

Hi everyone!!

CS 374A Review MT2

Electric Boogaloo

ACM @ UIUC

November 10, 2024



Disclaimers and Logistics

- **Disclaimer:** Some of us are CAs, but we have not seen the exam. We have no idea what the questions are. However, we've taken the course and reviewed Sarel's previous exams, so we have **suspicions** as to what the questions will be like.
- This review session is being recorded. Recordings and slides will be distributed on EdStem after the end.
- **Agenda:** We'll quickly review all topics likely to be covered, then go through a practice exam, then review individual topics by request.
 - Questions are designed to be written in the same style as ~~Kari~~^{Sarel's} previous exams but to be *slightly* harder, so don't worry if you don't get everything right away!
- Please let us know if we're going too fast/slow, not speaking loud enough/speaking too loud, etc.
- If you have a question anytime during the review session, please ask! Someone else almost surely has a similar question.
- We'll provide a feedback form at the end of the session.

Recursion

- **Definition:** Reducing the problem to a smaller instance of itself, where eventually we can terminate in a base case.
 - Think: If we have a problem of size n , we want to continuously reduce to a problem smaller than n .
 - Example: Tower of Hanoi

Template

```
1: procedure AMAZINGRECURSIVEALGO( $n$ )
2:   if  $n ==$  [some base case] then
3:     return [value]
4:   else
5:     return AmazingRecursiveAlgo( $n - 1$ )
```

- Similar to **induction!**

Recursion: Runtime Analysis

- **General Form:**

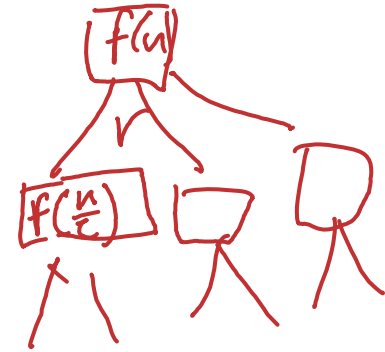
$$T(n) = \underbrace{r}_{\text{\# of subproblems}} \cdot \overbrace{T\left(\frac{n}{c}\right)}^{\text{work at each subproblem}} + \underbrace{f(n)}_{\text{work at current level}}$$

- Describes how the amount of work changes between each level of recursion.
- We can solve for a **time complexity** that describes the scaling behaviour of the algorithm at hand.

Recursion: Runtime Analysis

- **General Form:**

$$T(n) = \underbrace{r}_{\text{\# of subproblems}} \cdot \overbrace{T\left(\frac{n}{c}\right)}^{\text{work at each subproblem}} + \underbrace{f(n)}_{\text{work at current level}}$$



- Describes how the amount of work changes between each level of recursion.
- We can solve for a **time complexity** that describes the scaling behaviour of the algorithm at hand.

- **Master's Theorem**

Master's Theorem

Decreasing: $r \cdot f(n/c) = \kappa \cdot f(n)$ where $\kappa < 1 \implies T(n) = O(f(n))$

Equal: $r \cdot f(n/c) = \kappa \cdot f(n)$ where $\kappa = 1 \implies T(n) = O(f(n) \cdot \log_c n)$

Increasing: $r \cdot f(n/c) = \kappa \cdot f(n)$ where $\kappa > 1 \implies T(n) = O(n^{\log_c r})$

Recursion: Runtime Analysis

- **General Form:**

$$T(n) = \underbrace{r}_{\text{\# of subproblems}} \cdot \overbrace{T\left(\frac{n}{c}\right)}^{\text{work at each subproblem}} + \underbrace{f(n)}_{\text{work at current level}}$$

- Describes how the amount of work changes between each level of recursion.
- We can solve for a **time complexity** that describes the scaling behaviour of the algorithm at hand.

- **Master's Theorem**

Master's Theorem

Decreasing: $r \cdot f(n/c) = \kappa \cdot f(n)$ where $\kappa < 1 \implies T(n) = O(f(n))$

Equal: $r \cdot f(n/c) = \kappa \cdot f(n)$ where $\kappa = 1 \implies T(n) = O(f(n) \cdot \log_c n)$

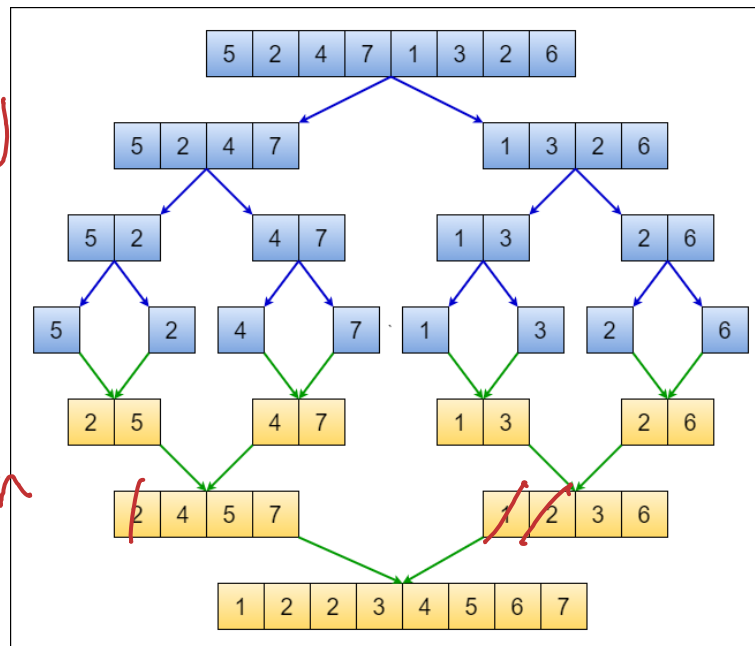
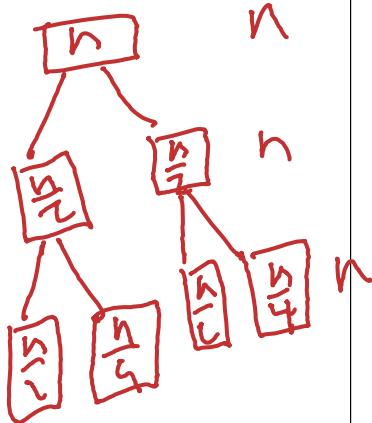
Increasing: $r \cdot f(n/c) = \kappa \cdot f(n)$ where $\kappa > 1 \implies T(n) = O(n^{\log_c r})$

- **Intuition:** If each level contains more work than the level below it, then the root level will dominate. If each level contains the same amount of work, then we have $\log_c n$ levels with $f(n)$ work. If each level contains less work than the work below it, then the leaf nodes will dominate.

Divide and Conquer Algos: Merge Sort

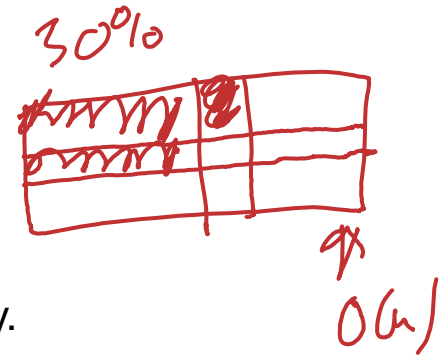
- **Purpose:** Sort an arbitrary array.
- **Time Complexity:** $O(n \log n)$
- **Intuition:** Three phases: (a) split the array in half, (b) sort each side, (c) merge the sorted halves by repeatedly comparing smallest elements on each side not yet inserted.

$$T(n) = 2T(n/2) + O(n)$$

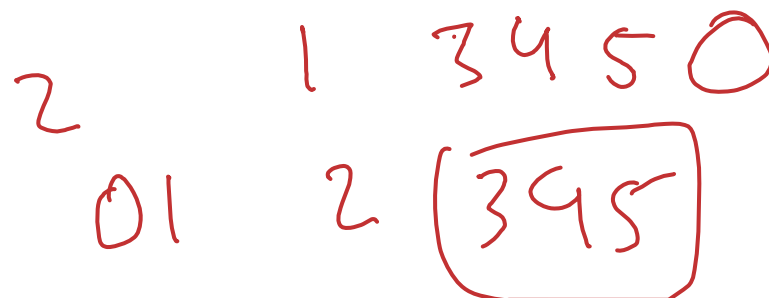


$$O(n \log n)$$

Divide and Conquer Algos: Quickselect



- **Purpose:** Get the n^{th} smallest element in an arbitrary array.
- **Time Complexity:** Avg: $O(n)$ | Worst; $O(n^2)$, ($O(n)$ with MoM)
- **Intuition:** Pick a pivot P with a value P_V and rearrange the array such that all the elements that are less than P_V are to the left of P and all the elements that are greater than P_V are to the right of P , just like quick select. If the length of the array of elements that are less than P_V is greater than n , then we know that the n^{th} smallest element is to the left of P and we recurse on the left subarray. Otherwise, we know that the n^{th} smallest element is to the right of P and we recurse on the right subarray.
 - **Why the poor worst case performance?**
 - Again, because we can get unlucky and pick the worst possible pivot at every step.
 - We can guarantee linear performance with a better pivot-picking algorithm such as MEDIANOFMEDIANS
 - ▶ Finds element that larger than $\frac{3}{10}$ and smaller than $\frac{7}{10}$ of the array's elements.
 - ▶ Runs in $O(n)$ time



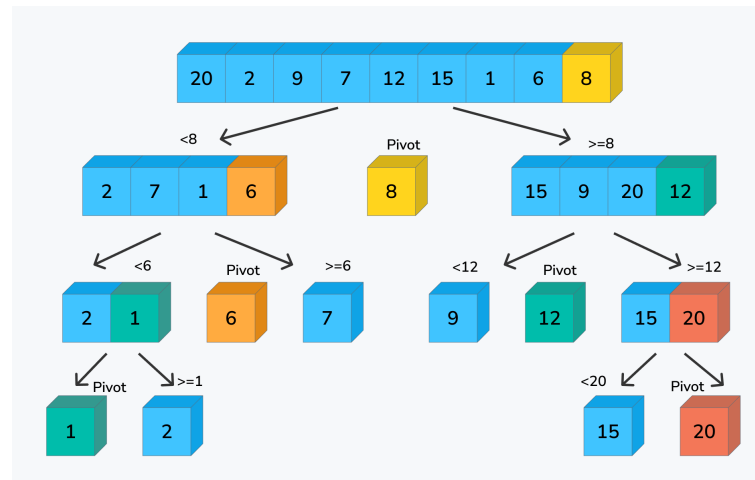
$$T(n) = O(n) + T\left(\frac{3n}{10}\right) + O(n)$$

$$T(n) = O(n)$$

4th smallest

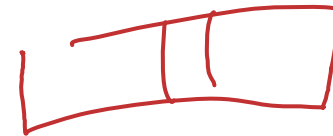
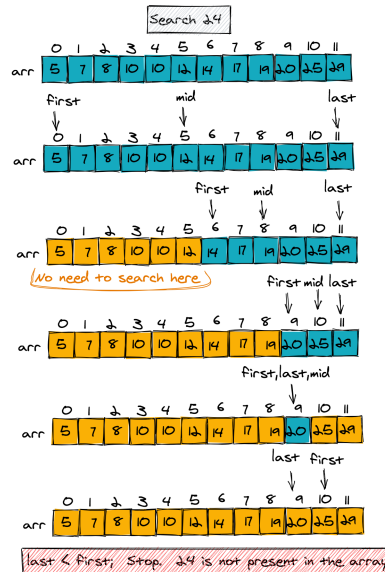
Divide and Conquer Algos: Quicksort

- **Purpose:** Sort an arbitrary array.
- **Time Complexity:** Avg: $O(n \log n)$ | Worst: $O(n^2)$ ($O(n \log n)$ deterministic with quickselect partitioning)
- **Intuition:** Pick a pivot and rearrange the array such that all the elements that are less than the pivot value are to the left of the pivot value and all the elements that are greater than the pivot value are to the right of the pivot value. Then sort each side.
 - **Why the poor worst case performance?**
 - Because we can get unlucky and pick the worst possible pivot at every step.



Divide and Conquer Algos: Binary Search

- **Purpose:** Find the existence of an element in a sorted array
- **Time Complexity:** $O(\log n)$
- **Intuition:** Say we are trying to find the value n . Pick the middle element M in the array. If $n > M$, the element must be to the right of n and we recurse on the right. Otherwise, we recurse on the left.



$$T(n) = T(n/2) + O(1)$$

$$T(n) = \log_2 n$$

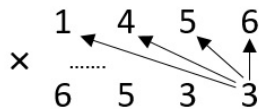
Divide and Conquer Algos: Karatsuba's

- **Purpose:** Multiplication
- **Time Complexity:** $O(n^{\log_2 3}) \approx O(n^{1.585})$
- 3 Phases:
 - **Divide:** Represent x as $x_0p^n + x_1$, y as $y_0p^n + y_1$
 - **Recurse:** Calculate x_0y_0 , x_1y_1 , $(x_0 + x_1)(y_0 + y_1)$
 - **Combine:** Use the three results to calculate our final answer.

$$A = a_1 \cdot 2^p + a_2$$

$$B = b_1 \cdot 2^p + b_2$$

$$T(n) = 3T\left(\frac{n}{2}\right) + O(n)$$



	(3×1)	(3×4)	(3×5)	(3×6)		
	(3×1)	(3×4)	(3×5)	(3×6)	0	
	(5×1)	(5×4)	(5×5)	(5×6)	0	0
	(6×1)	(6×4)	(6×5)	(6×6)	0	0
					0	0
						0

16 single-digit x single-digit multiplications

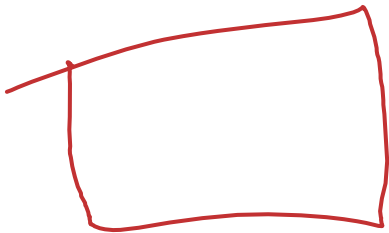
95,12,048

$$a_1 b_1 \cdot 2^{2p} + (a_1 b_2 + a_2 b_1) \cdot 2^{p-1} + a_2 b_2$$

$$\cancel{a_1 b_1 + a_2 b_2} + (a_2 b_1 + a_1 b_2)$$

Backtracking

- Technique to methodically explore the solutions to a problem via the reduction to said problem to a smaller variant of itself, a.k.a **recursion**.
- Intuitively, think of the problem space as a maze that we are trying to find the exit of. For each path, you would traverse until you reach a dead end, at which point you **back track** to try a different path.
- To find recurrence, think "What information about a subset of my current problem space would be really nice to know?"



Backtracking

- Technique to methodically explore the solutions to a problem via the reduction to said problem to a smaller variant of itself, a.k.a **recursion**.
- Intuitively, think of the problem space as a maze that we are trying to find the exit of. For each path, you would traverse until you reach a dead end, at which point you **back track** to try a different path.
- To find recurrence, think "What information about a subset of my current problem space would be really nice to know?" **Example:** Longest Increasing Subsequence
- "What is the length of a longest increasing subsequence in an arbitrary array?"

Backtracking

- Technique to methodically explore the solutions to a problem via the reduction to said problem to a smaller variant of itself, a.k.a **recursion**.
- Intuitively, think of the problem space as a maze that we are trying to find the exit of. For each path, you would traverse until you reach a dead end, at which point you **back track** to try a different path.
- To find recurrence, think "What information about a subset of my current problem space would be really nice to know?" **Example:** Longest Increasing Subsequence
- "What is the length of a longest increasing subsequence in an arbitrary array?"

$$\text{LIS}(i, j) = \begin{cases} 0 & \text{if } i = 0 \\ \text{LIS}(i - 1, j) & \text{if } A[i] \geq A[j] \\ \max \begin{cases} \text{LIS}(i - 1, j) \\ 1 + \text{LIS}(i - 1, i) \end{cases} & \text{else} \end{cases}$$

Handwritten notes:
 ... i
 A[j]

Backtracking

- Technique to methodically explore the solutions to a problem via the reduction to said problem to a smaller variant of itself, a.k.a **recursion**.
- Intuitively, think of the problem space as a maze that we are trying to find the exit of. For each path, you would traverse until you reach a dead end, at which point you **back track** to try a different path.
- To find recurrence, think "What information about a subset of my current problem space would be really nice to know?" **Example:** Longest Increasing Subsequence
- "What is the length of a longest increasing subsequence in an arbitrary array?"

$$\text{LIS}(i, j) = \begin{cases} 0 & \text{if } i = 0 \\ \text{LIS}(i - 1, j) & \text{if } A[i] \geq A[j] \\ \max \begin{cases} \text{LIS}(i - 1, j) \\ 1 + \text{LIS}(i - 1, i) \end{cases} & \text{else} \end{cases}$$

This kind of sucks; we're redoing computation that we've already done! What if instead, we computed all the subproblems beforehand, wrote down the solutions, then did the recursion?

Dynamic Programming

- It's backtracking, but we compute all of the subproblems iteratively.
 - This idea of "writing things down" as to not repeat computation is called **memoization**

Dynamic Programming

- It's backtracking, but we compute all of the subproblems iteratively.
 - This idea of "writing things down" as to not repeat computation is called **memoization**
- Alternatively, you can think about this recursively, except we check our memoization structure to see if we've computed anything before. If we have, we just use the computed result. Otherwise, we compute the subproblem.

@lvj_came

Dynamic Programming

- It's backtracking, but we compute all of the subproblems iteratively.
 - This idea of "writing things down" as to not repeat computation is called **memoization**
- Alternatively, you can think about this recursively, except we check our memoization structure to see if we've computed anything before. If we have, we just use the computed result. Otherwise, we compute the subproblem.
- For a DP solution, we need:
 1. English Description
 2. Recurrence
 3. Memoization Structure
 4. Solution Location
 5. Evaluation Order
 6. Runtime

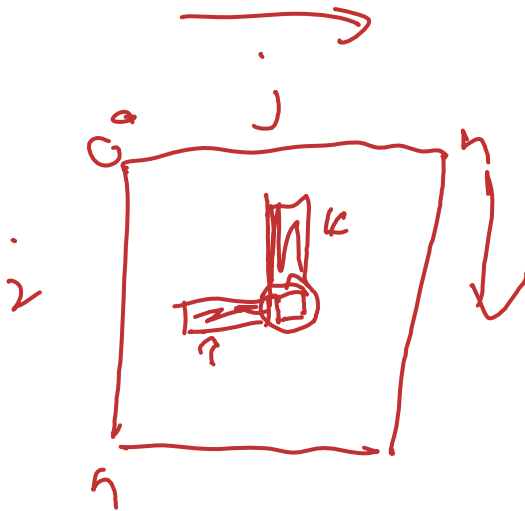
Dynamic Programming

- It's backtracking, but we compute all of the subproblems iteratively.
 - This idea of "writing things down" as to not repeat computation is called **memoization**
- Alternatively, you can think about this recursively, except we check our memoization structure to see if we've computed anything before. If we have, we just use the computed result. Otherwise, we compute the subproblem.
- For a DP solution, we need:
 1. English Description
 2. Recurrence
 3. Memoization Structure
 4. Solution Location
 5. Evaluation Order
 6. Runtime
- **How to solve a DP:**
 - Identify how we can take advantage of a recursive call on a smaller subset of the input space.
 - Identity base cases
 - Identity recurrences (they should cover all possible cases at each step)

Dynamic Programming

Let's look at the LIS example from before: "What is the length of a longest increasing subsequence in an arbitrary array?"

$$LIS(i, j) = \begin{cases} 0 & \text{if } i = 0 \\ LIS(i-1, j) & \text{if } A[i] > A[j] \\ \max \begin{cases} LIS(i-1, j) \\ 1 + LIS(i-1, i) \end{cases} & \text{otherwise} \end{cases}$$



we memo in a 2D array of size $[0 \dots n][0 \dots n]$

Dynamic Programming

Let's look at the LIS example from before: "What is the length of a longest increasing subsequence in an arbitrary array?"

procedure LIS-ITERATIVE($A[1..n]$):

$A \leftarrow [1 \dots n][1 \dots n]$

for all $i \leftarrow 1 \dots n$ **do**

for all $j \leftarrow i \dots n$ **do**

if $A[i] \leq A[j]$ **then**

$LIS[i][j] = 1$

else

$LIS[i][j] = 0$

for all $i \leftarrow 1 \dots n$ **do**

for all $j \leftarrow \tilde{2} \dots n$ **do**

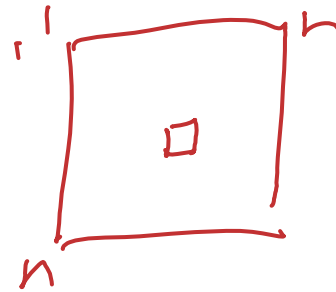
if $A[i] \geq A[j]$ **then**

$LIS[i][j] = LIS[i-1, j]$

else

$LIS[i][j] = \max \begin{cases} LIS[i-1, j] \\ LIS[i-1, i] + 1 \end{cases}$

return $LIS[n, n]$



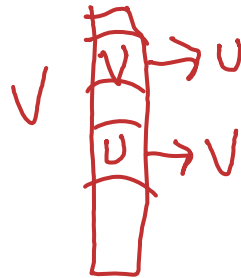
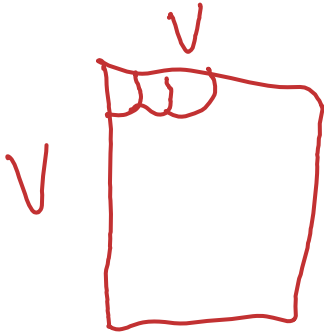
1, 2, 3, -4, 5

Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**

$$G = (V, E)$$

$$(U, V) \quad U \rightarrow V$$

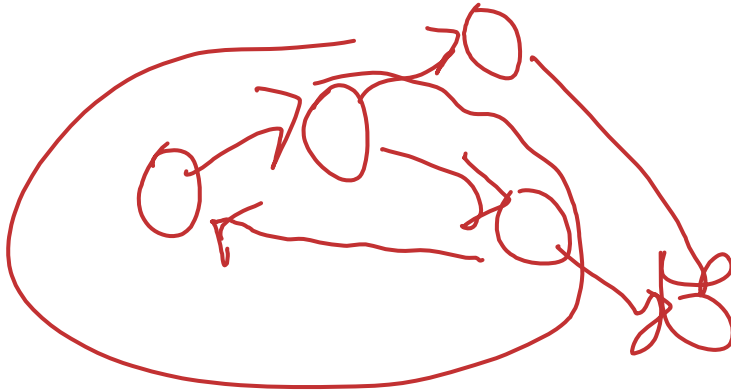


Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**
 - **Directed:** Each edge $(u, v) \in E$ now has a direction $u \rightarrow v$

Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**
 - **Directed:** Each edge $(u, v) \in E$ now has a direction $u \rightarrow v$
 - **Acyclic:** No cycles.



Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**
 - **Directed:** Each edge $(u, v) \in E$ now has a direction $u \rightarrow v$
 - **Acyclic:** No cycles.
- **Path:** A sequence of distinct vertices where each pair of consecutive vertices have an edge



Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**
 - **Directed:** Each edge $(u, v) \in E$ now has a direction $u \rightarrow v$
 - **Acyclic:** No cycles.
- **Path:** A sequence of distinct vertices where each pair of consecutive vertices have an edge
- **Cycle:** A sequence of distinct vertices where each pair of consecutive vertices have an edge **and** the first and last vertices are connected.

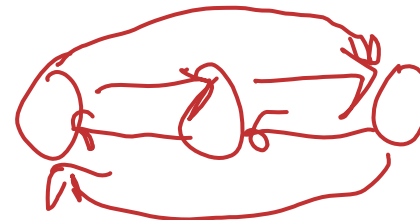
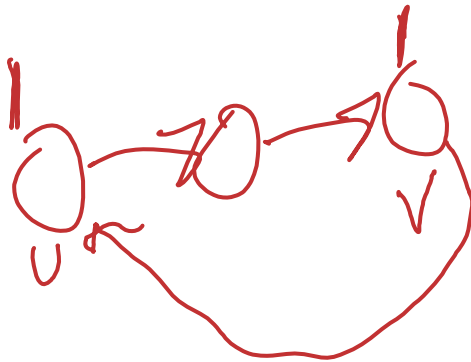
Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**
 - **Directed:** Each edge $(u, v) \in E$ now has a direction $u \rightarrow v$
 - **Acyclic:** No cycles.
- **Path:** A sequence of distinct vertices where each pair of consecutive vertices have an edge
- **Cycle:** A sequence of distinct vertices where each pair of consecutive vertices have an edge **and** the first and last vertices are connected.
- **Connected:** $u, v \in V$ are connected \iff there exists a path between u and v .



Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**
 - **Directed:** Each edge $(u, v) \in E$ now has a direction $u \rightarrow v$
 - **Acyclic:** No cycles.
- **Path:** A sequence of distinct vertices where each pair of consecutive vertices have an edge
- **Cycle:** A sequence of distinct vertices where each pair of consecutive vertices have an edge **and** the first and last vertices are connected.
- **Connected:** $u, v \in V$ are connected \iff there exists a path between u and v .
- **Strongly Connected:** $u, v \in V$ are strongly connected \iff there exists a path between u and v and from v to u .



Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**
 - **Directed:** Each edge $(u, v) \in E$ now has a direction $u \rightarrow v$
 - **Acyclic:** No cycles.
- **Path:** A sequence of distinct vertices where each pair of consecutive vertices have an edge
- **Cycle:** A sequence of distinct vertices where each pair of consecutive vertices have an edge **and** the first and last vertices are connected.
- **Connected:** $u, v \in V$ are connected \iff there exists a path between u and v .
- **Strongly Connected:** $u, v \in V$ are strongly connected \iff there exists a path between u and v and from v to u .
- **Connected Component (of u):** The set of all vertices connected to u .



Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**
 - **Directed:** Each edge $(u, v) \in E$ now has a direction $u \rightarrow v$
 - **Acyclic:** No cycles.
- **Path:** A sequence of distinct vertices where each pair of consecutive vertices have an edge
- **Cycle:** A sequence of distinct vertices where each pair of consecutive vertices have an edge **and** the first and last vertices are connected.
- **Connected:** $u, v \in V$ are connected \iff there exists a path between u and v .
- **Strongly Connected:** $u, v \in V$ are strongly connected \iff there exists a path between u and v and from v to u .
- **Connected Component (of u):** The set of all vertices connected to u .
- **Strongly Connected Component:** A set of vertices a strongly connected component if each pair of vertices are strongly connected.



Graphs

- **Definition:** A set of vertices V connected by a set of edges E . Individual edges are notated as (u, v) , where $u, v \in V$.
 - They are usually represented as **adjacency lists** or **adjacency matrices**
 - **Directed:** Each edge $(u, v) \in E$ now has a direction $u \rightarrow v$
 - **Acyclic:** No cycles.
- **Path:** A sequence of distinct vertices where each pair of consecutive vertices have an edge
- **Cycle:** A sequence of distinct vertices where each pair of consecutive vertices have an edge **and** the first and last vertices are connected.
- **Connected:** $u, v \in V$ are connected \iff there exists a path between u and v .
- **Strongly Connected:** $u, v \in V$ are strongly connected \iff there exists a path between u and v and from v to u .
- **Connected Component (of u):** The set of all vertices connected to u .
- **Strongly Connected Component:** A set of vertices a strongly connected component if each pair of vertices are strongly connected.

Graph Algorithms: Traversal

Graph Algorithms: Traversal

- **BFS:**

- **Purpose:** Reachability, Shortest Path (unweighted graph)
- **Implementation details:** Add your neighbours to a **queue**, pop from the queue to get next node
- **Runtime:** $O(V + E)$

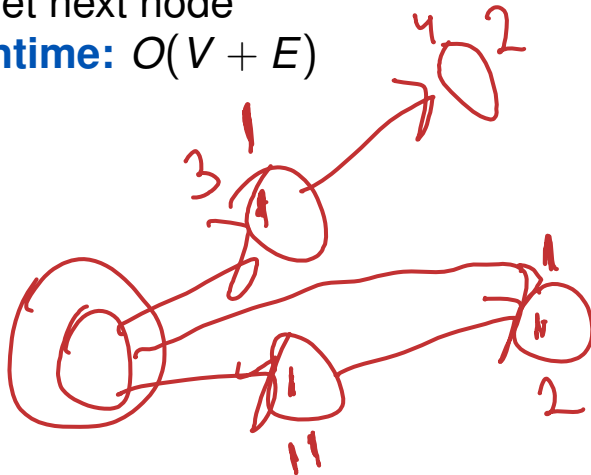
Graph Algorithms: Traversal

- **BFS:**

- **Purpose:** Reachability, Shortest Path (unweighted graph)
- **Implementation details:** Add your neighbours to a **queue**, pop from the queue to get next node
- **Runtime:** $O(V + E)$

- **DFS:**

- **Purpose:** Reachability, toposort
- **Implementation details:** Add your neighbours to a **stack**, pop from the stack to get next node
- **Runtime:** $O(V + E)$



Graph Algorithms: Shortest Path

- **Dijkstra's**

- **Purpose:** SSSP, no negative edges
- **Implementation:** Visit neighbours in **priority queue**
- **Runtime:** $O(m \log n)$ (with **Quake Heaps**, $O(m + n \log n)$)

$$|E| = m$$

$$|V| = n$$

Graph Algorithms: Shortest Path

- **Dijkstra's**

- **Purpose:** SSSP, no negative edges
- **Implementation:** Visit neighbours in **priority queue**
- **Runtime:** $O(m \log n)$ (with **Quake Heaps**, $O(m + n \log n)$)

- **Bellman-Ford:**

- **Purpose:** SSSP, yes negative weights. Will detect negative cycles.
- **Implementation:** Dynamic Programming recurrence:
 - ▶ $d(v, k)$ is the shortest-walk distance from s to v using at most k edges
 - ▶ $d(v, k) = \min \left(d(v, k - 1), \min_{u \rightarrow v} d(u, k - 1) + \ell(u \rightarrow v) \right)$
- **Runtime:** $O(mn)$

Graph Algorithms: Shortest Path

- **Dijkstra's**

- **Purpose:** SSSP, no negative edges
- **Implementation:** Visit neighbours in **priority queue**
- **Runtime:** $O(m \log n)$ (with **Quake Heaps**, $O(m + n \log n)$)

- **Bellman-Ford:**

- **Purpose:** SSSP, yes negative weights. Will detect negative cycles.
- **Implementation:** Dynamic Programming recurrence:
 - ▶ $d(v, k)$ is the shortest-walk distance from s to v using at most k edges
 - ▶ $d(v, k) = \min \left(d(v, k - 1), \min_{u \rightarrow v} d(u, k - 1) + \ell(u \rightarrow v) \right)$
- **Runtime:** $O(mn)$

- **Floyd-Warshall:**

- **Purpose:** APSP, yes negative edge weights
- **Implementation:** Dynamic Programming recurrence:
 - ▶ $d(u, v, i)$ is the shortest-path distance from u to v only going through vertices $1 \dots i$.
 - ▶ $d(u, v, i) = \min (d(u, k, k - 1), d(k, v, k - 1) +)$
- **Runtime:** $O(n^3)$

Graph Algorithms: MSTs

3 main algorithms:

- **Prim-Jarnik**: Keep a priority queue for edges to be added to the tree. Start the tree at some arbitrarily selected root vertex. When adding a vertex, add all of its neighbors to the queue. Runtime: $O(|E| \log |V|)$, $O(|V| \log |V| + |E|)$ using Quake heaps.

Graph Algorithms: MSTs

3 main algorithms:

- **Prim-Jarnik**: Keep a priority queue for edges to be added to the tree. Start the tree at some arbitrarily selected root vertex. When adding a vertex, add all of its neighbors to the queue. Runtime: $O(|E| \log |V|)$, $O(|V| \log |V| + |E|)$ using Quake heaps.
- **Kruskal**: Keep a disjoint-sets data structure to keep track of connected components. Sort the edges, then in order, add each edge if it connects two components. Runtime: $O(|E| \log |V|)$.

Graph Algorithms: MSTs

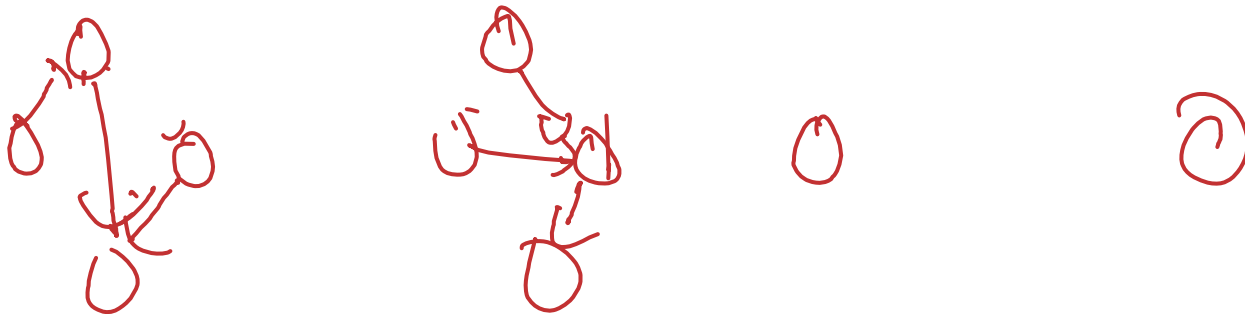
3 main algorithms:

- **Prim-Jarnik**: Keep a priority queue for edges to be added to the tree. Start the tree at some arbitrarily selected root vertex. When adding a vertex, add all of its neighbors to the queue. Runtime: $O(|E| \log |V|)$, $O(|V| \log |V| + |E|)$ using Quake heaps.
- **Kruskal**: Keep a disjoint-sets data structure to keep track of connected components. Sort the edges, then in order, add each edge if it connects two components. Runtime: $O(|E| \log |V|)$.
- **Borůvka**: No fancy data structures! Just find smallest edge going out of each vertex, then contract all edges that you selected! Runtime: $O(|E| \log |V|)$

Graph Algorithms: MSTs

3 main algorithms:

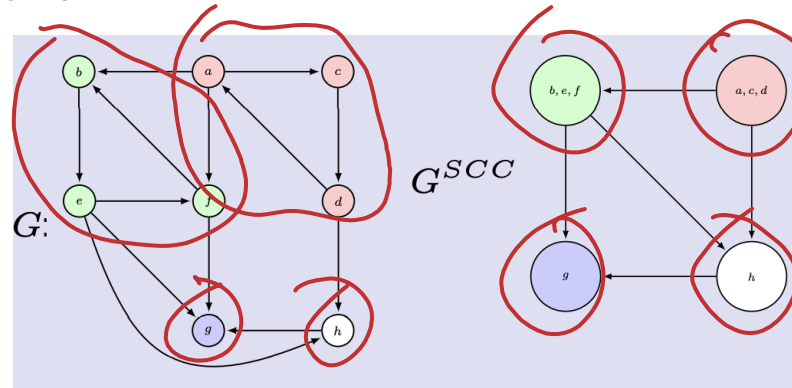
- **Prim-Jarnik**: Keep a priority queue for edges to be added to the tree. Start the tree at some arbitrarily selected root vertex. When adding a vertex, add all of its neighbors to the queue. Runtime: $O(|E| \log |V|)$, $O(|V| \log |V| + |E|)$ using Quake heaps.
- **Kruskal**: Keep a disjoint-sets data structure to keep track of connected components. Sort the edges, then in order, add each edge if it connects two components. Runtime: $O(|E| \log |V|)$.
- **Borůvka**: No fancy data structures! Just find smallest edge going out of each vertex, then contract all edges that you selected! Runtime: $O(|E| \log |V|)$
- Faster (but way more complicated algorithms) exist. **Yao** (1975): $O(|E| \log \log |V|)$ with a modification of Borůvka's (using linear-time median selection). **Karger-Klein-Tarjan** (1995): $O(|E|)$ in expectation, **Chazelle** (2000): $O(|E| \alpha(|V|, |E|))$ deterministic



Graph Algorithms: SCC

SCC-Finding Algorithms (Tarjan's, Kosuraju's)

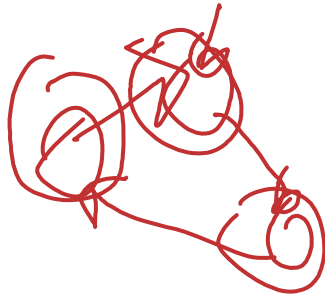
- **Purpose:** To identify (and collapse) SCCs in a (directed) graph
- **Runtime:** $O(V + E)$
- **Returns:** A metagraph that has one node for each SCC.



Graph Algorithms: Longest Path

Longest path in a Directed Acyclic Graph (DAG)

- **Purpose:** To find the longest simple path (no repeating vertices) by weight in a graph which is guaranteed to be a DAG¹.
- **Runtime**²: $O(V + E)$
- **Returns:** The sum of the weights of the longest path in the DAG.



¹Finding the longest path in other types of graphs is at least NP-hard.

²This is a relatively straight-forward DP on a DAG problem if you wish to derive it.

Graph Problems: General Stuff

How to solve graph problems:

1. Identify type of problem (Reachability, Shortest Path, SCC)
2. Construct new graph
 - Add sources/sinks
 - Add vertices via $V' = V \times \{\text{some set}\}$ (Useful for tracking states)
 - Add vertices via $E' = E \times \{\text{some set}\}$ (Useful for allowing/prohibit certain behaviour)
3. Apply some stock algorithm (**DO NOT MODIFY THE ALGORITHMS - MODIFY THE INPUTS!**)
4. Draw connection between how to result of the algorithm upon the new graph relates to the solution of the original question.

$G = (V, E)$ BFS \Rightarrow distances for all nodes

Traverse v , all $v \in V$ distance n , add to set.

$O(V+E)$ time,



Recurrences and Asymptotics

Give a tight asymptotic bound for each recurrence:

- $T(n) = 4 T(\frac{n}{2}) + n \log_2 n$

$n \log n$

$2n \log \frac{n}{2} \leftarrow \frac{n}{2} \log \frac{n}{2}$

$2^{(n(\log_2 n) - 1)}$

at layer l , $\log_2 \frac{n}{2^l} \log_2 \frac{n}{2^l} \rightarrow 2^l n (\log_2 n - l)$

$l = \log_2 n$

$2^{\log_2 n - 1} n (\log_2 n - 1) = \boxed{n^2}$

Recurrences and Asymptotics

Give a tight asymptotic bound for each recurrence:

- $T(n) = 4 T\left(\frac{n}{2}\right) + n \log_2 n$
- $T(n) = T\left(\frac{3n}{4}\right) + T\left(\frac{n}{4}\right) + 5n$

Recurrences and Asymptotics

Give a tight asymptotic bound for each recurrence:

- $T(n) = 4 T\left(\frac{n}{2}\right) + n \log_2 n$
- $T(n) = T\left(\frac{3n}{4}\right) + T\left(\frac{n}{4}\right) + 5n$
- $T(n) = 9 T\left(\frac{n}{3}\right) + n^2$

Recurrences and Asymptotics

Give a tight asymptotic bound for each recurrence:

- $T(n) = 4 T(\frac{n}{2}) + n \log_2 n$
- $T(n) = T(\frac{3n}{4}) + T(\frac{n}{4}) + 5n$
- $T(n) = 9 T(\frac{n}{3}) + n^2$

Group the following functions s.t. f and g are in the same group if $f(x) \sim \Theta(g(x))$, and sort the groups by runtime:

- n^n
- $\log \log n$
- 374^n
- $n!$
- $\log(n + n^{374})$
- $\log n^n$
- $n^{1.000001}$
- 2^n
- $n \log n^5$
- $\frac{1}{\log_n 2}$

Divide and Conquer

Consider the following (correct!) in-place sorting algorithm:

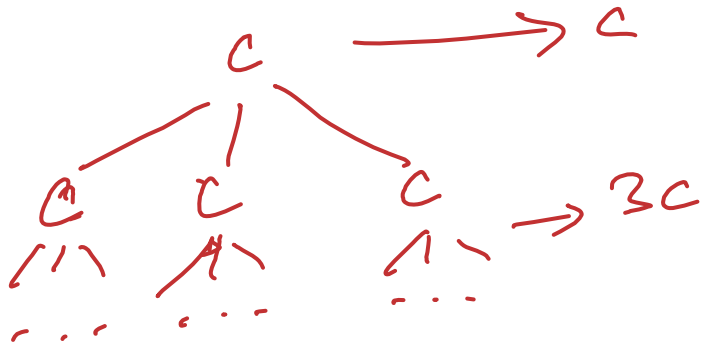
```

1: procedure STOOGESORT( $A[1 \dots n]$ )
2:   If  $A[1] > A[n]$ , swap them.
3:   if  $n \geq 3$  then
4:     STOOGESORT the initial  $2/3$  of  $A$ 
5:     STOOGESORT the final  $2/3$  of  $A$ 
6:     STOOGESORT the initial  $2/3$  of  $A$  (again)
  
```

$$T(n) = O(1) + 3T\left(\frac{2}{3}n\right)$$

$\frac{n}{1.5}$

Give a *tight* asymptotic bound on the runtime of StoogeSort, in terms of n .



$$\log_{3/2} n = \log_{1.5} n$$

levels

$$3^{\log_{1.5} n} = n$$

$$O(n^{\log_{1.5} 3})$$

Dynamic Programming

You are given an array of integers $A[1..2n]$. On each turn, you choose two arbitrary integers with no other numbered integers between them. You then earn points equal to the division of your two chosen numbers, and both chosen integers are removed. Once a square is removed, it cannot be chosen in any future turns. Your goal is to remove all of the numbers and to earn as many points as possible. Describe an algorithm to find the maximum number of points you can earn in a game.

4	3	6	5	9	1	2	3	7	3	4	8
4	3	6	5	9	1	2	3			4	8
4	3	6	5			2	3			4	8
4	3	6					3			4	8
4	3									4	8
										4	8

+ $\frac{7}{3}$ points!

+ $\frac{9}{1}$ points!

+ $\frac{5}{2}$ points!

+ $\frac{6}{3}$ points!

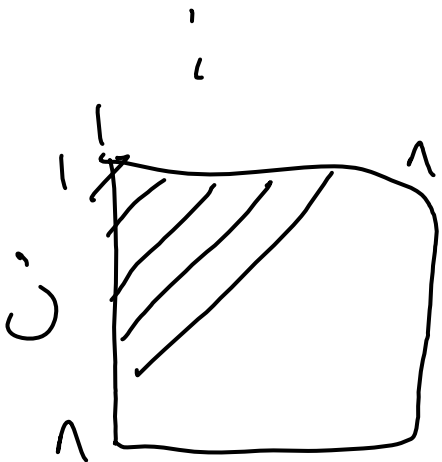
+ $\frac{4}{3}$ points!

+ $\frac{4}{8}$ points!

All done!

Dynamic Programming $i \% 2 = j \% 2 \rightarrow -\infty$

$$\text{MaxScore}(i, j) = \begin{cases} 0 & i \geq j \\ A[i] / A[j] & \text{if } i+1 = j \\ \max_{\text{for all } i \leq a < b \leq j} \{ A[a] / A[b] + \text{MaxScore}[i, a-1] + \text{MaxScore}[a+1, b-1] + \text{MS}[b+1, j] \} \end{cases}$$



$O(n^2)$

$O(n^2)$

$O(n^4)$

for $i=2 \rightarrow n$
for $j=i$

Dynamic Programming

This semester, Ryan, a brilliant all-knowing algorithms student, learned about the lovely social hours ACM puts on even Friday, providing free food to all that come. Unfortunately, Seth Jerickson, the almighty evil professor, has scheduled a 100 point exam every Saturday! This is only worsened by the fact that every time that Ryan goes to an ACM social hour, he falls into a deep food coma. This causes him to lose so many hours of studying and brain cells that his max score on this exam and every following exam that he takes after attending a social hour decreases by $P[i]$, where i is the current week number. Ryan can only lose up to M points in the semester before he fails the course. Alternatively, if Ryan does not go to ACM's social hour for that week, he will have enough time to study to obtain the maximum score he is able to achieve. Now, studying for 374 comes at the expensive of his happiness levels, and thus, he would like to attend as many social hours as possible. For each social hour he attends on week i , he will receive $H[i]$ amount of happiness. Furthermore, for each consecutive social hour he attends, there is a cumulative 2x multiplier on the happiness he gains.

For example, suppose that there are 10 weeks in the semester, and Ryan attends social hour on weeks 2, 4, and 5 (0-indexed). Let $H = [5, 9, 2, 4, 7, 2, 8, 0, 4, 2]$ and let $P = [9, 4, 2, 8, 6, 4, 3, 2, 7, 9]$. The amount of happiness Ryan achieves is $0 + 0 + 2 + 0 + 7 + 2 \times 2 + 0 + 0 + 0 + 0$. His exam scores for the 10 exams would be 100, 100, 98, 100, 92, 88, 100, 100, 100, 100. Given N weeks of the semester, where each week i has:

- A happiness value $H[i]$ gained for attending that week's social
- A penalty value $P[i]$ subtracted from his current maximum score

Describe and analyze an algorithm that determines the maximum happiness Ryan can achieve while still passing the exam.

Dynamic Programming

Graphs

A QSP - $O(V^3)$

SSSP from all MTS ($O(E \log V - |M|)$)

You are in Austin, TX, which is lit at night by "moon towers", high powerful lamps which illuminate multiple blocks. You are given a weighted undirected graph $G = (V, E)$ as a map of Austin, as well as a list of moon towers represented as pairs (p, v) where $v \in V$ is the location of the tower and $p \in \mathbb{R}_+$ is the power of the moon tower.

- (a) A vertex v is considered *lit* if there exists a moontower m s.t. $d(m, v) < p(m)$. Describe and analyze an algorithm to find all *lit* vertices.

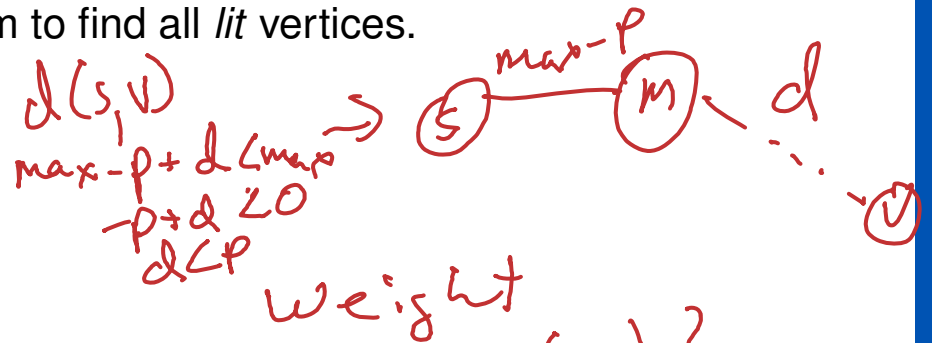
$$G' = (V', E')$$

$$V' = V \cup \{s\}$$

$$E' = E \cup \left\{ s \rightarrow m, \text{max MT} - p(m) \right\}$$

Dijkstra, check if $d(s, v) < \text{max MT}$

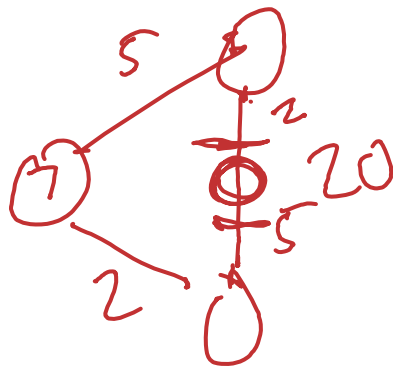
$$O(E \log V)$$



Graphs

You are in Austin, TX, which is lit at night by “moon towers”, high powerful lamps which illuminate multiple blocks. You are given a weighted undirected graph $G = (V, E)$ as a map of Austin, as well as a list of moon towers represented as pairs (p, v) where $v \in V$ is the location of the tower and $p \in \mathbb{R}_+$ is the *power* of the moon tower.

- (a) A vertex v is considered *lit* if there exists a moontower m s.t. $d(m, v) < p(m)$. Describe and analyze an algorithm to find all *lit* vertices.
- (b) An edge $u \leftrightarrow v$ is *lit* if for all $0 \leq h \leq \ell(u \leftrightarrow v)$, there exists a moontower m s.t. either $d(m, u) + h < p(m)$ or $d(m, v) + \ell(u \leftrightarrow v) - h < p(m)$. Describe and analyze an algorithm to find all *lit* edges.



for $u \leftrightarrow v \in E$
 if $\max_{MT} p - d(s, u)$
 $\cdot \max_{MT} p - d(s, v)$
 $< \ell(u, v)$:
 good!
 else:
 bad!

Graphs

You are in Austin, TX, which is lit at night by “moon towers”, high powerful lamps which illuminate multiple blocks. You are given a weighted undirected graph $G = (V, E)$ as a map of Austin, as well as a list of moon towers represented as pairs (p, v) where $v \in V$ is the location of the tower and $p \in \mathbb{R}_+$ is the *power* of the moon tower.

- (a) A vertex v is considered *lit* if there exists a moontower m s.t. $d(m, v) < p(m)$. Describe and analyze an algorithm to find all *lit* vertices.
- (b) An edge $u \leftrightarrow v$ is *lit* if for all $0 \leq h \leq \ell(u \leftrightarrow v)$, there exists a moontower m s.t. either $d(m, u) + h < p(m)$ or $d(m, v) + \ell(u \leftrightarrow v) - h < p(m)$. Describe and analyze an algorithm to find all *lit* edges.
- (c) Using your answers to parts (a) and (b), describe and analyze an algorithm to calculate the shortest $s \rightarrow t$ path where at most k edges are unlit.

$$G' = (V', E')$$

$$V' = V \times \{0 \dots k\}$$

$$E' = \left\{ \begin{array}{l} (u, i) \rightarrow (v, i) \text{ if } u \rightarrow v \text{ lit} \\ (u, i) \rightarrow (u, i+1) \text{ if } u \rightarrow v \text{ unlit} \end{array} \right.$$

Dijkstra's

$$O(E' \log V')$$

$$\begin{array}{l} V' = kV \\ E' = kE \end{array}$$

$$\text{WTF } \min_d((s, 0), (t, i))$$

$$O(kE \log kV)$$

Graphs

- (a) A train station is considered *critical* if its breaking down would result in one city in the network no longer being able to reach another city in the network by any path. Describe and analyze an algorithm to, given a train station t and a network G , determine if t is a *critical* train station,

Graphs

- (a) A train station is considered *critical* if its breaking down would result in one city in the network no longer being able to reach another city in the network by any path. Describe and analyze an algorithm to, given a train station t and a network G , determine if t is a *critical* train station,
- (b) Describe and analyze an algorithm to find all *critical* train stations.

Feedback

- Further questions in ACM (Siebel 1104) after this
- Please fill out the feedback form:
`go.acm.illinois.edu/cs374a_mt2_feedback`

