5.

11 Optimisations for NP-Hard Problems

TODO: Lectures 6, 7, 8.