Lecture 20: Max-Flow II

Michael Dinitz

November 7, 2024
601.433/633 Introduction to Algorithms

Introduction

Last time:

- Max-Flow = Min-Cut
- Can compute max flow and min cut using Ford-Fulkerson: while residual graph has an $s \to t$ path, push flow along it.
  - Corollary: if all capacities integers, max-flow is integral
  - If max-flow has value F, time $O(F(m+n))$ (if all capacities integers)
  - Exponential time!

Today:

- Important setting where FF is enough: max bipartite matching
- Two ways of making FF faster: Edmonds-Karp

Max Bipartite Matching

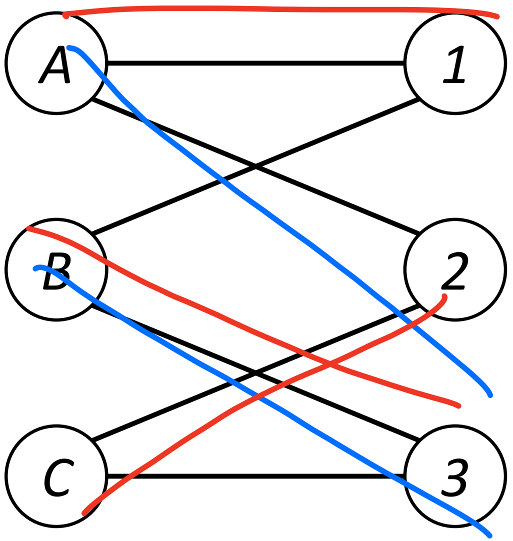
Setup

Definition

A graph $G = (V, E)$ is *bipartite* if V can be partitioned into two parts L, R such that every edge in E has one endpoint in L and one endpoint in R.

Definition

A *matching* is a subset $M \subseteq E$ such that $e \cap e' = \emptyset$ for all $e, e' \in M$ with $e \neq e'$ (no two edges share an endpoint)



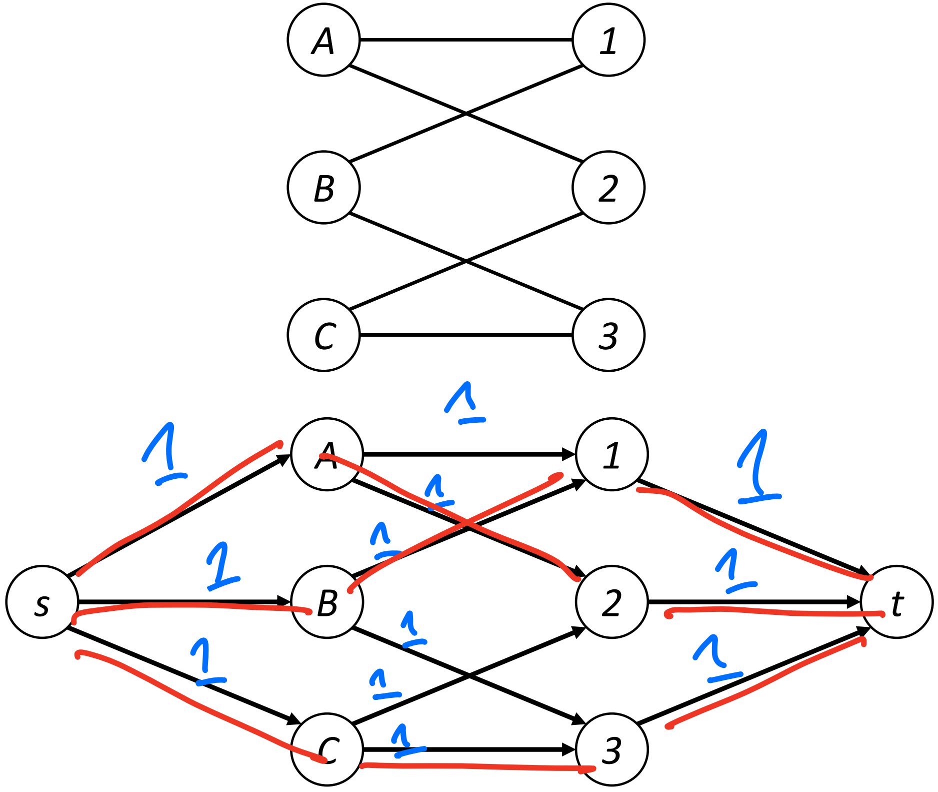
Bipartite Maximum Matching: Given bipartite graph $G = (V, E)$, find matching M maximizing $|M|$

- Extremely important problem, doesn't seem to have much to do with flow!

Algorithm

Give all edges capacity $\mathbf{1}$
Direct all edges from $\boldsymbol{L}$ to $\boldsymbol{R}$
Add source $\boldsymbol{s}$ and sink $\boldsymbol{t}$
Add edges of capacity $\mathbf{1}$ from $\boldsymbol{s}$ to $\boldsymbol{L}$
Add edges of capacity $\mathbf{1}$ from $\boldsymbol{R}$ to $\boldsymbol{t}$

Run FF to get flow $\boldsymbol{f}$
Return $\boldsymbol{M} = \{\boldsymbol{e} \in \boldsymbol{L} \times \boldsymbol{R} : \boldsymbol{f}(\boldsymbol{e}) > \mathbf{0}\}$



Correctness

Claim: M is a matching

Correctness

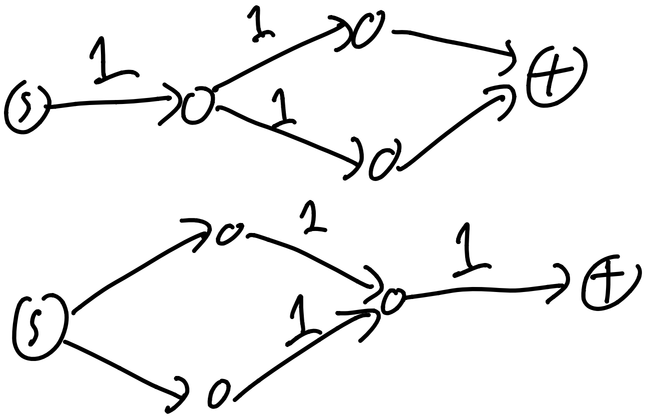
Claim: M is a matching

Proof: capacities in $\{0, 1\} \implies f(e) \in \{0, 1\}$ for all e (integrality)

Correctness

Claim: M is a matching

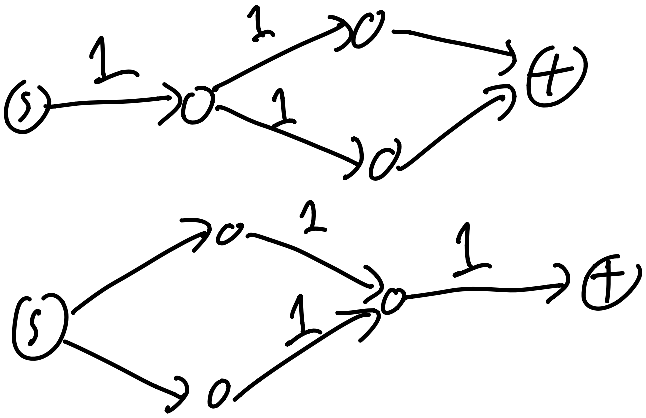
Proof: capacities in $\{0, 1\} \implies f(e) \in \{0, 1\}$ for all e (integrality)



Correctness

Claim: M is a matching

Proof: capacities in $\{0, 1\} \implies f(e) \in \{0, 1\}$ for all e (integrality)

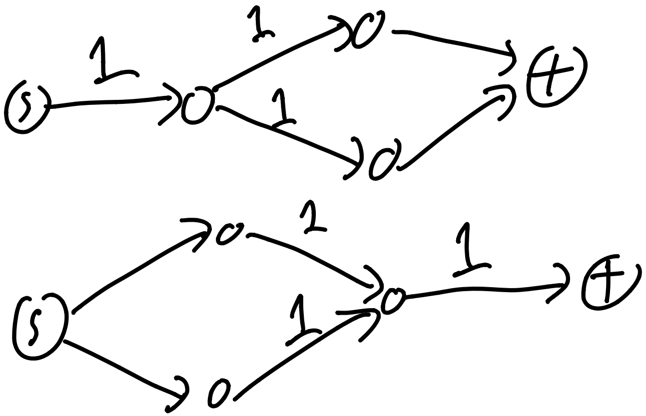


Claim: M is maximum matching

Correctness

Claim: M is a matching

Proof: capacities in $\{0,1\} \implies f(e) \in \{0,1\}$ for all e (integrality)



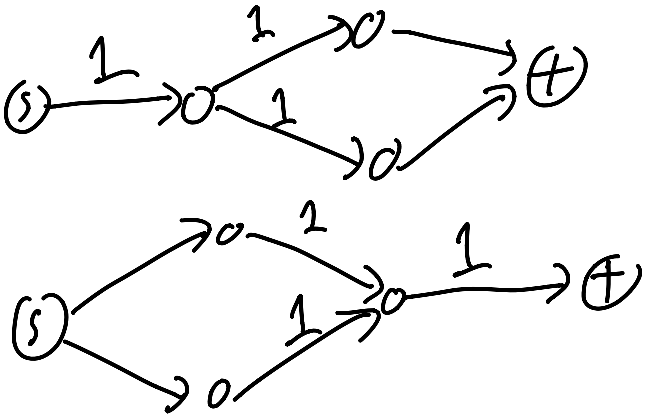
Claim: M is maximum matching

Proof: Suppose larger matching M'

Correctness

Claim: M is a matching

Proof: capacities in $\{0, 1\} \implies f(e) \in \{0, 1\}$ for all e (integrality)



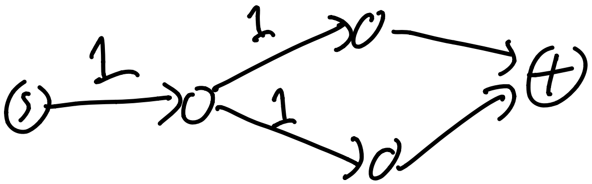
Claim: M is maximum matching

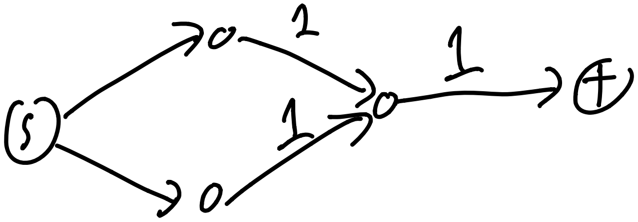
Proof: Suppose larger matching M'
Can send $|M'|$ flow using M'!

Correctness

Claim: M is a matching

Proof: capacities in $\{0,1\} \implies f(e) \in \{0,1\}$ for all e (integrality)





Claim: M is maximum matching

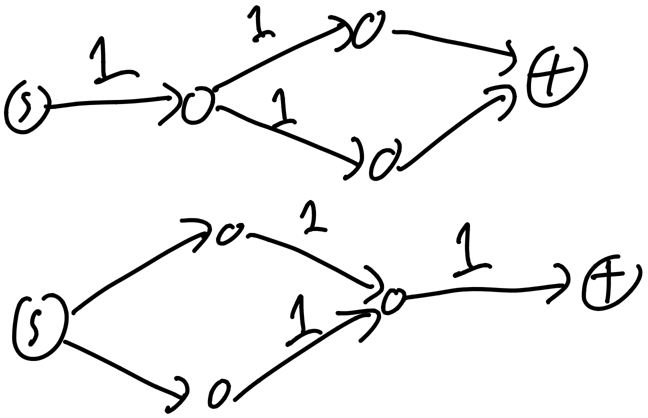
Proof: Suppose larger matching M'
Can send $|M'|$ flow using M'!

- $f'(s,u) = 1$ is u matched in M', otherwise 0
- $f'(v,t) = 1$ if v matched in M', otherwise 0
- $f'(u,v) = 1$ if $\{u,v\} \in M'$, otherwise 0

Correctness

Claim: M is a matching

Proof: capacities in $\{0,1\} \implies f(e) \in \{0,1\}$ for all e (integrality)



Claim: M is maximum matching

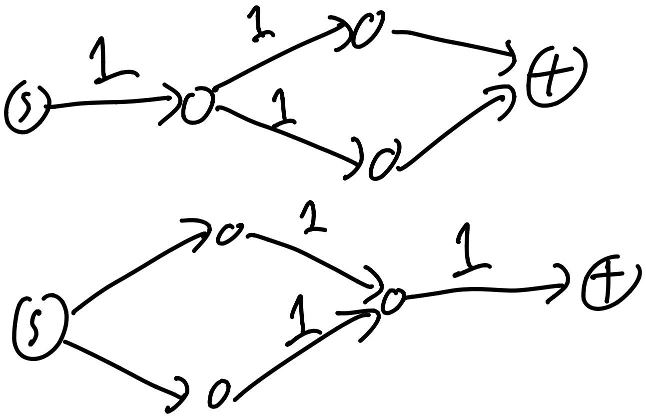
Proof: Suppose larger matching M'
Can send $|M'|$ flow using M'!

- $f'(s,u) = 1$ is u matched in M', otherwise 0
- $f'(v,t) = 1$ if v matched in M', otherwise 0
- $f'(u,v) = 1$ if $\{u,v\} \in M'$, otherwise 0
- $|f'| = |M'| > |M| = |f|$

Correctness

Claim: $\boldsymbol{M}$ is a matching

Proof: capacities in $\{0,1\} \implies \boldsymbol{f}(\boldsymbol{e}) \in \{0,1\}$ for all $\boldsymbol{e}$ (integrality)



Claim: $\boldsymbol{M}$ is maximum matching

Proof: Suppose larger matching $\boldsymbol{M}'$
Can send $|\boldsymbol{M}'|$ flow using $\boldsymbol{M}'$!

- $\boldsymbol{f}'(\boldsymbol{s},\boldsymbol{u}) = \mathbf{1}$ is $\boldsymbol{u}$ matched in $\boldsymbol{M}'$, otherwise $\mathbf{0}$
- $\boldsymbol{f}'(\boldsymbol{v},\boldsymbol{t}) = \mathbf{1}$ if $\boldsymbol{v}$ matched in $\boldsymbol{M}'$, otherwise $\mathbf{0}$
- $\boldsymbol{f}'(\boldsymbol{u},\boldsymbol{v}) = \mathbf{1}$ if $\{\boldsymbol{u},\boldsymbol{v}\} \in \boldsymbol{M}'$, otherwise $\mathbf{0}$
- $|\boldsymbol{f}'| = |\boldsymbol{M}'| > |\boldsymbol{M}| = |\boldsymbol{f}|$
- Contradiction

Running Time

Running Time:

- $O(n+m)$ to make new graph
- $|f| = |M| \le n/2$ iterations of FF

$\Longrightarrow$ $O(n(m+n)) = O(mn)$ time (assuming $m \ge \Omega(n)$)

Exensions

Many extensions:

- Max-weight bipartite matching
- Min-cost perfect matching
- Matchings in general graphs
- ...

Exensions

Many extensions:

- Max-weight bipartite matching
- Min-cost perfect matching
- Matchings in general graphs
- ...

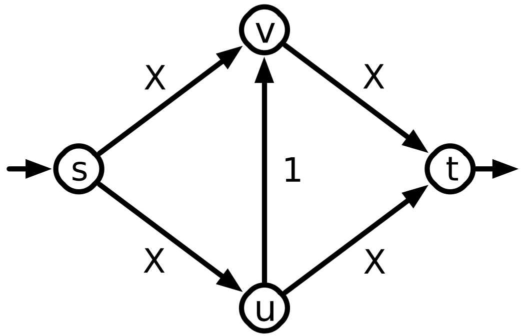
Still active area of study!

- Michael Dinitz, Sungjin Im, Thomas Lavastida, Benjamin Moseley, Sergei Vassilvitskii. *Faster Matchings via Learned Duals*. NeurIPS 2021.
- Michael Dinitz, George Li, Quanquan Liu, Felix Zhou. *Differentially Private Matchings*. Submitted (Monday), hopefully on arXiv soon.

Edmonds-Karp

Intuition

Bad example for Ford-Fulkerson:

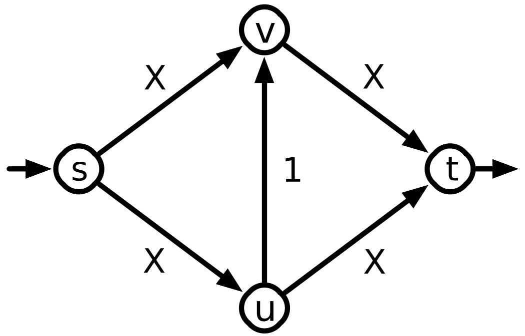


A bad example for the Ford-Fulkerson algorithm.

If Ford-Fulkerson chooses bad augmenting paths, super slow!

Intuition

Bad example for Ford-Fulkerson:



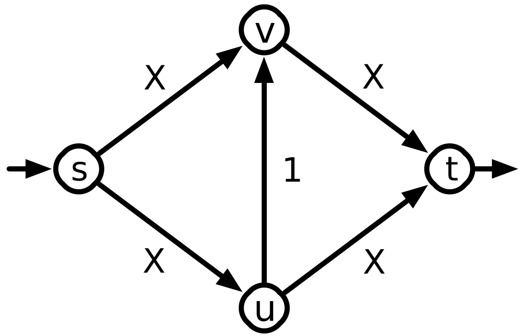
A bad example for the Ford-Fulkerson algorithm.

If Ford-Fulkerson chooses bad augmenting paths, super slow!

Obvious idea: Choose better paths!

Intuition

Bad example for Ford-Fulkerson:

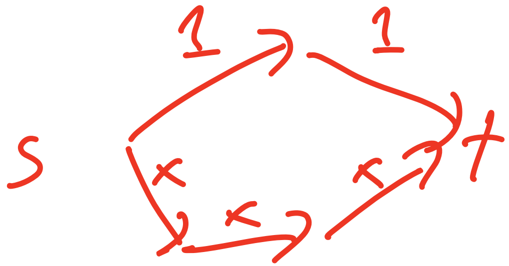


A bad example for the Ford-Fulkerson algorithm.

If Ford-Fulkerson chooses bad augmenting paths, super slow!

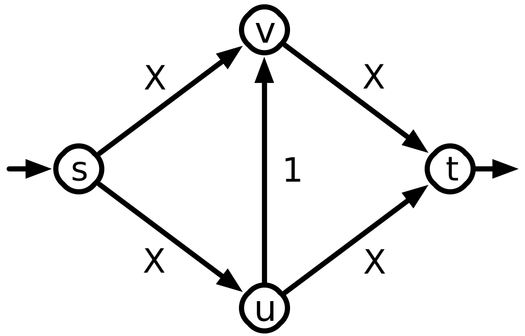
Obvious idea: Choose better paths!

Obvious path to pick:



Intuition

Bad example for Ford-Fulkerson:



A bad example for the Ford-Fulkerson algorithm.

If Ford-Fulkerson chooses bad augmenting paths, super slow!

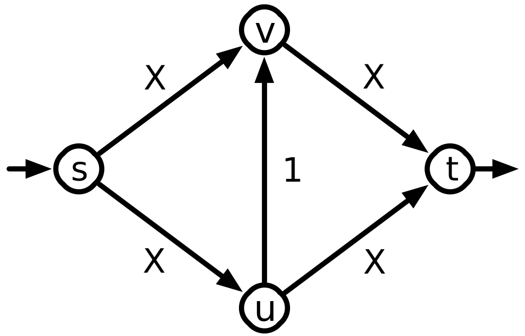
Obvious idea: Choose better paths!

Obvious path to pick:

$$\underset{\text{augmenting paths } P}{\arg\max} \min_{e \in P} c_f(e).$$

Intuition

Bad example for Ford-Fulkerson:



A bad example for the Ford-Fulkerson algorithm.

If Ford-Fulkerson chooses bad augmenting paths, super slow!

Obvious idea: Choose better paths!

Obvious path to pick:

$$\arg\max_{\text{augmenting paths } P} \min_{e \in P} c_f(e).$$

- "Widest" path: push as much flow as possible each iteration

Edmonds-Karp #1

Edmonds-Karp #1: Ford-Fulkerson, always choose "widest" path.

- Correct, since FF. Running time?

Edmonds-Karp #1

Edmonds-Karp #1: Ford-Fulkerson, always choose "widest" path.

- Correct, since FF. Running time?

Lemma

In any graph with max $s-t$ flow F, there exists a path from s to t with capacity at least F/m

Edmonds-Karp #1

Edmonds-Karp #1: Ford-Fulkerson, always choose "widest" path.

- Correct, since FF. Running time?

Lemma

In any graph with max $s-t$ flow F, there exists a path from s to t with capacity at least F/m

Proof.

Let $X = \{e \in E : c(e) < F/m\}$.

Edmonds-Karp #1

Edmonds-Karp #1: Ford-Fulkerson, always choose "widest" path.

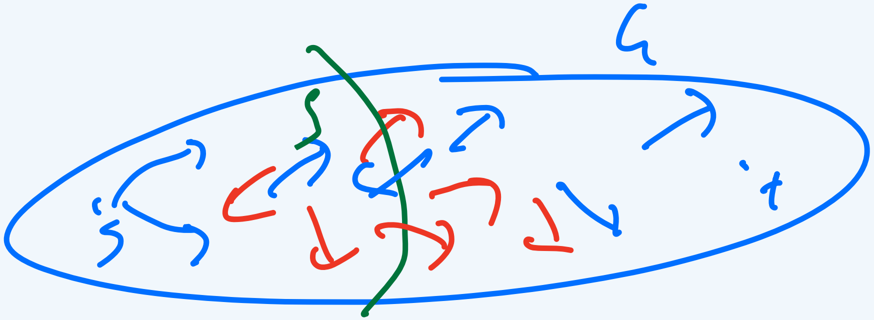
- Correct, since FF. Running time?

Lemma

In any graph with max $s-t$ flow F, there exists a path from s to t with capacity at least F/m

Proof.

Let $X = \{e \in E : c(e) < F/m\}$.
If no $s \to t$ path in $G \setminus X$, then X an (edge) cut. Let S = nodes reachable from s in $G \setminus X$.



Edmonds-Karp #1

Edmonds-Karp #1: Ford-Fulkerson, always choose "widest" path.

- Correct, since FF. Running time?

Lemma

In any graph with max $\boldsymbol{s-t}$ flow $\boldsymbol{F}$, there exists a path from $\boldsymbol{s}$ to $\boldsymbol{t}$ with capacity at least $\boldsymbol{F/m}$

Proof.

Let $X = \{e \in E : c(e) < F/m\}$.
If no $s \to t$ path in $G \setminus X$, then X an (edge) cut. Let S = nodes reachable from s in $G \setminus X$.

$$cap(S, \bar{S}) \le cap(X) = \sum_{e \in X} c(e) < m \cdot (F/m) = F$$

Edmonds-Karp #1

Edmonds-Karp #1: Ford-Fulkerson, always choose "widest" path.

- Correct, since FF. Running time?

Lemma

In any graph with max $s - t$ flow F, there exists a path from s to t with capacity at least F/m

Proof.

Let $X = \{e \in E : c(e) < F/m\}$.
If no $s \to t$ path in $G \setminus X$, then X an (edge) cut. Let S = nodes reachable from s in $G \setminus X$.

$$cap(S, \bar{S}) \leq cap(X) = \sum_{e \in X} c(e) < m \cdot (F/m) = F$$

$\implies$ min (s, t) cut $\leq cap(S, \bar{S}) < F$.

Edmonds-Karp #1

Edmonds-Karp #1: Ford-Fulkerson, always choose "widest" path.

- Correct, since FF. Running time?

Lemma

In any graph with max $s-t$ flow F, there exists a path from s to t with capacity at least F/m

Proof.

Let $X = \{e \in E : c(e) < F/m\}$.
If no $s \to t$ path in $G \setminus X$, then X an (edge) cut. Let S = nodes reachable from s in $G \setminus X$.

$$cap(S, \bar{S}) \le cap(X) = \sum_{e \in X} c(e) < m \cdot (F/m) = F$$

$\implies$ min (s,t) cut $\le cap(S, \bar{S}) < F$. Contradiction.

Edmonds-Karp #1

Edmonds-Karp #1: Ford-Fulkerson, always choose "widest" path.

- Correct, since FF. Running time?

Lemma

In any graph with max $s - t$ flow F, there exists a path from s to t with capacity at least F/m

Proof.

Let $X = \{e \in E : c(e) < F/m\}$.
If no $s \to t$ path in $G \setminus X$, then X an (edge) cut. Let S = nodes reachable from s in $G \setminus X$.

$$cap(S, \bar{S}) \le cap(X) = \sum_{e \in X} c(e) < m \cdot (F/m) = F$$

$\implies$ min (s,t) cut $\le cap(S, \bar{S}) < F$. Contradiction.
$\implies$ $\exists s \to t$ path P in $G \setminus X$: every edge of P has capacity at least F/m $\square$

Edmonds-Karp #1

Edmonds-Karp #1: Ford-Fulkerson, always choose "widest" path.

- Correct, since FF. Running time?

Lemma

In any graph with max $s-t$ flow F, there exists a path from s to t with capacity at least F/m

Proof.

Let $X = \{e \in E : c(e) < F/m\}$.
If no $s \to t$ path in $G \setminus X$, then X an (edge) cut. Let S = nodes reachable from s in $G \setminus X$.

$$cap(S, \bar{S}) \le cap(X) = \sum_{e \in X} c(e) < m \cdot (F/m) = F$$

$\implies$ min (s,t) cut $\le cap(S, \bar{S}) < F$. Contradiction.
$\implies$ $\exists s \to t$ path P in $G \setminus X$: every edge of P has capacity at least F/m □

Does this implies at most m iterations?

EK 1 Running Time

Theorem

If F is the value of the maximum flow and all capacities are integers, $\#$ iterations of EK1 is at most $O(m \log F)$

EK 1 Running Time

Theorem

If F is the value of the maximum flow and all capacities are integers, # iterations of EK1 is at most $O(m \log F)$

How much flow remains to be be sent after iteration i?

EK 1 Running Time

Theorem

If F is the value of the maximum flow and all capacities are integers, # iterations of EK1 is at most $O(m \log F)$

How much flow remains to be be sent after iteration i?

- $i = 0$:

EK 1 Running Time

Theorem

If F is the value of the maximum flow and all capacities are integers, # iterations of EK1 is at most $O(m \log F)$

How much flow remains to be be sent after iteration i?

- $i = 0$: F

EK 1 Running Time

Theorem

If F is the value of the maximum flow and all capacities are integers, # iterations of EK1 is at most $O(m \log F)$

How much flow remains to be be sent after iteration i?

- $i = 0$: F
- $I = 1$: Sent at least F/m, so at most $F - F/m = F(1 - 1/m)$ remaining

EK 1 Running Time

Theorem

If F is the value of the maximum flow and all capacities are integers, # iterations of EK1 is at most $O(m \log F)$

How much flow remains to be be sent after iteration i?

- $i = 0$: F
- $I = 1$: Sent at least F/m, so at most $F - F/m = F(1 - 1/m)$ remaining
- $i = 2$: Sent at least R/m if R was remaining after iteration 1, so at most $R - R/m = R(1 - 1/m) \leq F(1 - 1/m)^2$ remaining

EK 1 Running Time

Theorem

If F is the value of the maximum flow and all capacities are integers, # iterations of EK1 is at most $O(m \log F)$

How much flow remains to be be sent after iteration i?

- $i = 0$: F
- $I = 1$: Sent at least F/m, so at most $F - F/m = F(1 - 1/m)$ remaining
- $i = 2$: Sent at least R/m if R was remaining after iteration 1, so at most $R - R/m = R(1 - 1/m) \leq F(1 - 1/m)^2$ remaining

By induction: after iteration i, at most $F(1 - 1/m)^i$ flow remaining to be sent.

EK 1 Running Time

Theorem

If F is the value of the maximum flow and all capacities are integers, # iterations of EK1 is at most $O(m \log F)$

How much flow remains to be be sent after iteration i?

- $i = 0$: F
- $I = 1$: Sent at least F/m, so at most $F - F/m = F(1 - 1/m)$ remaining
- $i = 2$: Sent at least R/m if R was remaining after iteration 1, so at most $R - R/m = R(1 - 1/m) \leq F(1 - 1/m)^2$ remaining

By induction: after iteration i, at most $F(1 - 1/m)^i$ flow remaining to be sent.

Super useful inequality: $1 + x \leq e^x$ for all $x \in \mathbb{R}$

EK 1 Running Time

Theorem

If F is the value of the maximum flow and all capacities are integers, # iterations of EK1 is at most $O(m \log F)$

How much flow remains to be be sent after iteration i?

- $i = 0$: F
- $I = 1$: Sent at least F/m, so at most $F - F/m = F(1 - 1/m)$ remaining
- $i = 2$: Sent at least R/m if R was remaining after iteration 1, so at most $R - R/m = R(1 - 1/m) \leq F(1 - 1/m)^2$ remaining

By induction: after iteration i, at most $F(1 - 1/m)^i$ flow remaining to be sent.
Super useful inequality: $1 + x \leq e^x$ for all $x \in \mathbb{R}$
$\implies$ If $i > m \ln F$, amount remaining to be sent at most

$$F(1 - 1/m)^i < F(1 - 1/m)^{m \ln F} \leq F(e^{-1/m})^{m \ln F} = F \cdot e^{-\ln F} = 1$$

But all capacities integers, so must be finished!

Finishing EK1

Modified version of Dijkstra: find widest path in $\boldsymbol{O(m \log n)}$ time

- Total time $\boldsymbol{O(m \log n \cdot m \log F) = O(m^2 \log n \log F)}$
- Polynomial time!

Finishing EK1

Modified version of Dijkstra: find widest path in $O(m \log n)$ time

- Total time $O(m \log n \cdot m \log F) = O(m^2 \log n \log F)$
- Polynomial time!

Question: can we get running time independent of F?

- *Strongly* polynomial-time algorithm.

Edmonds-Karp #2

Again use Ford-Fulkerson, but pick *shortest* augmenting path (unweighted)

- Ignore capacities, just find augmenting path with fewest hops!
- Easy to compute with BFS in $O(m + n)$ time.

Edmonds-Karp #2

Again use Ford-Fulkerson, but pick *shortest* augmenting path (unweighted)

- Ignore capacities, just find augmenting path with fewest hops!
- Easy to compute with BFS in $\boldsymbol{O(m+n)}$ time.

Main question: how many iterations?

Edmonds-Karp #2

Again use Ford-Fulkerson, but pick *shortest* augmenting path (unweighted)

- Ignore capacities, just find augmenting path with fewest hops!
- Easy to compute with BFS in $O(m+n)$ time.

Main question: how many iterations?

Theorem

EK2 has at most $O(mn)$ iterations, so at most $O(m^2 n)$ running time (if $m \geq n$)

Proof (sketch) of EK2

Idea: prove that distance from $\boldsymbol{s}$ to $\boldsymbol{t}$ (unweighted) goes up by at least one every $\leq \boldsymbol{m}$ iterations.

Proof (sketch) of EK2

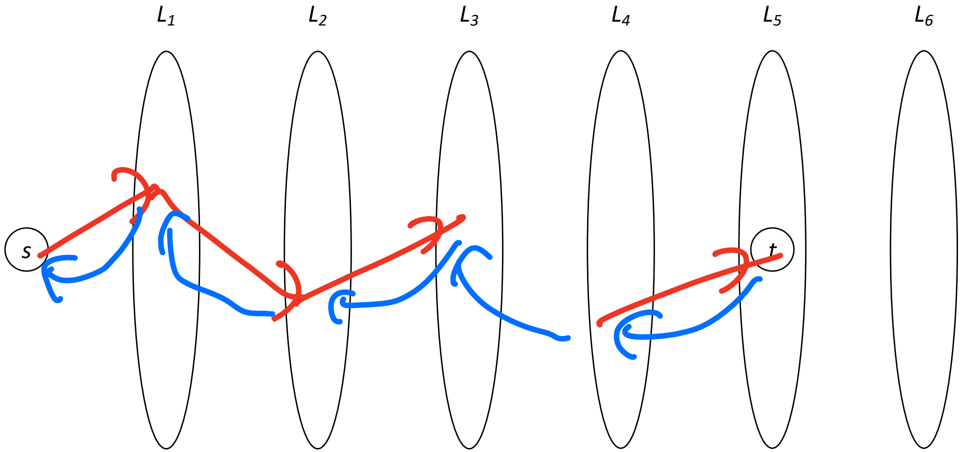
Idea: prove that distance from s to t (unweighted) goes up by at least one every $\leq m$ iterations.

- Distance initially $\geq 1 \implies$ distance $> n$ after at most mn iterations
- Only distance larger than n is ∞: no $s \to t$ path

$\implies$ Terminates after at most mn iterations.

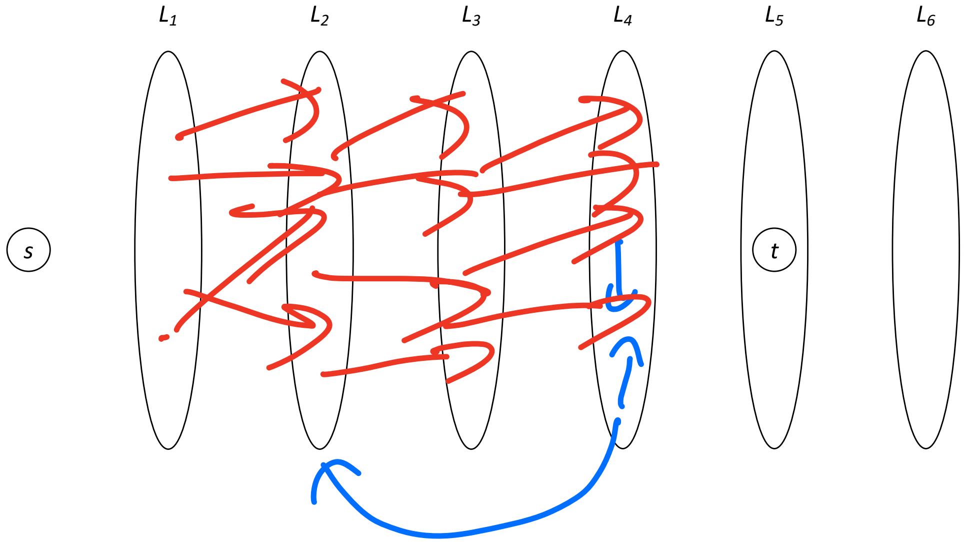
Proof (sketch) of EK2 (continued)

Suppose $s \to t$ distance is d.
"Lay out" residual graph in levels by BFS (distance from s)



Proof (sketch) of EK2 (continued)

Suppose $s \to t$ distance is d.
"Lay out" residual graph in levels by BFS (distance from s)

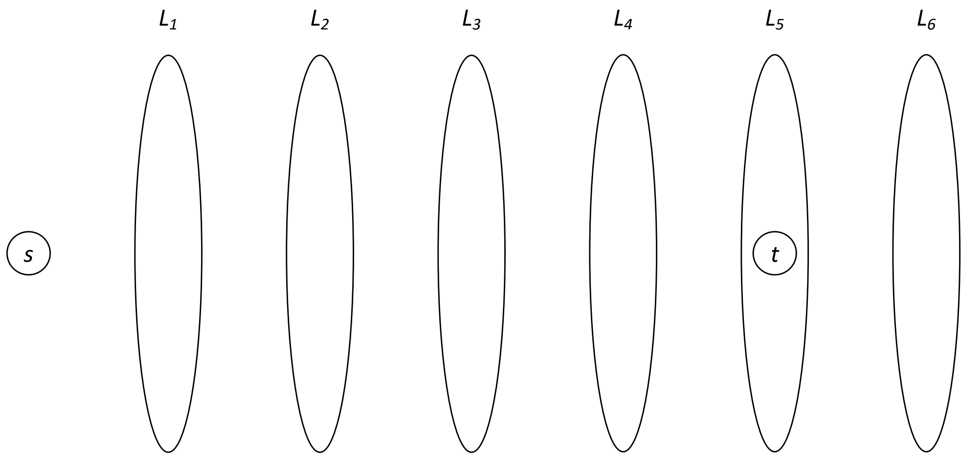


Edge types:

- Forward edges: **1** level
- Edges inside level
- Backwards edges

Proof (sketch) of EK2 (continued)

Suppose $s \to t$ distance is d.
"Lay out" residual graph in levels by BFS (distance from s)



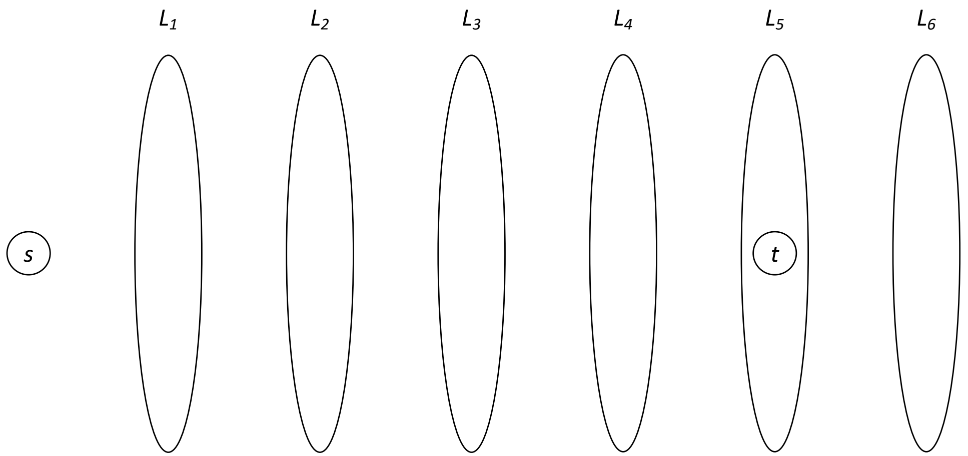
Edge types:

- Forward edges: **1** level
- Edges inside level
- Backwards edges

What happens when we choose a *shortest* augmenting path?

Proof (sketch) of EK2 (continued)

Suppose $\boldsymbol{s} \to \boldsymbol{t}$ distance is $\boldsymbol{d}$.
"Lay out" residual graph in levels by BFS (distance from $\boldsymbol{s}$)



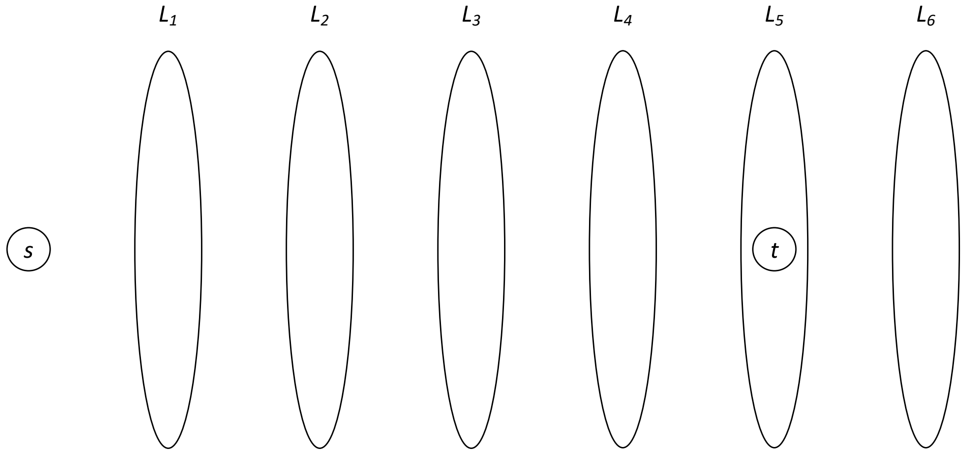
Edge types:

- Forward edges: **1** level
- Edges inside level
- Backwards edges

What happens when we choose a *shortest* augmenting path? Only uses forward edges!

Proof (sketch) of EK2 (continued)

Suppose $s \to t$ distance is d.
"Lay out" residual graph in levels by BFS (distance from s)



Edge types:

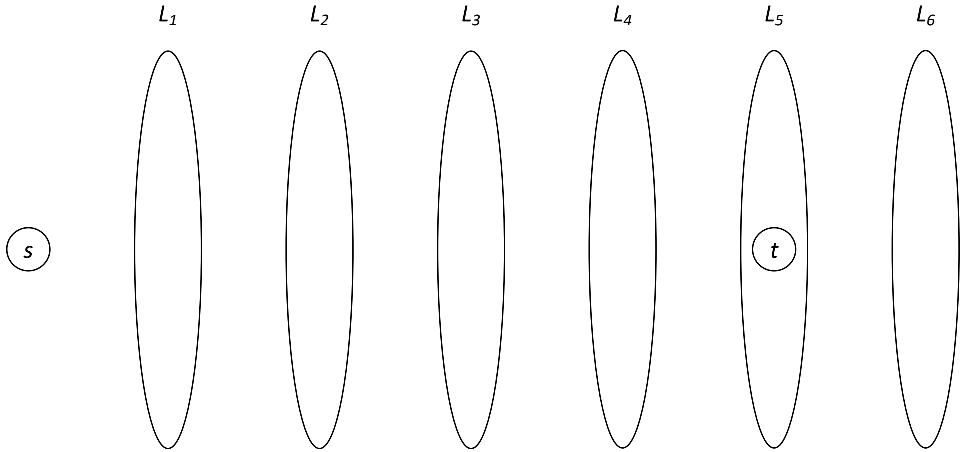
- Forward edges: **1** level
- Edges inside level
- Backwards edges

What happens when we choose a *shortest* augmenting path? Only uses forward edges!

- At least **1** forward edge gets removed, replaced with backwards edge.
- No backwards edges turned forward

Proof (sketch) of EK2 (continued)

Suppose $s \to t$ distance is d.
"Lay out" residual graph in levels by BFS (distance from s)



Edge types:

- Forward edges: 1 level
- Edges inside level
- Backwards edges

What happens when we choose a *shortest* augmenting path? Only uses forward edges!

- At least 1 forward edge gets removed, replaced with backwards edge.
- No backwards edges turned forward

So after m iterations (same layout): no path using only forward edges $\implies$ distance larger than d!

Finishing EK2

So at most mn iterations. Each iteration unweighted shortest path: BFS, time $O(m + n)$

Finishing EK2

So at most mn iterations. Each iteration unweighted shortest path: BFS, time $O(m+n)$

Total time: $O(mn(m+n)) = O(m^2 n)$. Independent of F!

Extensions

Many better algorithms for max-flow: *blocking flows* (Dinitz's algorithm (not me)), *push-relabel* algorithms, etc.

- CLRS has a few of these.
- State of the art:
  - Strongly polynomial: $\boldsymbol{O(mn)}$. Orlin [2013] & King, Rao, Tarjan [1994]
  - Weakly Polynomial: $\boldsymbol{O(m^{1+o(1)} \log U)}$ (where $\boldsymbol{U}$ is maximum capacity). Chen, Kyng, Liu, Peng, Gutenberg and Sachdeva [2022]

Many other variants of flows, some of which are just $\boldsymbol{s - t}$ max flow in disguise!

- Min-Cost Max-Flow: every edge also has a cost. Find minimum cost max-flow. Can be solved with just normal max flow!