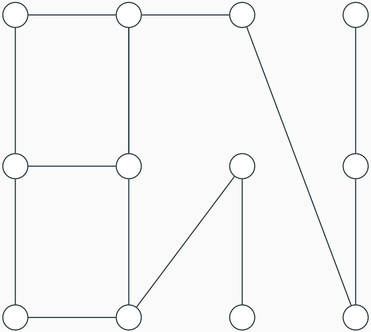
Hopcroft-Karp Algorithm for Bipartite Matching

(Matching Results Chris Likes)

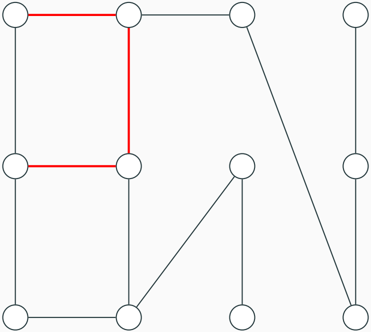
Maximum Matching

Given a graph $G(V, E)$, a matching is a set of edges $M \subseteq E$, such that no two edges of M share a common node.



Maximum Matching

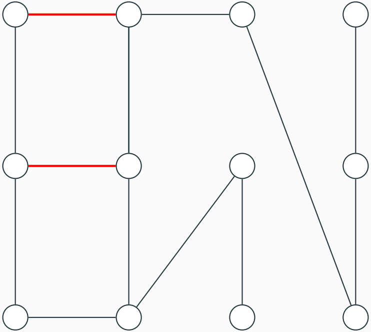
Given a graph $G(V, E)$, a matching is a set of edges $M \subseteq E$, such that no two edges of M share a common node.



not a matching

Maximum Matching

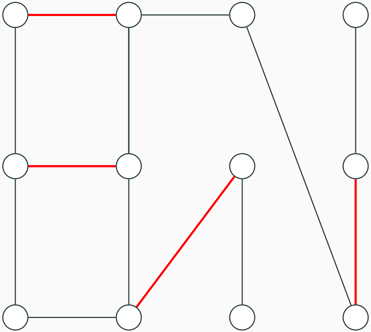
Given a graph $G(V, E)$, a matching is a set of edges $M \subseteq E$, such that no two edges of M share a common node.



a valid matching

Maximum Matching

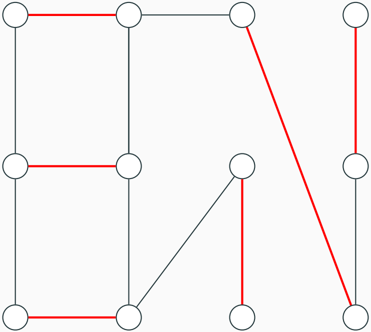
Given a graph $G(V, E)$, a matching is a set of edges $M \subseteq E$, such that no two edges of M share a common node.



a maximal matching

Maximum Matching

Given a graph $G(V, E)$, a matching is a set of edges $M \subseteq E$, such that no two edges of M share a common node.



a maximum matching

Relationship Between Maximum and Maximal

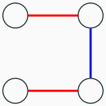
For any two maximal matchings A and B, we have

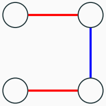
$|A| \leq 2|B|$ **and** $|B| \leq 2|A|$.

For any two maximal matchings A and B, we have

$$|A| \leq 2|B| \textbf{ and } |B| \leq 2|A|.$$

- Each edge in $B \setminus A$ is adjacent to at most two edges in $A \setminus B$.



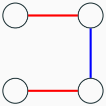


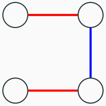
Relationship Between Maximum and Maximal

For any two maximal matchings A and B, we have

$$|A| \leq 2|B| \text{ and } |B| \leq 2|A|.$$

- Each edge in $B \setminus A$ is adjacent to at most two edges in $A \setminus B$.
- Each edge in $A \setminus B$ is adjacent to some edge in $B \setminus A$.



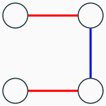


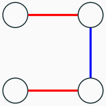
Relationship Between Maximum and Maximal

For any two maximal matchings A and B, we have

$|A| \leq 2|B|$ **and** $|B| \leq 2|A|$.

- Each edge in $B \setminus A$ is adjacent to at most two edges in $A \setminus B$.
- Each edge in $A \setminus B$ is adjacent to some edge in $B \setminus A$.
- Therefore $|A \setminus B| \leq 2|B \setminus A|$.



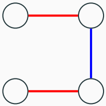


Relationship Between Maximum and Maximal

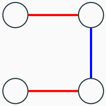
For any two maximal matchings A and B, we have

$$|A| \leq 2|B| \textbf{ and } |B| \leq 2|A|.$$

- Each edge in $B \setminus A$ is adjacent to at most two edges in $A \setminus B$.
- Each edge in $A \setminus B$ is adjacent to some edge in $B \setminus A$.
- Therefore $|A \setminus B| \leq 2|B \setminus A|$.



$$|A| \;=\; |A \cap B| + |A \setminus B|$$

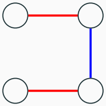


Relationship Between Maximum and Maximal

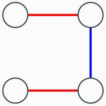
For any two maximal matchings A and B, we have

$$|A| \leq 2|B| \text{ and } |B| \leq 2|A|.$$

- Each edge in $B \setminus A$ is adjacent to at most two edges in $A \setminus B$.
- Each edge in $A \setminus B$ is adjacent to some edge in $B \setminus A$.
- Therefore $|A \setminus B| \leq 2|B \setminus A|$.



$$\begin{aligned} |A| &= |A \cap B| + |A \setminus B| \\ &\leq |A \cap B| + 2|B \setminus A| \end{aligned}$$

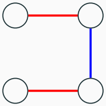


Relationship Between Maximum and Maximal

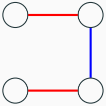
For any two maximal matchings A and B, we have

$$|A| \leq 2|B| \text{ and } |B| \leq 2|A|.$$

- Each edge in $B \setminus A$ is adjacent to at most two edges in $A \setminus B$.
- Each edge in $A \setminus B$ is adjacent to some edge in $B \setminus A$.
- Therefore $|A \setminus B| \leq 2|B \setminus A|$.



$$\begin{aligned} |A| &= |A \cap B| + |A \setminus B| \\ &\leq |A \cap B| + 2|B \setminus A| \\ &\leq 2|A \cap B| + 2|B \setminus A| \end{aligned}$$

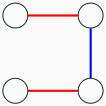


Relationship Between Maximum and Maximal

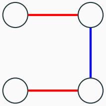
For any two maximal matchings A and B, we have

$|A| \leq 2|B|$ **and** $|B| \leq 2|A|$.

- Each edge in $B \setminus A$ is adjacent to at most two edges in $A \setminus B$.
- Each edge in $A \setminus B$ is adjacent to some edge in $B \setminus A$.
- Therefore $|A \setminus B| \leq 2|B \setminus A|$.



$$\begin{aligned} |A| &= |A \cap B| + |A \setminus B| \\ &\leq |A \cap B| + 2|B \setminus A| \\ &\leq 2|A \cap B| + 2|B \setminus A| \\ &= 2|B \cap A| + 2|B \setminus A| \end{aligned}$$

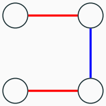


Relationship Between Maximum and Maximal

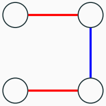
For any two maximal matchings A and B, we have

$$|A| \leq 2|B| \textbf{ and } |B| \leq 2|A|.$$

- Each edge in $B \setminus A$ is adjacent to at most two edges in $A \setminus B$.
- Each edge in $A \setminus B$ is adjacent to some edge in $B \setminus A$.
- Therefore $|A \setminus B| \leq 2|B \setminus A|$.



$$\begin{aligned} |A| &= |A \cap B| + |A \setminus B| \\ &\leq |A \cap B| + 2|B \setminus A| \\ &\leq 2|A \cap B| + 2|B \setminus A| \\ &= 2|B \cap A| + 2|B \setminus A| \\ &= 2|B| \end{aligned}$$



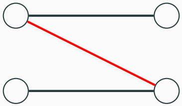
Algorithms for Maximum Cardinality Matching

General: $O(m\sqrt{n})$, Micali, Vazirani, 1980
$O(n^{\omega})$[1] Mucha, Sankowski, 2004

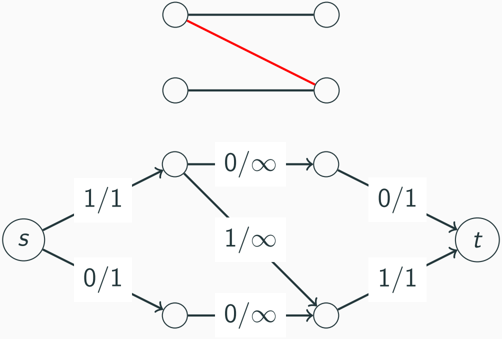
Bipartite: $O(m\sqrt{n})$, Hopcroft, Karp, 1973
$O(n^{\omega})$, Mucha, Sankowski, 2004
$\tilde{O}(m^{10/7})$, Madry, 2013

[1] $\omega \approx 2.37286...$

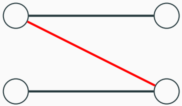
Augmenting Paths in Flows

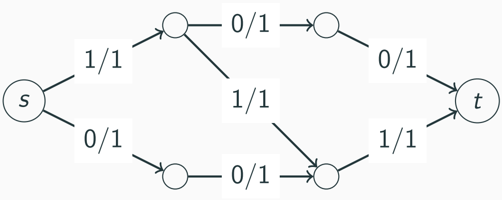


Augmenting Paths in Flows

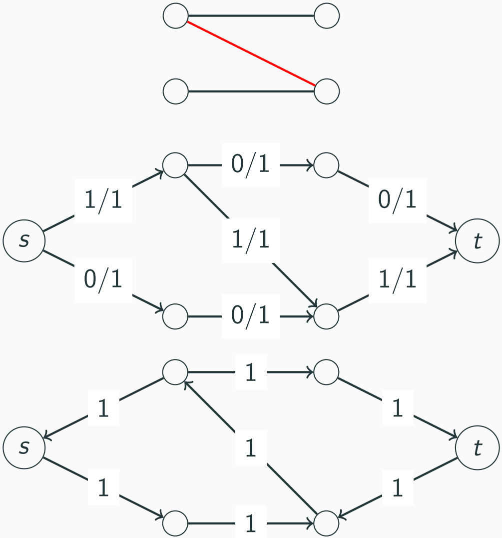


Augmenting Paths in Flows

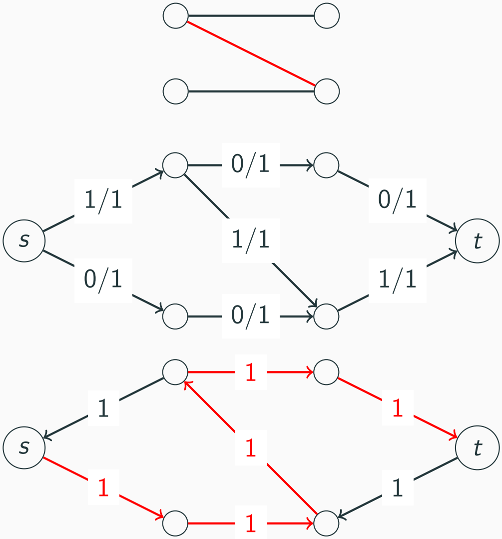




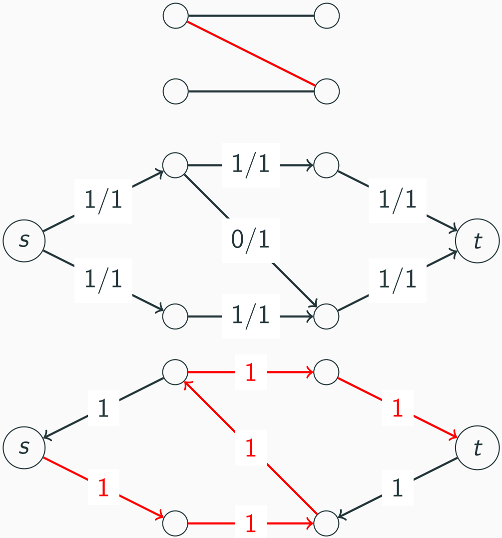
Augmenting Paths in Flows



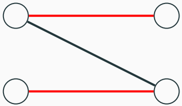
Augmenting Paths in Flows

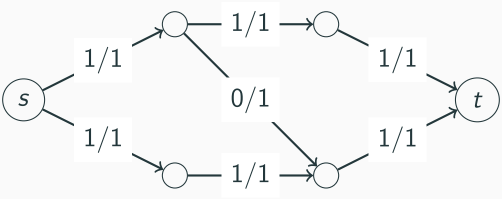


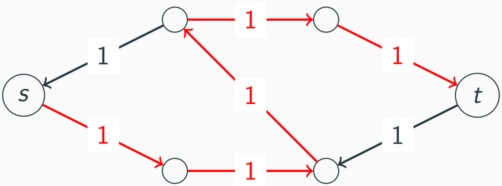
Augmenting Paths in Flows



Augmenting Paths in Flows







Basics

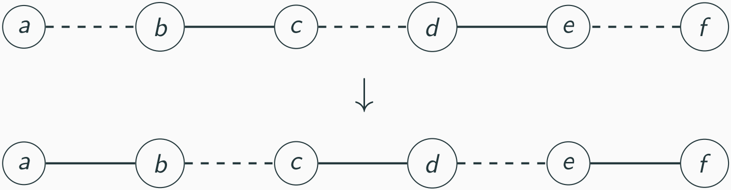
Augmenting Paths: A path starting and ending with two free nodes such that matched and unmatched edges alternate.

Basics

Augmenting Paths: A path starting and ending with two free nodes such that matched and unmatched edges alternate.

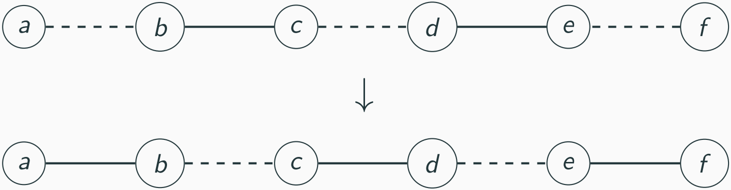
Basics

Augmenting Paths: A path starting and ending with two free nodes such that matched and unmatched edges alternate.



Basics

Augmenting Paths: A path starting and ending with two free nodes such that matched and unmatched edges alternate.



Given a matching M and an augmenting path p, we obtain a larger matching $M' = M \bigoplus p$ with $|M'| = |M| + 1$.

Optimum and Augmenting Paths

Lemma 0.5 (Berge's Theorem)

M is an optimal matching if and only if there exist no augmenting paths with respect to M.

$\Rightarrow$: Augmenting paths increase matching by 1.

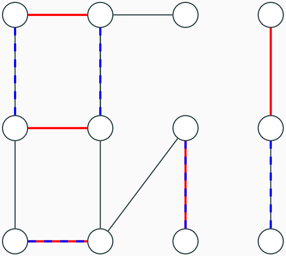
$\Leftarrow$: Let M' be an optimum.

$M' \bigoplus M$ can consist of:

1. Isolated vertices
2. Even length cycles
3. Even length paths
3. Odd length paths

M' and M agree in size on everything except for the odd length paths.

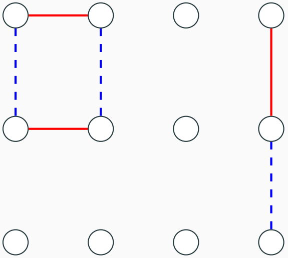
Augmenting paths are odd length paths.



Optimum and Augmenting Paths

Lemma 0.5 (Berge's Theorem)

M is an optimal matching if and only if there exist no augmenting paths with respect to M.



$\Rightarrow$**:** Augmenting paths increase matching by 1.

$\Leftarrow$**:** Let M' be an optimum.

$M' \bigoplus M$ can consist of:

1. Isolated vertices
2. Even length cycles
3. Even length paths
3. Odd length paths

M' and M agree in size on everything except for the odd length paths.

Augmenting paths are odd length paths.

First Attempt

Successive Augmenting Path

Compute a maximal matching M

For $i = 1$ **to** $n/2$

  For every unmatched node v

    Find an augmenting path p starting at v.

    $M = M \bigoplus p$

Running time: $O(mn^2)$

Correctness: Berge's Theorem

Giving the Algorithm a Bit of Structure

Lemma 1

Let M be a matching where the shortest augmenting path with respect to M has length k. Let P be a maximal set of edge disjoint augmenting paths of length k. Then the shortest augmenting path with respect to $M' = M \bigoplus P$ has length at least $k + 2$.

Let π be an augmenting path with respect to M'.

If π does not intersect any path from P, its length is greater than k.

Otherwise its existence contradicts the assumption that P is maximal or that the paths in P are shortest.

Proof of Lemma 1 Continued

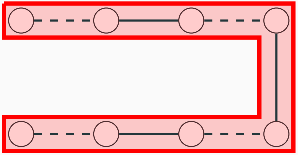
Let π intersect with path $p \in P$ in $M \bigoplus P$.

Then $\pi \bigoplus p$ were two augmenting paths a and b with respect to M.

Proof of Lemma 1 Continued

Let π intersect with path $p \in P$ in $M \bigoplus P$.

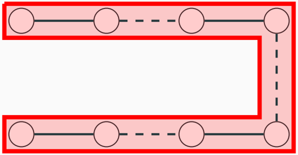
Then $\pi \bigoplus p$ were two augmenting paths a and b with respect to M.



Proof of Lemma 1 Continued

Let π intersect with path $p \in P$ in $M \bigoplus P$.

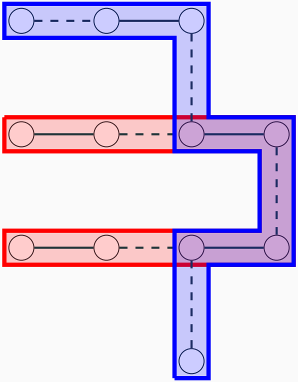
Then $\pi \bigoplus p$ were two augmenting paths a and b with respect to M.



Proof of Lemma 1 Continued

Let π intersect with path $p \in P$ in $M \bigoplus P$.

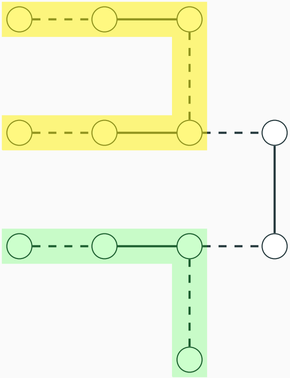
Then $\pi \bigoplus p$ were two augmenting paths a and b with respect to M.



Proof of Lemma 1 Continued

Let π intersect with path $p \in P$ in $M \bigoplus P$.

Then $\pi \bigoplus p$ were two augmenting paths a and b with respect to M.

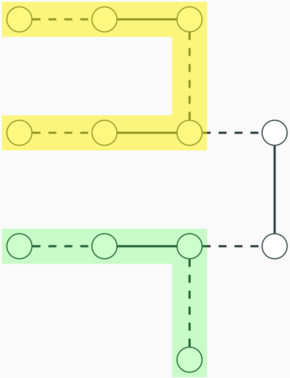


Proof of Lemma 1 Continued

Let π intersect with path $p \in P$ in $M \bigoplus P$.

Then $\pi \bigoplus p$ were two augmenting paths a and b with respect to M.

$|a| + |b| < |p| + |\pi| \leq 2|p|$



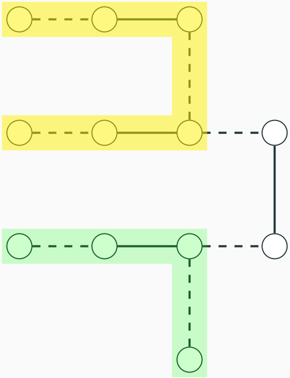
Proof of Lemma 1 Continued

Let π intersect with path $p \in P$ in $M \bigoplus P$.

Then $\pi \bigoplus p$ were two augmenting paths a and b with respect to M.

$|a| + |b| < |p| + |\pi| \leq 2|p|$

Contradiction to the assumption that p was shortest



Proof of Lemma 1 Continued

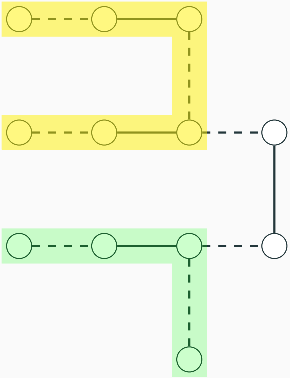
Let π intersect with path $p \in P$ in $M \bigoplus P$.

Then $\pi \bigoplus p$ were two augmenting paths a and b with respect to M.

$|a| + |b| < |p| + |\pi| \leq 2|p|$

Contradiction to the assumption that p was shortest

By induction, a similar claim holds if π intersects with $p_1, \ldots p_j \in P$



Second Attempt

Successive Shortest Augmenting Path

Compute a maximal matching M

For $i = 1$ **to** n

$P \leftarrow \emptyset$

    While augmenting paths of length $2i + 1$ exist

        Find an augmenting path p of length $2i + 1$.

        Add p to P

        Remove p and incident edges from the graph G

    $M = M \bigoplus P$

    Restore G

Naive way to determine P: Run a BFS for every free node.

Second Attempt

Successive Shortest Augmenting Path

Compute a maximal matching M

For $i = 1$ **to** n

$P \leftarrow \emptyset$

  While augmenting paths of length $2i + 1$ exist

    Find an augmenting path p of length $2i + 1$.

    Add p to P

    Remove p and incident edges from the graph G

  $M = M \bigoplus P$

  Restore G

Naive way to determine P: Run a BFS for every free node.

We can do better ☺

Second Attempt

Successive Shortest Augmenting Path

Compute a maximal matching M

For $i = 1$ **to** n

$P \leftarrow \emptyset$

  While augmenting paths of length $2i + 1$ exist

    Find an augmenting path p of length $2i + 1$.

    Add p to P

    Remove p and incident edges from the graph G

  $M = M \bigoplus P$

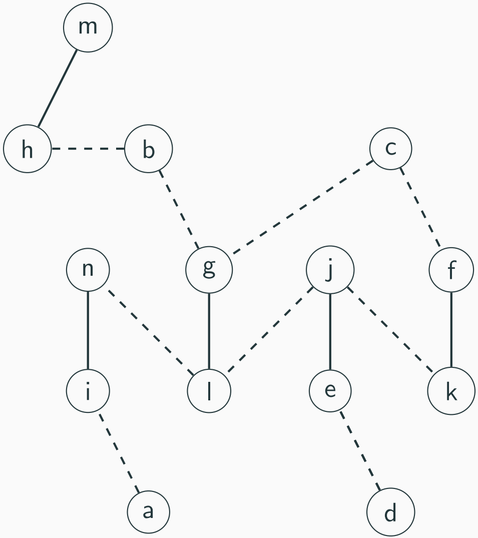
  Restore G

Naive way to determine P: Run a BFS for every free node.

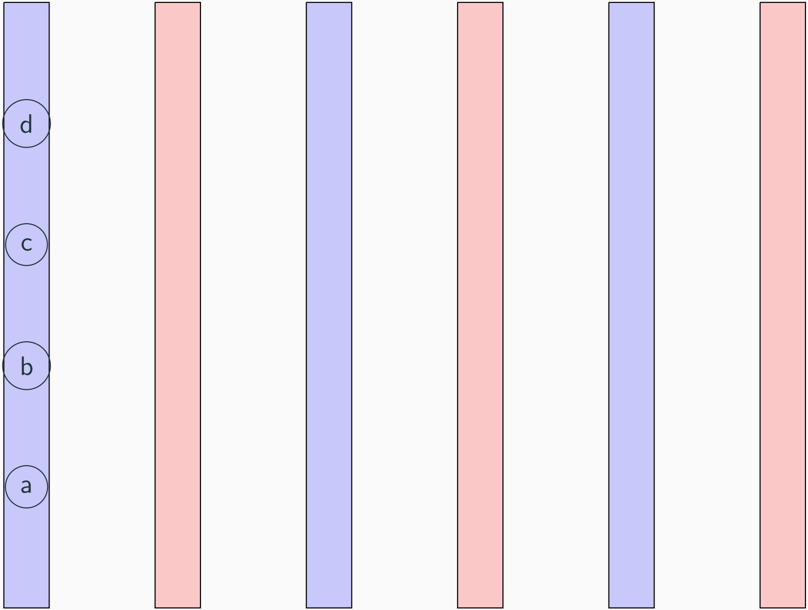
We can do better ☺

Compute P with just one *BFS* and one *DFS*

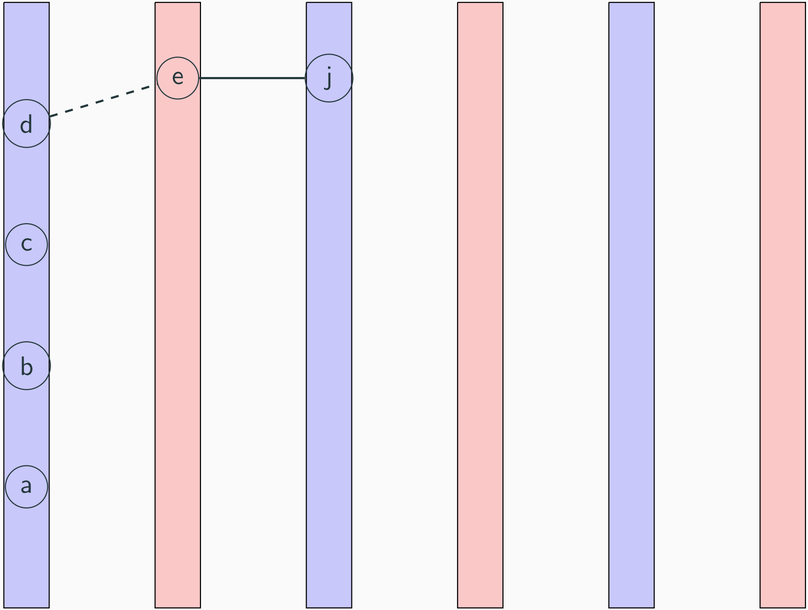
Running Example



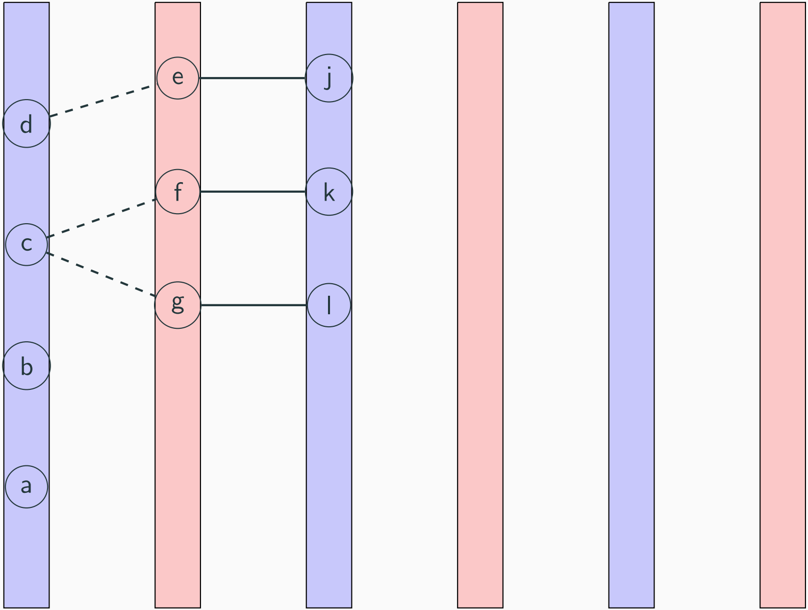
Hopcroft-Karp Trees



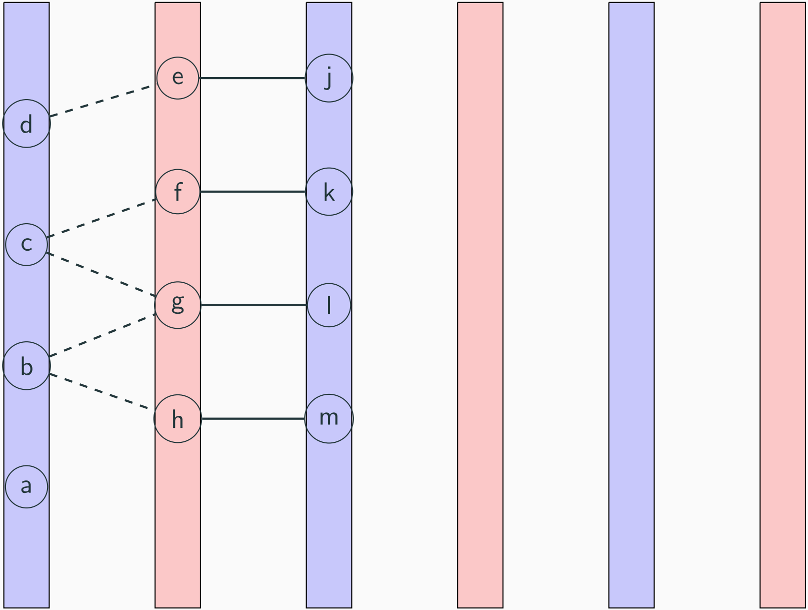
Hopcroft-Karp Trees



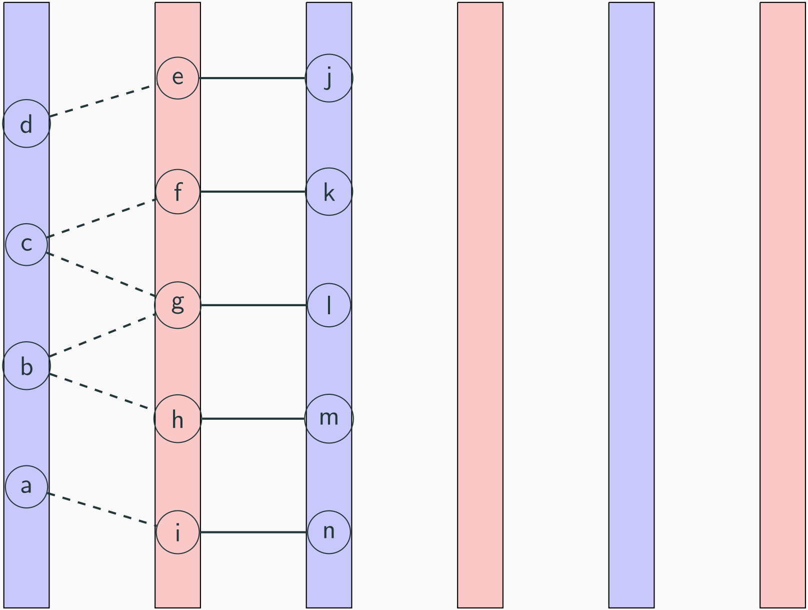
Hopcroft-Karp Trees



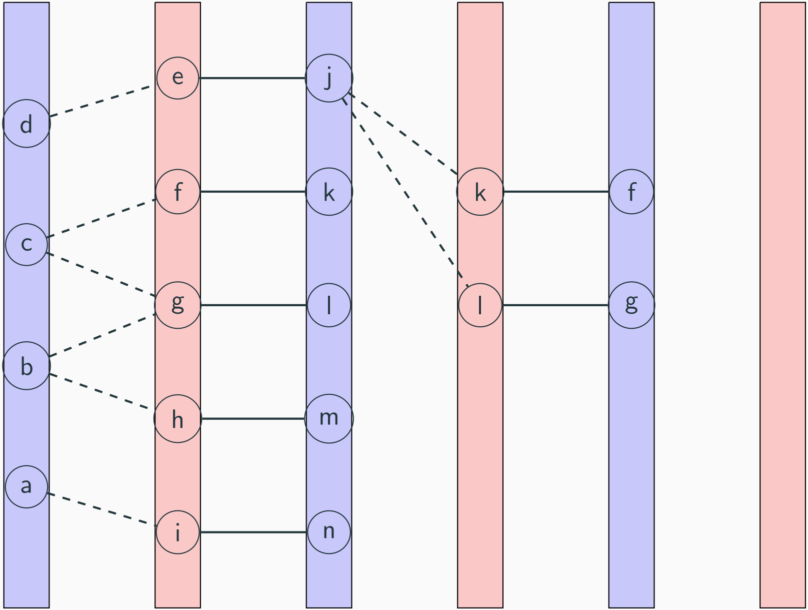
Hopcroft-Karp Trees



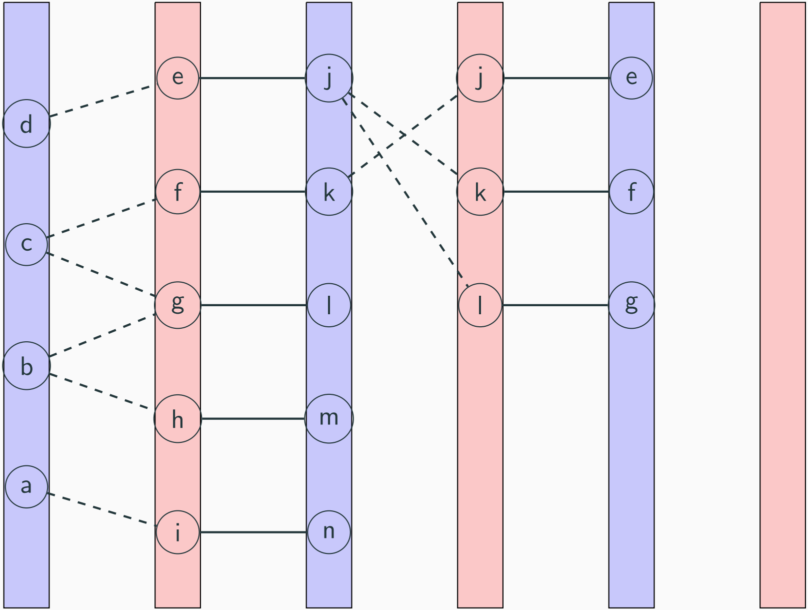
Hopcroft-Karp Trees



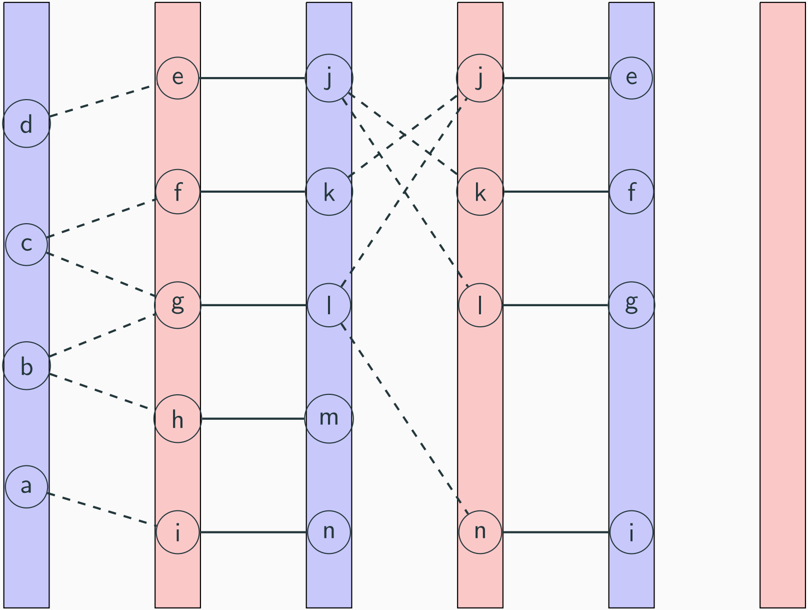
Hopcroft-Karp Trees



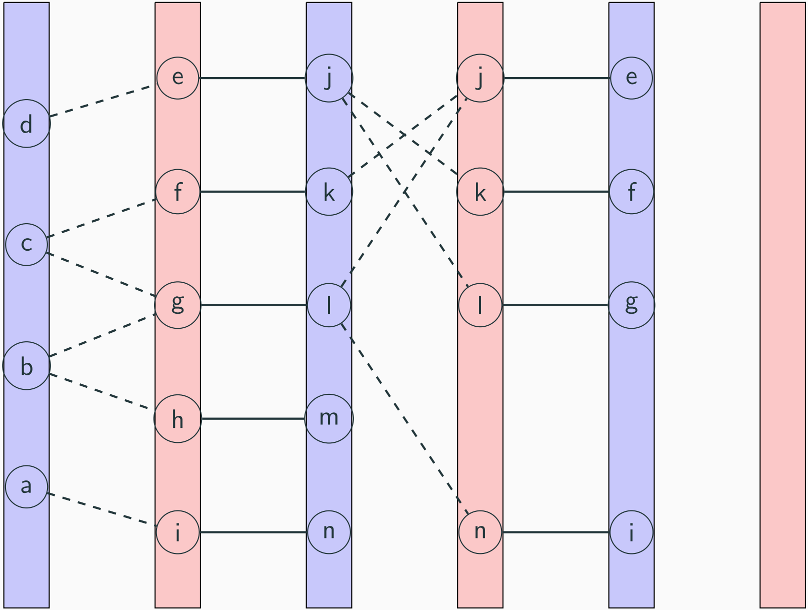
Hopcroft-Karp Trees



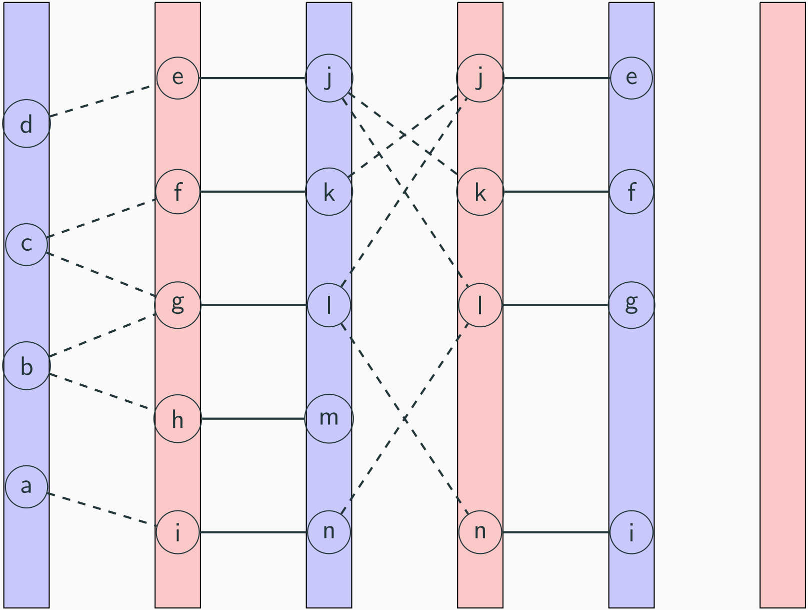
Hopcroft-Karp Trees



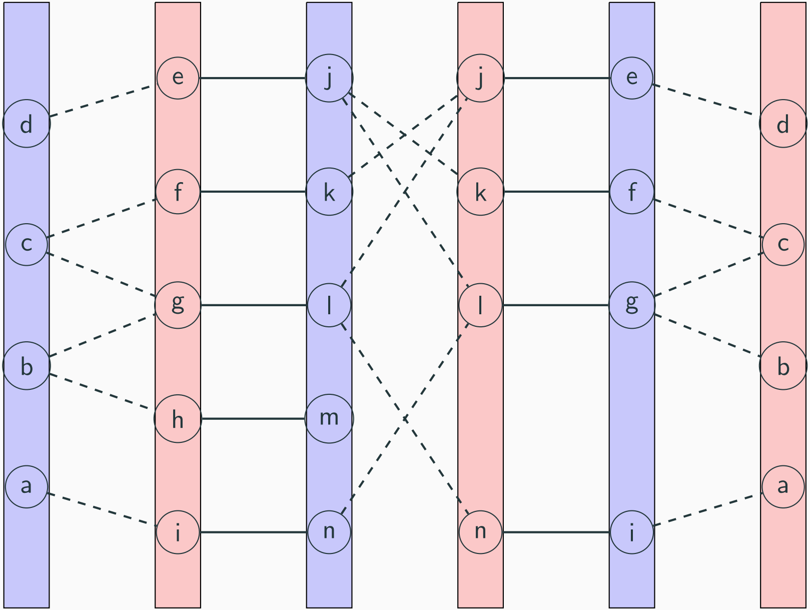
Hopcroft-Karp Trees



Hopcroft-Karp Trees



Hopcroft-Karp Trees



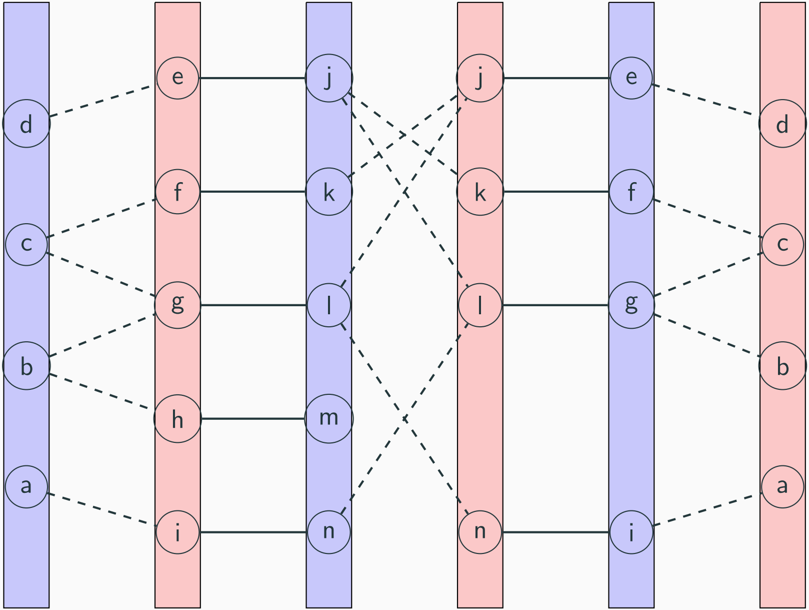
Formal Intermission

Hopcroft-Karp Tree Initialization

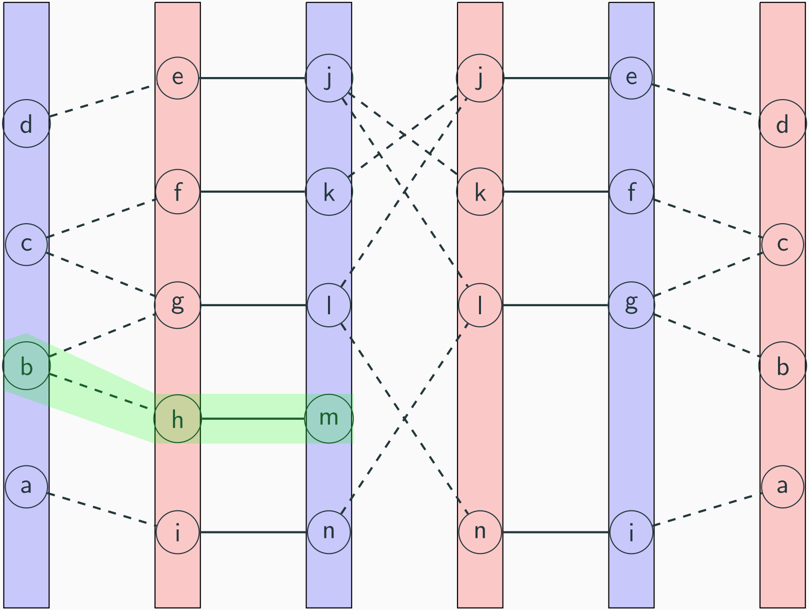
Put all free nodes into a queue
Run alternating path BFS
Place no node into more than one even (red) or odd (blue) level

Running Time: $O(m + n)$

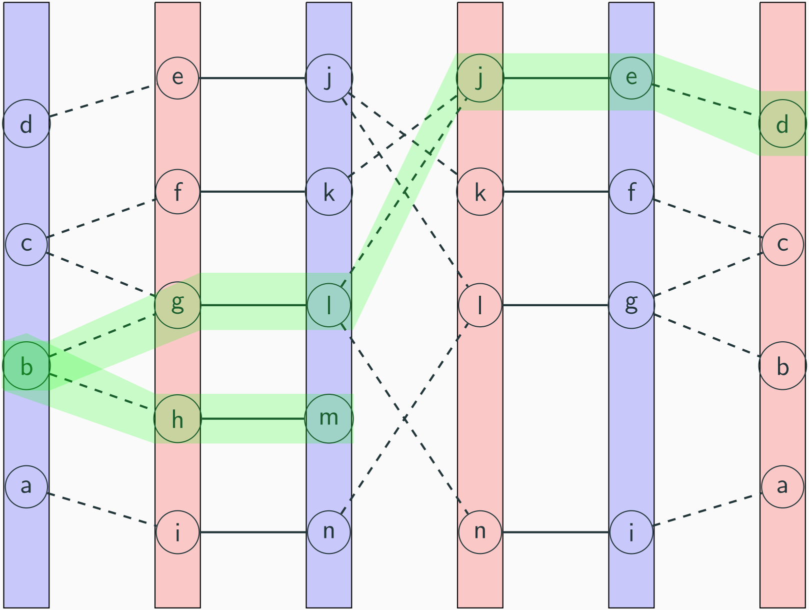
Hopcroft-Karp Trees



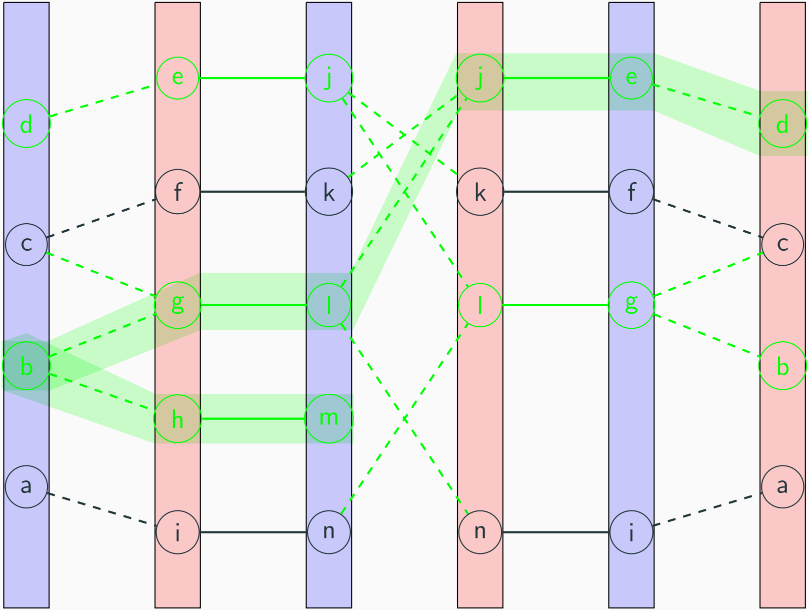
Hopcroft-Karp Trees



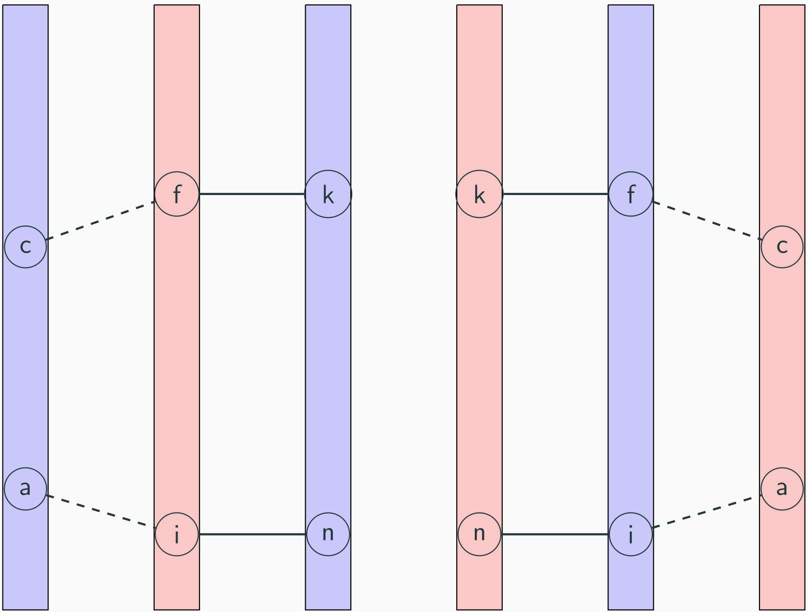
Hopcroft-Karp Trees



Hopcroft-Karp Trees



Hopcroft-Karp Trees



Disjoint Augmenting Path Detection

Hopcroft-Karp Tree Initialization

Put all free nodes into a queue

Run alternating path BFS

Place no node into more than one even (red) or odd (blue) level

Path Extraction

$P \leftarrow \emptyset$

For all free nodes in the first level

  Run a DFS in the Hopcroft-Karp Tree

  If path p is found, add p to P and stop.

  Remove every node and any incident edges of the DFS traversal.

Running Time: $O(m + n)$

Analysis of Hopcroft-Karp Trees

Claim

We can find a maximal set of disjoint augmenting paths in $O(m+n)$ time.

Hopcroft-Karp Trees contain *all* augmenting paths up to a given depth.

A DFS from a free node u finds an augmenting path starting at u, or none exist.

If a node a was visited by a DFS traversal, but not included in an augmenting path, then no augmenting path containing a exists (that does not overlap with P).

If a DFS rooted at u yields an augmenting path p, then no augmenting path disjoint from p at v can use one of the nodes in p.

Third Attempt

Hopcroft-Karp

Compute a maximal matching M

For $i = 1$ **to** n

$\quad P \leftarrow \emptyset$

$\quad$ Find a maximal set P of disjoint augmenting paths of length $2i + 1$.

$\quad M = M \bigoplus P$

Running Time: $O(m \cdot n)$

Third Attempt

Hopcroft-Karp

Compute a maximal matching M

For $i = 1$ **to** n

  $P \leftarrow \emptyset$

  Find a maximal set P of disjoint augmenting paths of length $2i + 1$.

  $M = M \bigoplus P$

Running Time: $O(m \cdot n)$

This is (more or less) the final algorithm.

Third Attempt

Hopcroft-Karp

Compute a maximal matching M

For $i = 1$ **to** n

  $P \leftarrow \emptyset$

  Find a maximal set P of disjoint augmenting paths of length $2i + 1$.

  $M = M \bigoplus P$

Running Time: $O(m \cdot n)$

This is (more or less) the final algorithm.

I promised you $O(m\sqrt{n})$.

Detailed View on Berge's Theorem

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

Detailed View on Berge's Theorem

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

Let M' be an optimum matching.

Detailed View on Berge's Theorem

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

Let M' be an optimum matching.

Consider $M' \bigoplus M$. It contains at most $|M'| - |M|$ augmenting paths.

Detailed View on Berge's Theorem

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

Let M' be an optimum matching.

Consider $M' \bigoplus M$. It contains at most $|M'| - |M|$ augmenting paths.

Each of these paths has length at least k.

Detailed View on Berge's Theorem

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

Let M' be an optimum matching.

Consider $M' \bigoplus M$. It contains at most $|M'| - |M|$ augmenting paths.

Each of these paths has length at least k.

The paths are disjoint, so their total length is at most n.

Detailed View on Berge's Theorem

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

Let M' be an optimum matching.

Consider $M' \bigoplus M$. It contains at most $|M'| - |M|$ augmenting paths.

Each of these paths has length at least k.

The paths are disjoint, so their total length is at most n.

There are are no more then $\frac{n}{k}$ such paths.

Applications

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

Applications

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

We run Hopcroft-Karp until $k \geq \sqrt{n}$.

Applications

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

We run Hopcroft-Karp until $k \geq \sqrt{n}$.

We have at most $\frac{n}{\sqrt{n}} = \sqrt{n}$ paths left.

Applications

Lemma 2

Let M' be an optimum matching and let M be any matching. If the length of the shortest augmenting path with respect to M is k, then $|M'| - |M| \leq \frac{n}{k}$.

We run Hopcroft-Karp until $k \geq \sqrt{n}$.

We have at most $\frac{n}{\sqrt{n}} = \sqrt{n}$ paths left.

The algorithm now only computes $\sqrt{n}$ remaining iterations of BFS/DFS searches.

Final Algorithm

Hopcroft-Karp

Compute a maximal matching M

For $i = 1$ **to** $\sqrt{n}$

$P \leftarrow \emptyset$

Find a maximal set P of disjoint augmenting paths in Hopcroft-Karp trees of depth $2i + 1$.

$M = M \bigoplus P$

For $i = 1$ **to** $\sqrt{n}$

$P \leftarrow \emptyset$

Find a maximal set P of disjoint augmenting paths in Hopcroft-Karp trees of depth n.

$M = M \bigoplus P$