Mechanism Design

CMPUT 355: Games, Puzzles, and Algorithms

Lecture Outline

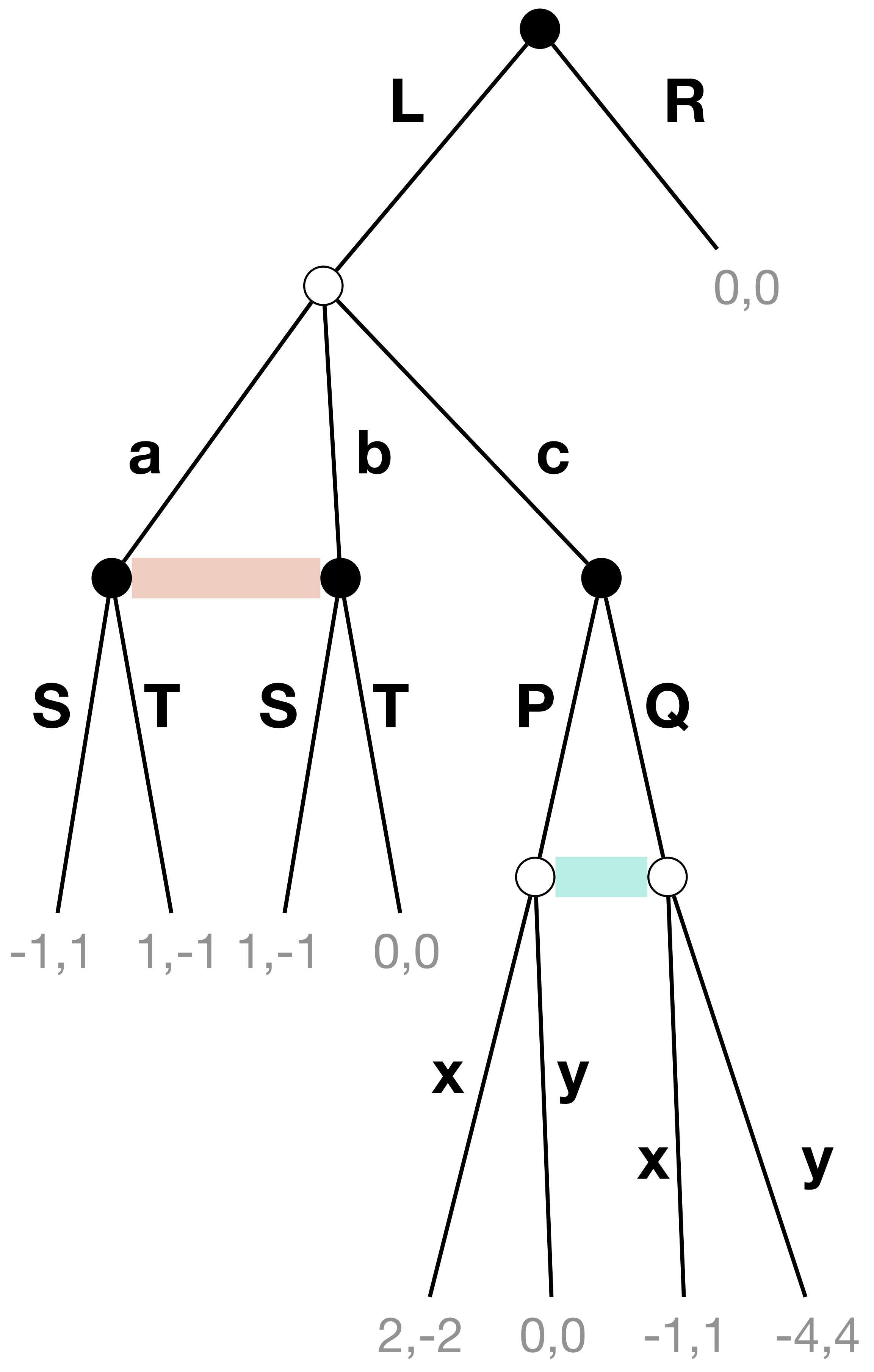
1. Logistics & Recap
2. Allocations Problem
3. Mechanism Design
4. The Revelation Principle

Final Exam Logistics

- **Final exam: Wed April 15** (one week from today)
  - **CCIS L2-190** (this room)
  - **1:00pm** (*not* this time)
  - Format: like a **quiz**, but longer! (3-4 times as long as a quiz)
  - Cumulative: covers the **whole semester**
  - *Very slight* emphasis on post-quiz-5 (maybe one extra question)
- **Practice material:**
  - No practice final (but: finals from past offerings are available here)
  - Practice questions #1-5 will be very relevant to the final
  - **Practice questions #6** released yesterday; solutions tomorrow (**Apr 9**)
- **Next lecture:** Review session

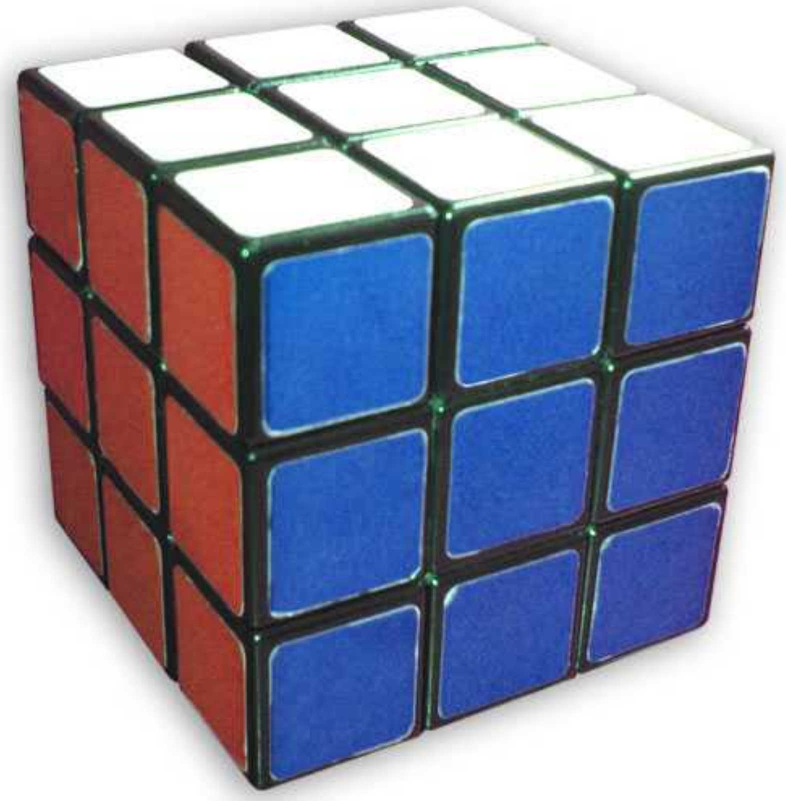
Recap: Imperfect Information Games

- **All** information is revealed in **perfect information games**; **no** information is revealed in **normal form games**
- Middle ground: **Imperfect information extensive form games**
- Players have **partial information** about which actions were played
  - **Information sets:** states that a player cannot distinguish
- **Pure strategies:** map from information set (not state) to action
  - **Mixed strategies:** randomly choose a **pure strategy** to play according to
  - **Behavioral strategies:** randomly choose an **action** at each **information set**
- **Kuhn's theorem:** **Mixed** and **behavioral** strategies equivalent in games of **perfect recall**



Scenario: Allocating a precious object

- **The Center** has a precious vintage object that it wishes to give to whoever values it the most.
  - For simplicity, assume there is no after-market etc.
- There are n agents, each of which is **economically rational** (**meaning what?**)
- Each agent has a **value** for the object that is known only to **themselves**
- **Question:** How should the **Center** decide who gets the object?
- **Question:** Can we represent this scenario using **game theory**? (**how?**)



Bayesian Games

- A **Bayesian game** is an imperfect information game in which:
  1. **Nature** assigns some private information (the **type**) to each agent
  2. Each agent takes some action **simultaneously**
  3. The utilities for an agent **depends on the types** (either their own, or all the types)
- In our allocation scenario, the standard setup is:
  1. **Nature** assigns each agent a **valuation** i.i.d. from a **known distribution**
  2. Each agent does something in a game designed by the **Center**
  3. The **Center** chooses who gets the object
  4. Every agent gets a utility of 0, except the agent who gets the object
  5. The agent who gets the object gets their value for the object
     (* possibly with some adjustments)

Mechanism Design

- **Mechanism design:** The Center **chooses a game** for the agents to play such that the Center's **goal** is guaranteed to be achieved in **equilibrium**
    - In our setting, Center's goal: give the object to agent with highest valuation
- The Center specifies:
    1. The **actions** that the agents can take
    2. What **outcomes** will occur for each profile of actions
- The Center does **not** specify the agents' **utilities** for the outcomes
- It **won't work** for the Center to just ask the agents for their valuations and then give the object to the agent that reports the highest valuation (**why?**)
- **Sealed-bid first-price auction:**
    - The Center asks everyone to report their value
    - The Center gives the item to the agent who reports the highest value, and charges them that value

Questions:

1. What are the actions in a sealed-bid auction?
2. What are the outcomes in a sealed-bid auction?
3. What should an agent i with valuation v_i **bid** in a sealed-bid first-price auction?

Second-Price Auctions

- It's not obvious **how to bid** in a **first-price sealed-bid auction**
  - There must be some equilibrium (**why?**)
  - It turns out the equilibrium is to **underreport** your value by a **specific amount**
  - If **other people** aren't playing the equilibrium, then maybe not a best response for **you** to play the equilibrium (**why?**)
- Alternative: **second-price auction**
  1. Everyone reports their valuation
  2. Bidder with highest reported valuation gets the item
  3. Bidder with highest reported valuation pays the second-highest reported valuation
- In a second-price auction, it is a **dominant strategy** to report your **true valuation**
  - (i) **Over-reporting is sub-optimal:** someone could bid higher than your value but lower than your bid
  - (ii) **Under-reporting is sub-optimal:** someone could bid lower than your value but higher than your bid
- **Dominant strategy truthful** auctions (or "**strategyproof** auctions") are **easy** to bid in
  - Reporting the truth is the right action *whether other people bid truthfully or not*

Revelation Principle

- **Not every goal** can be achieved by mechanism design (we'll see an example soon)
- But there are infinitely many games!
  How could you ever prove that **no game** can elicit the desired behaviour?

Theorem (Revelation Principle):
If there exists **any** mechanism which achieves a goal in equilibrium, then there exists a **truthful**, **direct** mechanism which achieves the same goal.

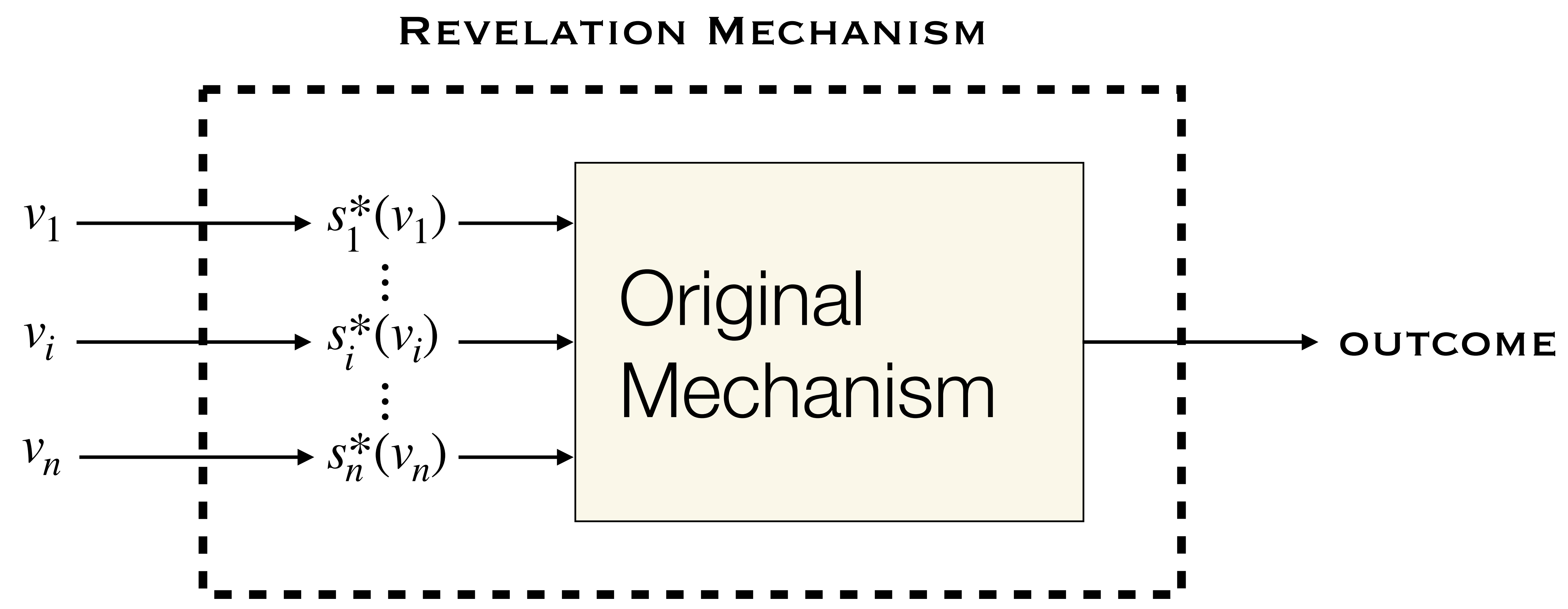
Corollary: To prove that a goal cannot be achieved by any mechanism, it is sufficient to prove that it cannot be achieved by any truthful direct mechanism.

Definitions: Truthful & Direct
In a **direct mechanism**, each agent's **actions** are simply reports of their **type**.
Agents report their own types simultaneously.

A direct mechanism is **truthful** if it is an **equilibrium** for agents to report their **true type**.

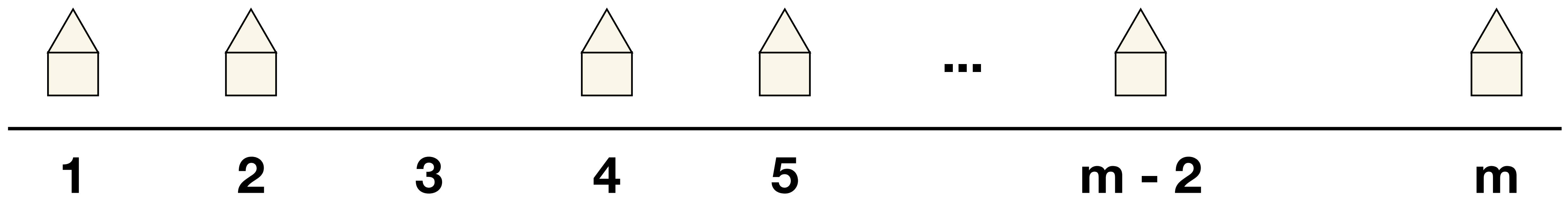
Revelation Principle Proof Sketch

- Let s_i^* be agent i's equilibrium strategy in the **original mechanism** M
- So the outcome will be $M\left(s_1^*(v_1), \ldots, s_n^*(v_n)\right)$
- **Revelation mechanism** R **simulates** M:
  1. Agents report a valuation v_i
  2. Revelation mechanism **computes** $s_i^*(v_i)$ for each player
  3. Outcome is $R(v_1, \ldots, v_n) = M\left(s_1^*(v_1), \ldots s_n^*(v_n)\right)$



Scenario: Locating a Library

- n agents live on a street with evenly-spaced addresses
  - i.e., 2 is twice as far from 4 as it is from 3
- Each agent has **single-peaked preferences:**
  They would like the library to be as **close** to them as possible
- The **Center** wants to locate a library at the **median** of the agents' preferences:
  1. Ask each agent for their **favourite** location
  2. Locate the library at the **median** of the reported locations

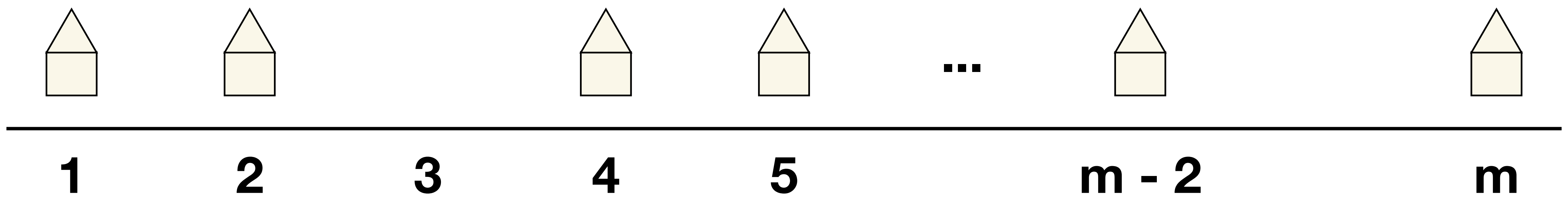


Questions:

1. Is this a **direct** mechanism?
2. Is this a **truthful** mechanism?
   Why or why not?

Locating a Library 2.0

- n agents live on a street with evenly-spaced addresses
  - i.e., 2 is twice as far from 4 as it is from 3
  - Not every address has an agent living there
- Each agent has **single-peaked preferences:**
  They would like the library to be as **close** to them as possible
- The **Center** wants to locate a library at the **mean** of the agents' preferences:
  1. Ask each agent for their **favourite** location
  2. Locate the library at the **mean** of the reported locations



Questions:

1. Is this a **direct** mechanism?
2. Is this a **truthful** mechanism?
   Why or why not?

Auctions Revisited: Outcomes & Payments

- A mechanism that simply queries valuations and then chooses the maximum is **not truthful**
  - Anyone who reports less than the maximum is failing to best respond; i.e., they have a **deviation**
- The reason that auctions work is that they add something to the outcome: **payments**
- The outcome that an auction chooses is a **combination** of **allocation** (who gets the item) and **payments** (how much do they pay)
- So it's possible for me to prefer "not getting the item" to "getting the item and paying $9000"

Summary

- **Mechanism design** is game theory in reverse:
  - The **Center** **designs a game** for agents to play
  - Agents are assumed to act "rationally" (i.e., play an equilibrium)
  - In equilibrium, the **Center's** **goals** (e.g., allocate to highest valuation, maximize revenue) are satisfied
  - E.g.: Auctions for allocating precious objects to agents with unknown valuations
- Some goals cannot be satisfied by **any** mechanism
- **Revelation principle:** If any mechanism can achieve a goal, then there exists a direct, truthful mechanism that achieves the goal
- **Payments** can make many unachievable goals achievable