CSE 421 Section 3

Problem solving with greedy algorithms

Administrivia

Announcements & Reminders

- **HW1**
  - Regrade requests are open
  - Answer keys available on Ed

- **HW2**
  - Was due yesterday, 10/9
  - Remember the **late problems** policy (NOT assignments)
    - Total of up to **10 late problem days**
    - At most **2 late days per problem**

- **HW3**
  - Due Wednesday 10/16 @ 11:59pm

How to write an algorithm

Problem solving strategy overview

Read and **summarize** the problem

Decide to use **known algorithm** or **techniques from scratch**

not covered this section

no idea

have idea

Solve examples to get ideas

Check that idea **isn't easily falsified** or **slow**

Write pseudocode, proof, and running time analysis

Getting started

Getting started

Read and **summarize** the problem

↓

Decide to use **known algorithm** or **techniques from scratch**

not covered this section

no idea

have idea

Solve examples to get ideas

Check that idea **isn't easily falsified** or **slow**

Write pseudocode, proof, and running time analysis

Problem summary

When reading a long word problem, it is useful to **summarize** it. A common way is:

Input: …
Expected output: …

- mathematical definitions of any special words used above

Problem summary

When reading a long word problem, it is useful to **summarize** it. A common way is:

Example

Input: Two sets P and R of n people each, with preference lists

Expected output: A stable matching

- **preference list:** an ordered list of people in the other set
- **stable matching:** a perfect matching for which there is no (p, r) where p and r prefer each other over their current match

Problem 1 – Line covering

Your new towing company wants to be prepared to help along the highway during the next snowstorm. You have a list of integers $t_1, t_2, \dots, t_n$ in increasing order, representing mile markers on the highway where you think it is likely someone will need a tow (entrances/exits, merges, rest stops, etc.). To ensure you can help quickly, you want to place your tow trucks so that from every marker, at least one truck is at most 3 miles away. Find a minimum length list of sites where you can place tow trucks to satisfy the requirement, written as a list of integers $a_1, a_2, \dots, a_m$ in increasing order. Note that the sites that you pick need not be a subset of the marked locations.

a) Write a summary of the problem.

Feel free to work with the people around you!

Solution

Problem 1 – Line covering

a) Write a summary of the problem.

Input: A list of increasing integers $t_1, t_2, \dots, t_n$

Expected output: A shortest list of increasing integers $a_1, \dots, a_m$ covering the input

- **cover:** for all $i \in \{1, \dots, n\}$, there exists $j \in \{1, \dots m\}$ such that $|t_i - a_j| \leq 3$.

Reduction vs. techniques from scratch

Read and **summarize** the problem

Known algorithms

- Stable matching
- Graph algorithms
- …etc.
- Network
- Linear pr

Techniques from scratch

- Greedy algorithms
- Divide and conquer (week 4)
- Dynamic programming (week 5)

None of these seem right for today's problem,
so we'll try a greedy algorithm!

Does the problem remind me of an algorithm I've seen in class?

Generating ideas

Generating ideas

Read and **summarize** the problem

↓

Decide to use **known algorithm** or **techniques from scratch**

not covered this section

no idea

have idea

Solve examples to get ideas ⇄ Check that idea **isn't easily falsified** or **slow**

↓

Write pseudocode, proof, and running time analysis

Generating ideas

If using **techniques** from scratch:
(for today, greedy algorithms only)

no idea

have idea

Solve many examples by hand

- In the beginning, don't worry about general strategy
- Think about what **patterns** appear
- If your brain is magically solving small examples, try bigger ones

idea! →

← **yes :(**

Ask yourself questions

- Can I break my strategy with a nasty example?
- Does my strategy ever waste time? Can I optimize it?

↓ **all good!**

Ideas for greedy algorithms

- What's a **greedy algorithm**?
  - Follows a rule to keep picking something
  - Doesn't consider the future
  - Doesn't go back to fix things

- Coming up with many greedy ideas should be easy. Finding the correct greedy idea will usually require trial and error or insight.

Problem 1 – Line covering

b) We will practice generating ideas.
   i. Solve these by hand. Don't worry too much about greedy strategies yet.

1, 2, 4, 10, 12

0, 1, 3, 5, 7, 8, 13, 14

Feel free to work with the people around you!

Solution

Problem 1 – Line covering

b) We will practice generating ideas.
   i. Solve these by hand. Don't worry too much about greedy strategies yet.

1, 2, 4, 10, 12

2 trucks. Many solutions, for example at 2 and 11.

0, 1, 3, 5, 7, 8, 13, 14

3 trucks. Many solutions, for example at 0, 7, and 13.

Problem 1 – Line covering

1, 2, 4, 10, 12

0, 1, 3, 5, 7, 8, 13, 14

ii. Suppose you came up with the greedy idea:

"Put a truck on the first uncovered marker."

Check that this idea works on the above examples. Then, try to break this idea by coming up with an example where it doesn't work.

Feel free to work with the people around you!

Solution

Problem 1 – Line covering

1, 2, 4, 10, 12

0, 1, 3, 5, 7, 8, 13, 14

ii. Suppose you came up with the greedy idea:

"Put a truck on the first uncovered marker."

Check that this idea works on the above examples. Then, try to break this idea by coming up with an example where it doesn't work.

0, 6 can be covered by one truck at 3, this method gives two trucks.

(many other examples)

Problem 1 – Line covering

iii. Come up with a new greedy idea that solves your new example. Does the idea work? If not, continue the process until you have a working idea.

Feel free to work with the people around you!

Solution

Problem 1 – Line covering

iii. Come up with a new greedy idea that solves your new example. Does the idea work? If not, continue the process until you have a working idea.

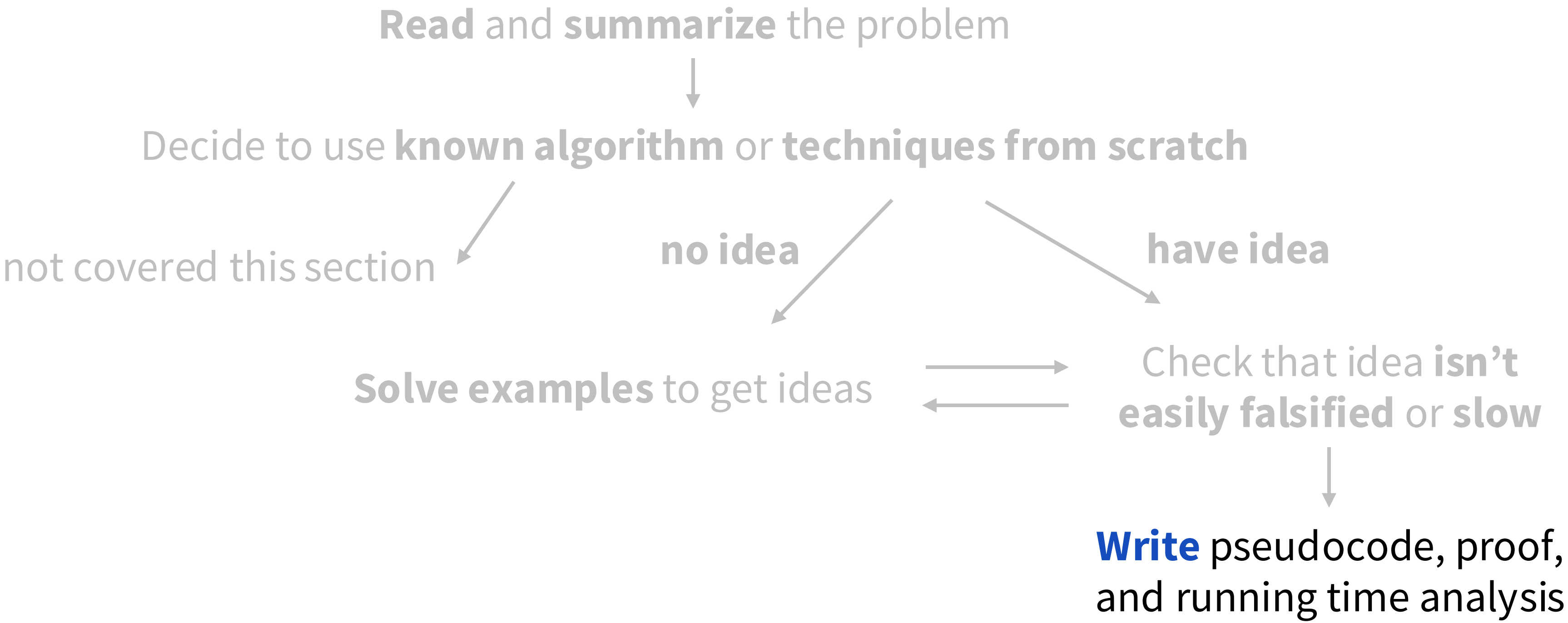
Sample final idea:

Place a truck at the farthest location that still covers the next uncovered marker.

That is, if t_i is the next uncovered marker, place a truck at $t_i + 3$.

Writing up your idea

Writing up your idea



Writing up your idea

Once you have an **efficient, working idea**:

1. Translate it into **pseudocode**.
   - More precise than English, but easier to understand than code.
   - No hard rules, but see handout from last week for common styles.
2. **Prove** the pseudocode correct.
   - We'll cover greedy-specific tips today!
3. Write up the **running time** analysis.

Problem 1 – Line covering

c) Write the pseudocode for the solution.

Previous Solution

Problem 1 – Line covering

c) Write the pseudocode for the solution.

Input: A list of increasing integers $t_1, t_2, \ldots, t_n$
Expected output: A shortest list of increasing integers $a_1, \ldots, a_m$ covering the input

Idea: Place a truck at the farthest location that still covers the next uncovered marker. That is, if t_i is the next uncovered marker, place a truck at $t_i + 3$.

Feel free to work with the people around you!

Solution

Problem 1 – Line covering

c) Write the pseudocode for the solution.

Input: A list of increasing integers $t_1, t_2, \dots, t_n$

Expected output: A shortest list of increasing integers $a_1, \dots, a_m$ covering the input

1. Let $i = 1$ and $j = 1$.
2. While $i \leq n$,
    a. Let $a_j = t_i + 3$.
    b. Repeatedly increment i until $t_i > a_j + 3$ (or $i > n$).
    c. Increment j.
3. Return the list of all a_j.

Solution

Problem 1 – Line covering

c) Write the pseudocode for the solution.

Input: A list of increasing integers $t_1, t_2, \ldots, t_n$

Expected output: A shortest list of increasing integers $a_1, \ldots, a_m$ covering the input

1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

Extra tip: Avoid changing values (excluding the iteration counter) whenever you can do so without increasing big-O runtime. This way, proofs are easier to write:

"i at the start of iteration j" → "i_j"

"i at the end of iteration j" → "i_{j+1}"

Algorithm proofs refresher

- As always, prove **validity**, **termination**, and **correctness**.

- Correctness always means:
  "My algorithm's output matches the problem summary's expected output."

- For greedy algorithms, correctness means "My output is an **optimal solution**."
  In other words, two things to prove:
  - "Output is a valid solution."
    - "The list $a_1, \ldots, a_m$ is in increasing order and covers all markers."
  - "Output is optimal."
    - "All other valid solutions use at least m trucks."

Algorithm proofs refresher

For optimality, there are some common strategies:

- **“Greedy stays ahead”**: For all other solutions, show by induction that at every step, your solution is at least as good.

- **“Exchange argument”**: For all other solutions that differ from yours, show how to replace a part of the other solution, so that the quality improves or stays the same (but never decreases).

- **“Structural argument”**: (less common) Find a “hard subset” of the input that immediately implies why other solutions must also be as bad as yours (or worse).

Problem 1 – Line covering

d) Write a proof that your pseudocode is correct.

Each line is valid:

Termination:

“The output is in increasing order.”:

“The output covers all markers.”:

Previous Solution

Problem 1 – Line covering

d) Write a proof that your pseudocode is correct.

Each line is valid:

Termination:

"The output is in increasing order.":

"The output covers all markers.":

1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

Focus on these easier parts first, and feel free to work with the people around you!

Solution

Problem 1 – Line covering

d) Write a proof that your pseudocode is correct.

Each line is valid:
Evident.

1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

Solution

Problem 1 – Line covering

d) Write a proof that your pseudocode is correct.

Termination:

We have $t_{i_j} = a_j - 3 < a_j + 3 < t_{i_{j+1}}$, thus $i_j \neq i_{j+1}$, so i increases every iteration and there are at most n iterations. Line 2b's inline "repeat" ends in at most n iterations as well, since we stop if $i_{j+1} > n$.

1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

Solution

Problem 1 – Line covering

d) Write a proof that your pseudocode is correct.

1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

"The output is in increasing order.":

Again, $t_{i_j} < t_{i_{j+1}}$, thus we conclude that $a_j = t_{i_j} + 3 < t_{i_{j+1}} + 3 = a_{j+1}$.

Solution

Problem 1 – Line covering

d) Write a proof that your pseudocode is correct.

"The output covers all markers.":

We increment i_{j+1} to $i_{j+1} + 1$ if and only if $t_{i_{j+1}}$ is covered by a_j. Since we exit the loop when $i_j > n$, every marker is covered.

1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

Problem 1 – Line covering

d) Write a proof that your pseudocode is correct.

Now for the harder part. For this section, try to write a "greedy stays ahead" proof!

"All other valid solutions use at least m trucks."

i. What is the "greedy stays ahead" claim?

Previous Solution

Problem 1 – Line covering

d) Write a proof that your pseudocode is correct.

Now for the harder part. For this section, try to write a "greedy stays ahead" proof!

"All other valid solutions use at least m trucks."

i. What is the "greedy stays ahead" claim?

1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

Feel free to work with the people around you!

Solution

Problem 1 – Line covering

i. What the "greedy stays ahead" claim?

Let $o_1, \dots, o_M$ be any other valid solution. We will show for all j:

$P(j) =$ "Sites $a_1, \dots, a_j$ cover all t_i that are covered by $o_1, \dots, o_j$ (and possibly more)."

There are actually many possible "greedy stays ahead" claims. Another option is:

$P(j) =$ "Sites $a_1, \dots, a_j$ covers at least as many t_i as $o_1, \dots, o_j$ covers."

The one we chose will be a bit natural to prove, since it describes the situation a bit more exactly.

Problem 1 – Line covering

ii. Prove the “greedy stays ahead” claim using induction.

Feel free to work with
the people around you!

Previous Solution

Problem 1 – Line covering

ii. Prove the "greedy stays ahead" claim using induction.

$P(j) =$ "Sites $a_1, \dots, a_j$ cover all t_i that are covered by $o_1, \dots, o_j$ (and possibly more)."

Feel free to work with the people around you!

1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

Solution

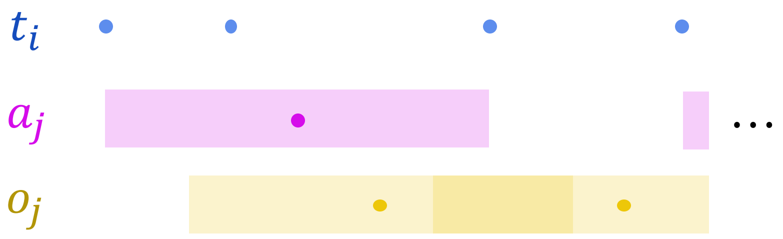
Problem 1 – Line covering

ii. Prove the "greedy stays ahead" claim using induction.

$P(j)$ = "Sites $a_1, \dots, a_j$ cover all t_i that are covered by $o_1, \dots, o_j$ (and possibly more)."

Base case: We will show $P(1)$, that a_1 covers all t_i that are covered by o_1.

1. We set $a_1 = t_1 + 3$.
2. If $o_1 > a_1$, then o_1 does not cover t_1, and neither do $o_2, \dots, o_M > o_1$, contradiction.



Solution

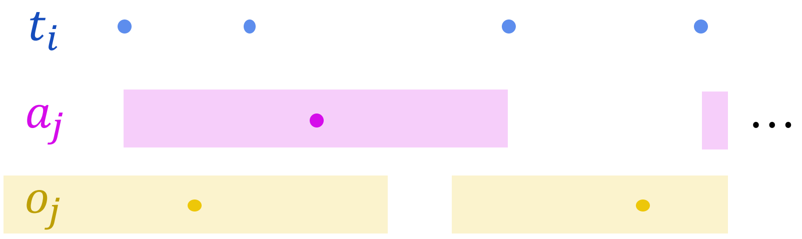
Problem 1 – Line covering

ii. Prove the "greedy stays ahead" claim using induction.

$P(j)$ = "Sites $a_1, \ldots, a_j$ cover all t_i that are covered by $o_1, \ldots, o_j$ (and possibly more)."

Base case: We will show $P(1)$, that a_1 covers all t_i that are covered by o_1.

1. We set $a_1 = t_1 + 3$.
2. If $o_1 > a_1$, then o_1 does not cover t_1, and neither do $o_2, \ldots, o_M > o_1$, contradiction.
3. If $o_1 \leq a_1$, since t_1 is the smallest marker, a_1 covers everything that o_1 covers.



Solution

Problem 1 – Line covering

$P(j) =$ "Sites $a_1, \ldots, a_j$ cover all t_i that are covered by $o_1, \ldots, o_j$ (and possibly more)."

Inductive hypothesis: Suppose that $P(j)$ holds for all $j \leq k$.

Inductive step: We will show $P(k+1)$.

1. Note that for all j, t_{i_j} is the smallest marker not covered by $a_1, \ldots, a_{j-1}$. (This is a loop invariant, formally prove it by induction).
2. So when $j = k+1$, $t_{i_{k+1}}$ is not covered by $a_1, \ldots, a_k$, and nor by $o_1, \ldots, o_k$ by IH.

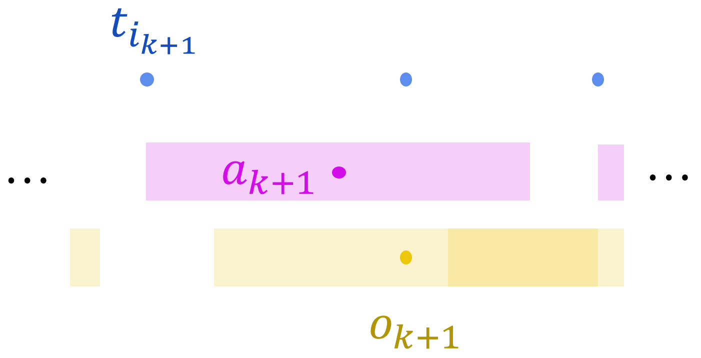
1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

Solution

Problem 1 – Line covering

$P(j) =$ "Sites $a_1, \dots, a_j$ cover all t_i that are covered by $o_1, \dots, o_j$ (and possibly more)."

3. We set $a_{k+1} = t_{i_{k+1}} + 3$.
4. If $o_{k+1} > a_{k+1}$, sites $o_1, \dots, o_k$ don't cover $t_{i_{k+1}}$ by what we just said, and nor do $o_{k+1}, \dots, o_M > t_{i_{k+1}} + 3$, contradiction.

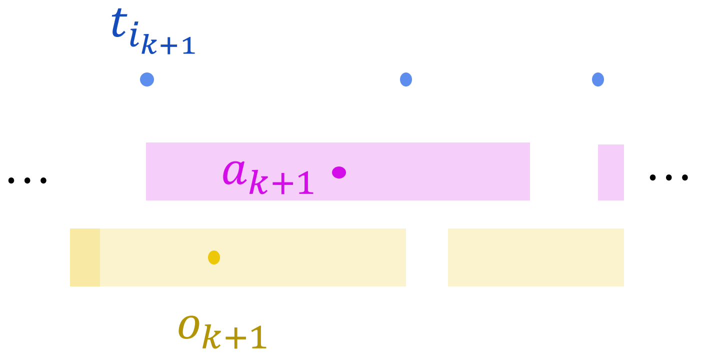


Solution

Problem 1 – Line covering

$P(j) =$ "Sites $a_1, \dots, a_j$ cover all t_i that are covered by $o_1, \dots, o_j$ (and possibly more)."

3. We set $a_{k+1} = t_{i_{k+1}} + 3$.
4. If $o_{k+1} > a_{k+1}$, sites $o_1, \dots, o_k$ don't cover $t_{i_{k+1}}$ by what we just said, and nor do $o_{k+1}, \dots, o_M > t_{i_{k+1}} + 3$, contradiction.
5. If $o_{k+1} \le a_{k+1}$, since $t_{i_{k+1}}$ is the smallest uncovered marker, a_{k+1} covers everything that o_{k+1} newly covers. Combined with IH, we get $P(k+1)$.



Problem 1 – Line covering

e) Analyze and prove the running time with big-O in a few sentences.

Previous Solution

Problem 1 – Line covering

e) Analyze and prove the running time with big-O in a few sentences.

1. Let $i_1 = 1$ and $j = 1$.
2. While $i_j \leq n$,
   a. Let $a_j = t_{i_j} + 3$.
   b. Let $i_{j+1} = i_j$, then repeatedly increment i_{j+1} until $t_{i_{j+1}} > a_j + 3$ (or $i_{j+1} > n$).
   c. Increment j.
3. Return the list of all a_j.

Feel free to work with the people around you!

Solution

Problem 1 – Line covering

e) Analyze and prove the running time with big-O in a few sentences.

Line 2b's inline repeat occurs n times across all iterations of the outer loop, and the rest of the outer loop takes constant time per iteration, for up to n iterations. Hence, the running time is $O(n)$.

Summary

Read and **summarize** the problem

Decide to use **known algorithm** or **techniques from scratch**

not covered this section

no idea

have idea

Solve examples to get ideas

Check that ideas are **correct** and **efficient**

Write pseudocode, proof, and running time analysis

Thanks for coming to section this week!