



CS 540 Introduction to Artificial Intelligence

Games I

University of Wisconsin-Madison
Spring 2026 Sections 1 & 2

Announcements

- **Homework:**

- HW8 online, due on Wednesday **April 15th at 11:59 PM**

- **Class roadmap:**

Games – Part I
Games – Part II
Reinforcement Learning

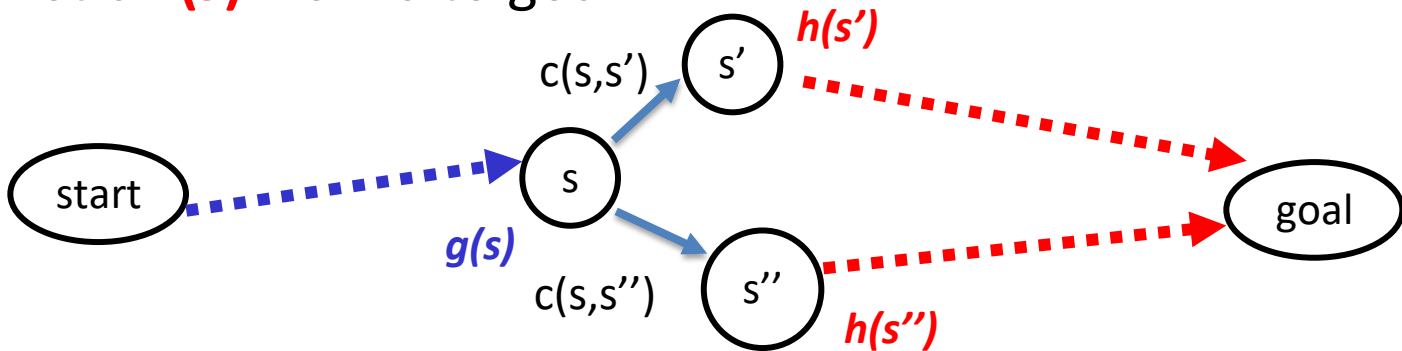
Outline

- Review/complete Informed Search
- Introduction to game theory
 - Properties of games, mathematical formulation
- Simultaneous-Move Games
 - Normal form, strategies, dominance, Nash equilibrium

Review: Informed Search

Informed search. Know:

- All uninformed search properties, plus
- Heuristic $h(s)$ from s to goal.



- Goal: **speed up search.**

Review: A* Search

- We store potential next states with a priority queue
- Expand the state with the smallest $g(s) + h(s)$
- One additional requirement:
 - Demand that $h(s) \leq h^*(s)$ where $h^*(s)$ is true cost from s to goal.
 - If heuristic has this property, it is called “admissible”
 - Still need $h(s) \geq 0$

A* Analysis

Some properties:

- Terminates!
- A* can use **lots of memory**:
 - $O(\# \text{ states})$.
- Will run out on large problems.
- Next, we will consider some alternatives to deal with this.

Break & Quiz

Q 2.1: Consider two heuristics for the 8 puzzle problem. h_1 is the number of tiles in wrong position. h_2 is the l_1 /Manhattan distance between the tiles and the goal location. How do h_1 and h_2 relate?

- A. h_2 dominates h_1
- B. h_1 dominates h_2
- C. Neither dominates the other

Break & Quiz

Q 2.1: Consider two heuristics for the 8 puzzle problem. h_1 is the number of tiles in wrong position. h_2 is the l_1 /Manhattan distance between the tiles and the goal location. How do h_1 and h_2 relate?

- **A. h_2 dominates h_1**
- B. h_1 dominates h_2
- C. Neither dominates the other

Break & Quiz

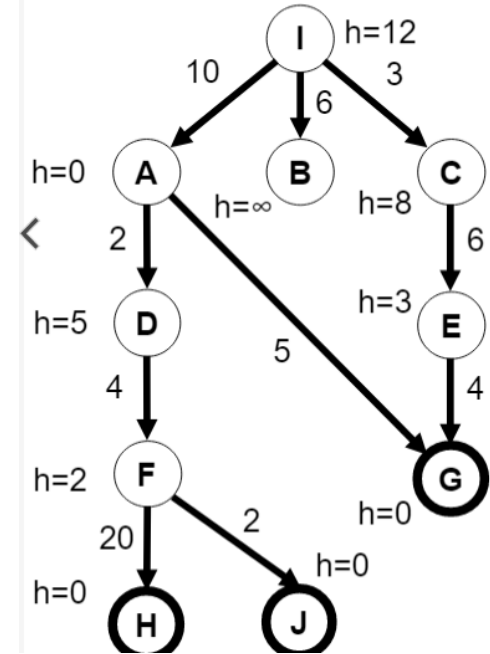
Q 2.1: Consider two heuristics for the 8 puzzle problem. h_1 is the number of tiles in wrong position. h_2 is the l_1 /Manhattan distance between the tiles and the goal location. How do h_1 and h_2 relate?

- **A. h_2 dominates h_1**
- **B. h_1 dominates h_2 (No: h_1 is a distance where each entry is at most 1, h_2 can be greater)**
- C. Neither dominates the other

Break & Quiz

Q 2.2: Consider the state space graph below. Goal states have bold borders. $h(s)$ is shown next to each node. What node will be expanded by A* after the initial state I?

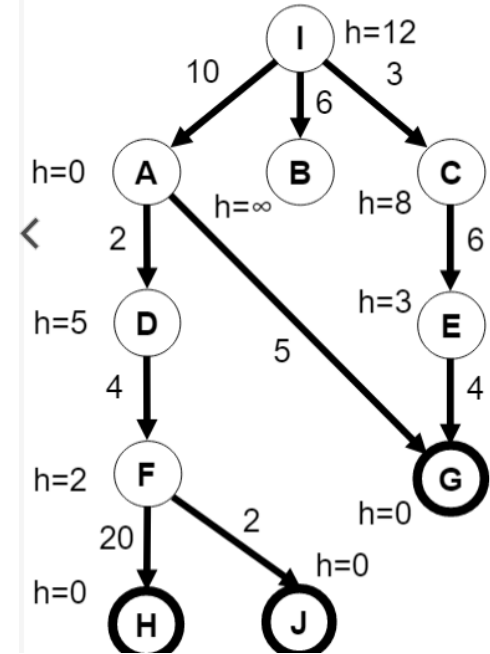
- A. A
- B. B
- C. C



Break & Quiz

Q 2.2: Consider the state space graph below. Goal states have bold borders. $h(s)$ is show next to each node. What node will be expanded by A* after the initial state I?

- **A. A**
- B. B
- C. C



IDA*: Iterative Deepening A*

Similar idea to our earlier iterative deepening.

- Bound the memory in search.
- At each phase, don't expand any node with $g(s) + h(s) > k$,
 - Assuming integer costs, do this for $k=0$, then $k=1$, then $k=2$, and so on
- Complete + optimal, might be costly time-wise
 - Revisit many nodes
- Lower memory use than A*

IDA*: Properties

How many restarts do we expect?

- With integer costs, optimal solution C^* , at most C^*

What about non-integer costs?

- Initial threshold k . Use the same rule for non-expansion
- Set new k to be the min $g(s) + h(s)$ for non-expanded nodes
- Worst case: restarted for each state

Beam Search

General approach (beyond A* too)

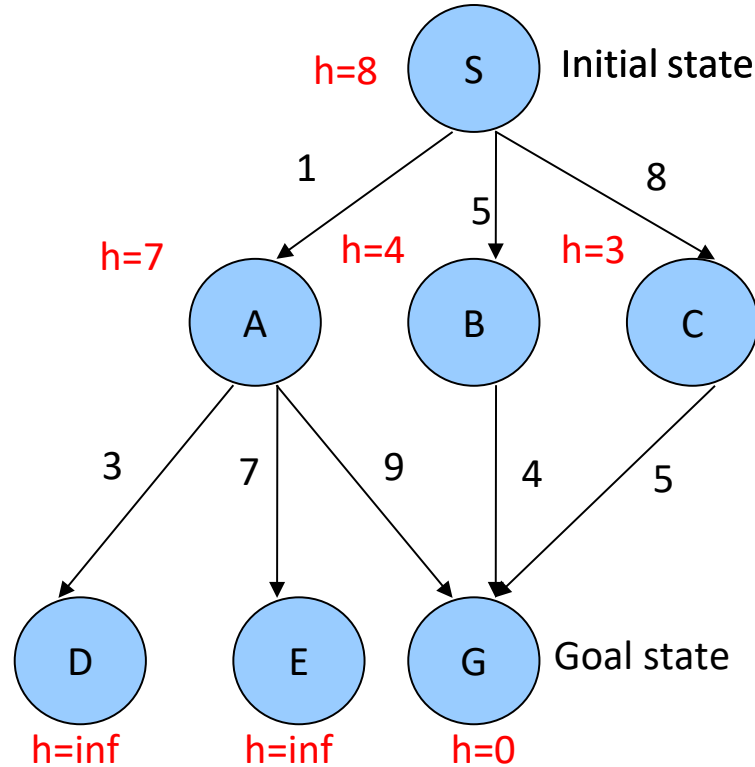
- Priority queue with fixed size k ; beyond k nodes, **discard!**
- **Upside**: good memory efficiency
- **Downside**: not complete or optimal

Variation:

- Priority queue with nodes that **are at most ϵ worse** than best node.

Recap and Examples

Example for A*:



Recap and Examples

Example for A*:

OPEN

S(0+8)

A(1+7) B(5+4) C(8+3)

B(5+4) C(8+3) D(4+inf) E(8+inf) G(10+0)

C(8+3) D(4+inf) E(8+inf) G(9+0)

C(8+3) D(4+inf) E(8+inf)

CLOSED

-

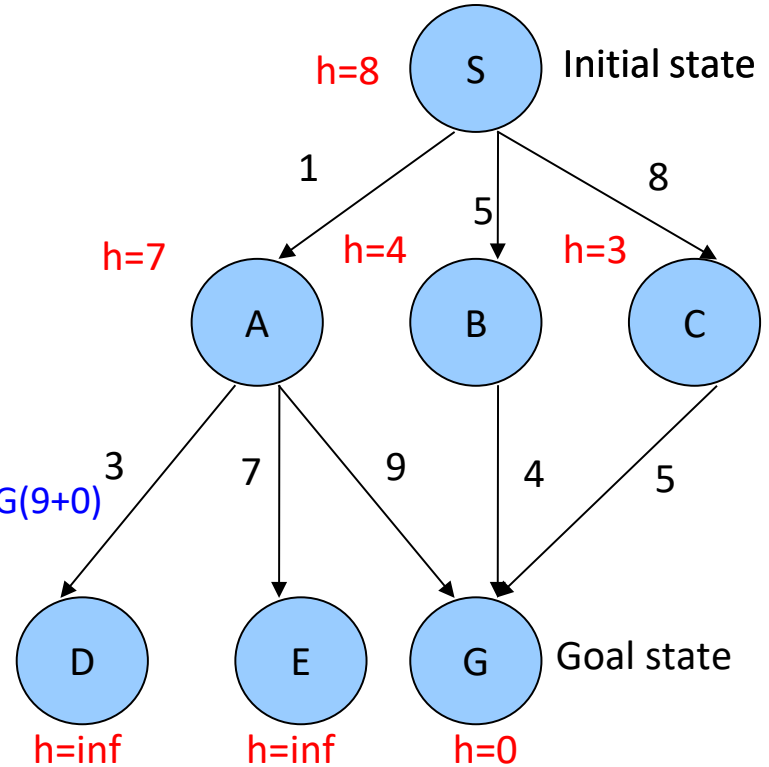
S(0+8)

A(1+7)

S(0+8) A(1+7) B(5+4)

S(0+8) A(1+7) B(5+4) G(9+0)

G → B → S



Recap and Examples

Example for IDA*:

Threshold = 8

PATH PREFIX

-

S

SA

SAH

SAHF

SAD

OPEN

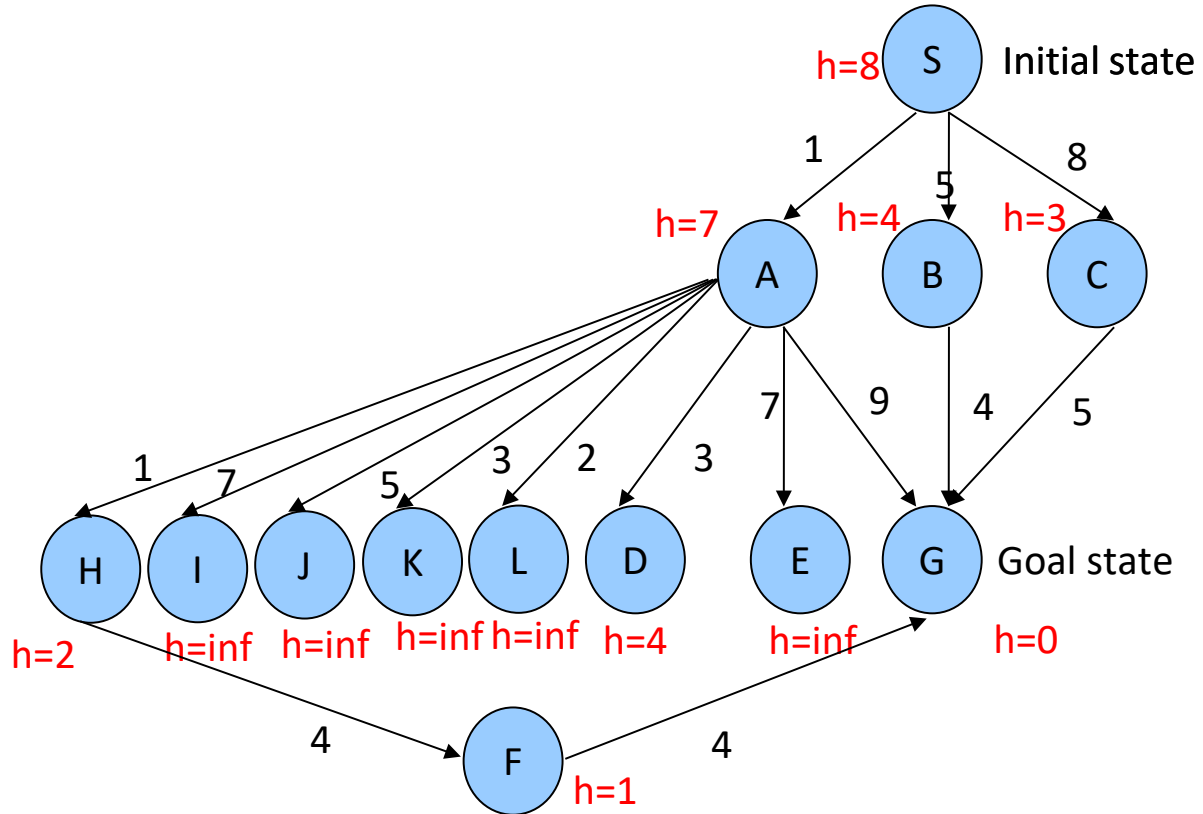
S(0+8)

A(1+7)

H(2+2) D(4+4)

D(4+4) F(6+1)

D(4+4)

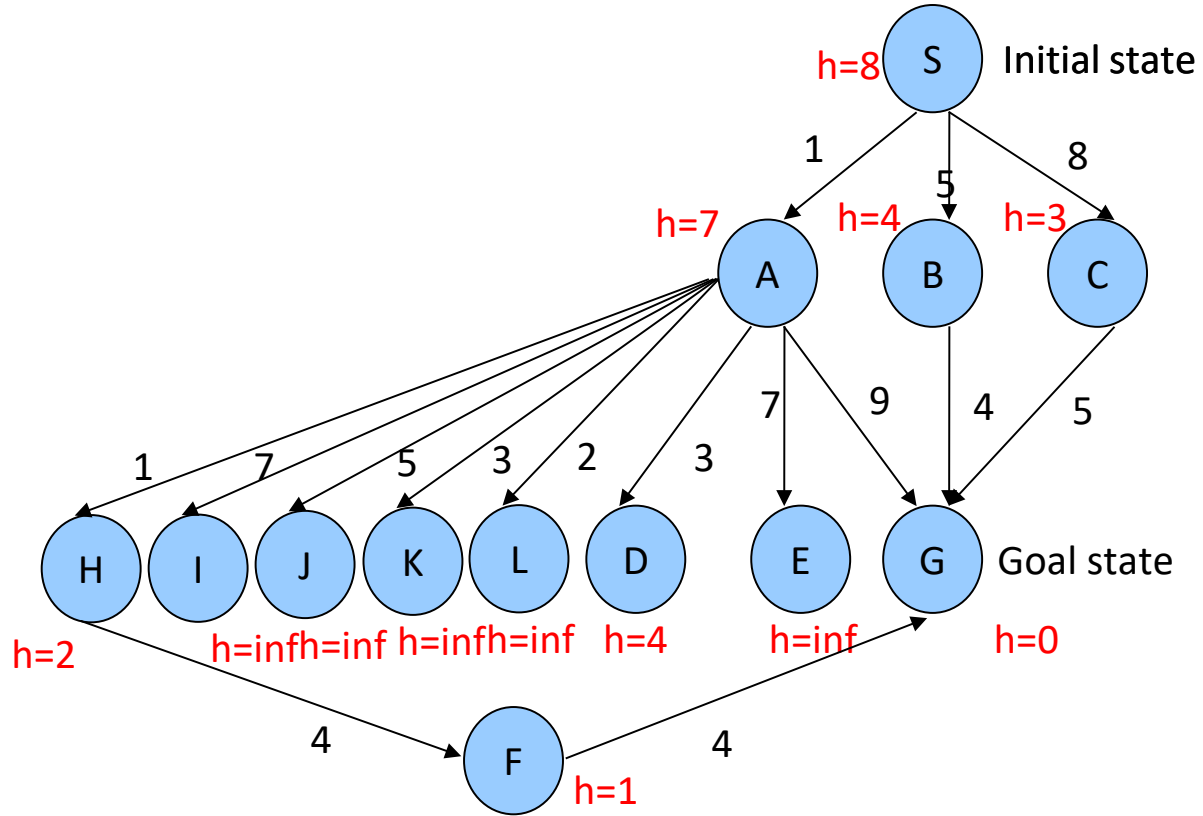


Recap and Examples

Example for IDA*:

Threshold = 9

PREFIX	OPEN
-	S(0+8)
S	A(1+7) B(5+4)
SA	B(5+4) H(2+2) D(4+4)
SAH	B(5+4) D(4+4) F(6+1)
SAHF	B(5+4) D(4+4)
SAD	B(5+4)
SB	G(9+0)
SBG	



Recap and Examples

Example for Beam Search: k=2

CURRENT

-

S

A

H

F

D

G

OPEN

S(0+8)

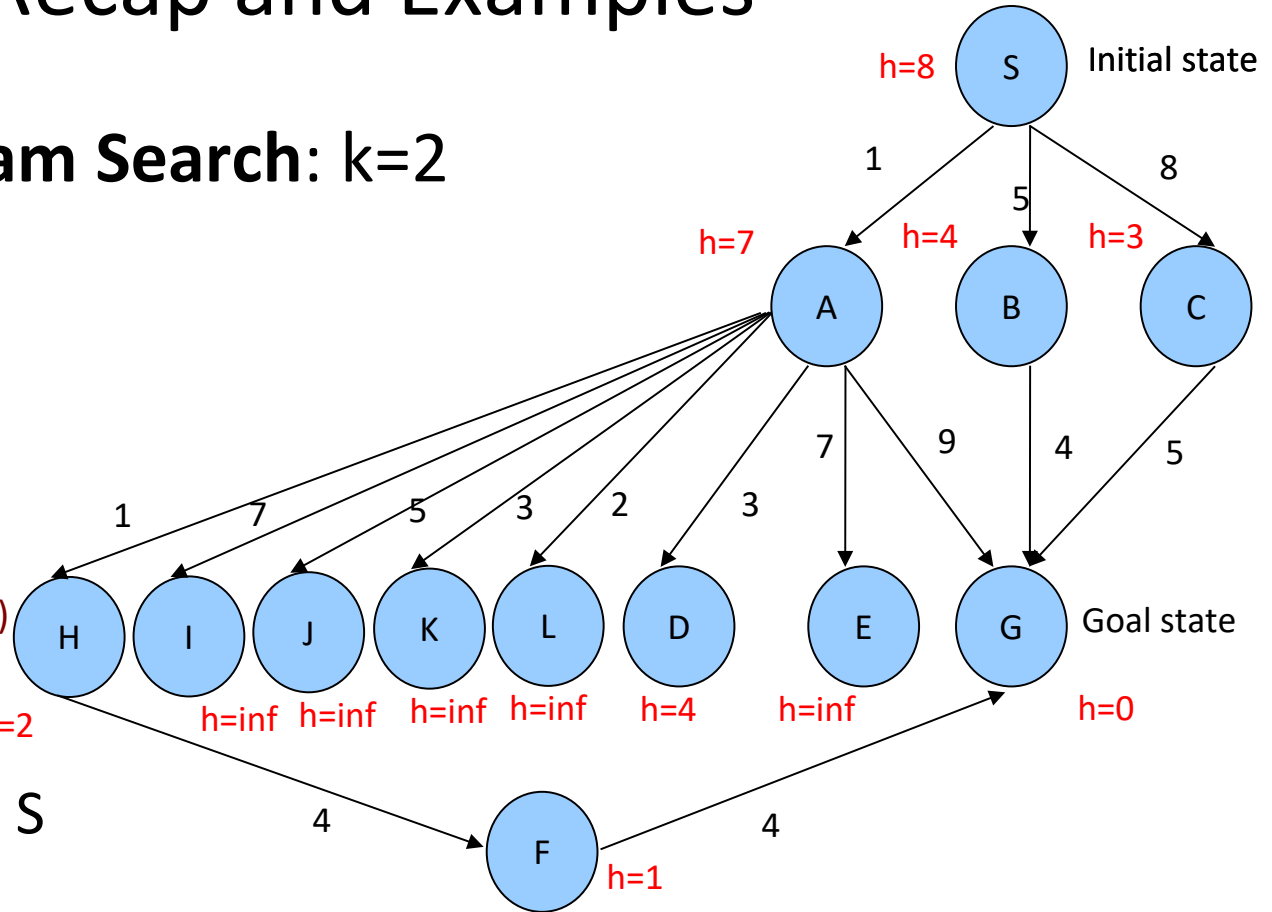
A(1+7) B(5+4)

H(2+2) D(4+4)

D(4+4) F(6+1)

D(4+4) G(10+0)

G(10+0)



G → F → H → A → S

Not optimal!

Summary

- Informed search: introduce heuristics
 - Not all approaches work: best-first greedy is bad
- A* algorithm
 - Properties of A*, idea of admissible heuristics
- Beyond A*
 - IDA*, beam search. Ways to deal with space requirements.

Suggested Readings

Textbook: Artificial Intelligence: A Modern Approach (4th edition).

Stuart Russell and Peter Norvig. Pearson, 2020.

- Section 3.5

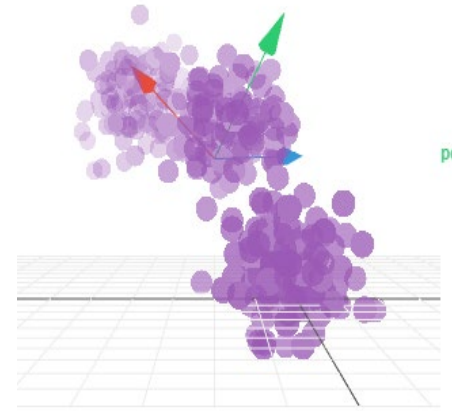


Games I

So Far in The Course

We looked at techniques:

- **Unsupervised:** See data, do something with it. Unstructured.
- **Supervised:** Train a model to make predictions. More structure (labels).
- **Planning and Games:** Much more structure.



Victor Powell



indoor

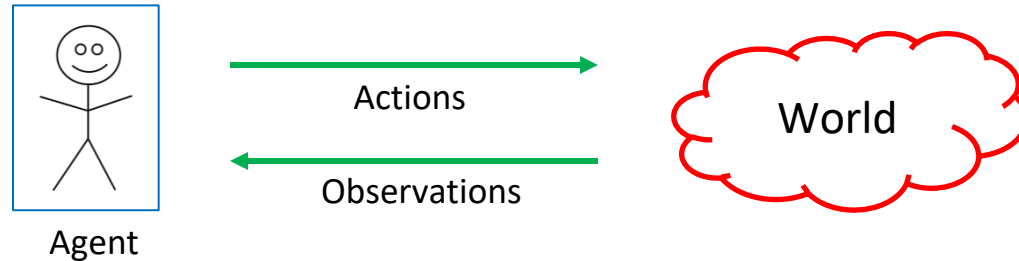


outdoor



More General Model

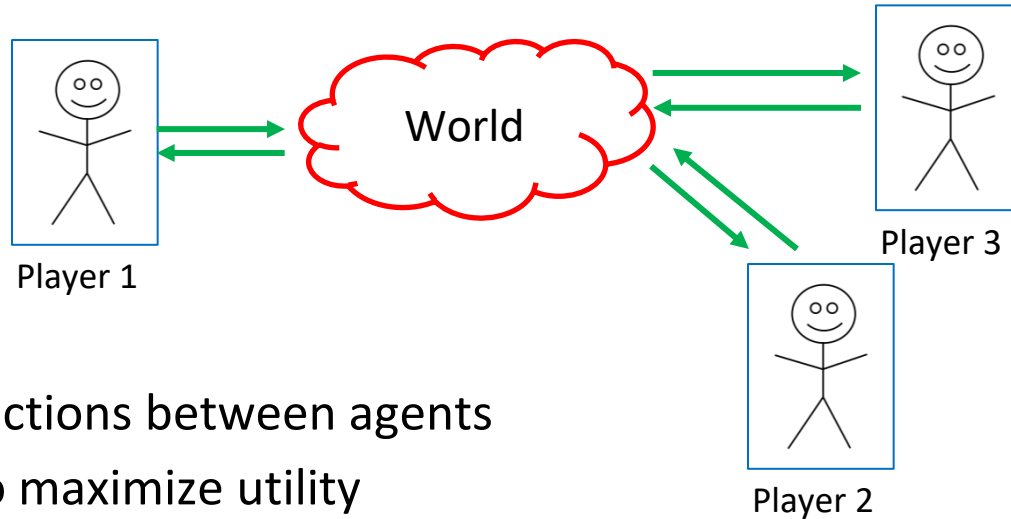
Suppose we have an **agent interacting** with the **world**



- Agent receives a reward based on state of the world
 - **Goal:** maximize reward / utility (\$\$\$)
 - Note: now **data** consists of actions, observations, and rewards
 - Setup for decision theory, reinforcement learning, planning

Games: Multiple Agents

Games setup: **multiple** agents



- Now: interactions between agents
- Still want to maximize utility
- Requires **strategic** decision making.

Modeling Games: Properties

Let's work through **properties** of games

- **Number** of agents/players
- Action space: **finite** or **infinite**
- **Deterministic** or **random**
- **Zero**-sum or **general**-sum
- **Sequential** or **simultaneous moves**

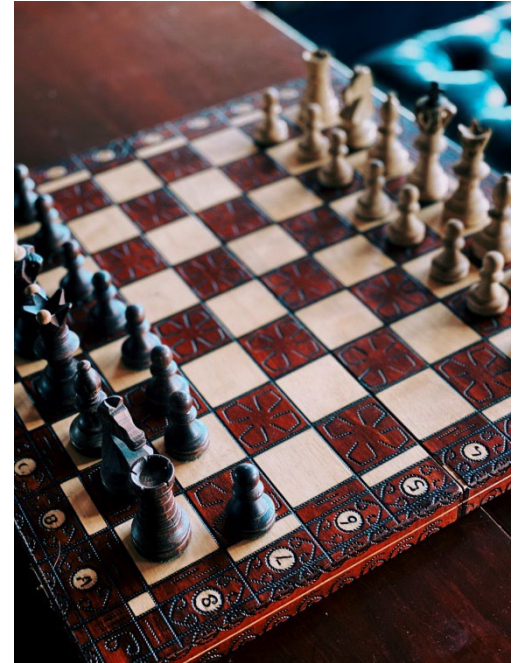


Wiki

Property 1: **Number** of players

1 or more players

- Usually interested in ≥ 2 players
- Typically a finite number of players



Property 2: Action Space

Action space: set of possible actions an agent can choose from.

Can be finite or infinite.

Examples:

- Rock-paper-scissors
- Tennis

Property 3: **Deterministic** or **Random**

- Is there **chance** in the game?
 - Poker
 - Chess
 - Scrabble



Property 4: **Sum of payoffs**

- Two basic types: zero sum vs. general sum.
- Zero sum: one player's win is the other's loss
 - Pure competition.
 - Example: rock-paper-scissors
- General sum
 - Example: driving to work, prisoner's dilemma

Property 5: **Sequential** or **Simultaneous Moves**

- Simultaneous: all players take action at the same time
- Sequential: take turns
 - But payoff is often only revealed at end of game

Quiz break:

Give the properties of the game shown on the right:

- Number of players?
- Deterministic or random?
- Sum of pay-offs?
- Finite or infinite action-space?
- Sequential or simultaneous?



Normal Form Game

Mathematical description of simultaneous games.

- n players $\{1, 2, \dots, n\}$
- Player i chooses strategy a_i from action space A_i .
- Strategy profile: $a = (a_1, a_2, \dots, a_n)$
- Player i gets rewards $u_i(a)$
 - **Note:** reward depends on other players!
- We consider the simple case where all reward functions are common knowledge.

Example of Normal Form Game

Prisoner's Dilemma

		Player 2	
		<i>Stay silent</i>	<i>Betray</i>
Player 1	<i>Stay silent</i>	-1, -1	-3, 0
	<i>Betray</i>	0, -3	-2, -2

- 2 players, 2 actions: yields 2x2 payoff matrix
- Strategy set: {Stay silent, betray}

Strictly Dominant Strategies

Let's analyze such games. Some strategies are better than others!

- Strictly dominant strategy: if a_i strictly better than b *regardless* of what other players do, a_i is **strictly dominant**
- I.e., $u_i(a_i, a_{-i}) > u_i(b, a_{-i}), \forall b \neq a_i, \forall a_{-i}$



All of the other entries
of a excluding i

- Sometimes a dominant strategy does not exist!

Strictly Dominant Strategies Example

Back to Prisoner's Dilemma

- Examine all the entries: betray strictly dominates
- Check:

		Player 2	
		<i>Stay silent</i>	<i>Betray</i>
Player 1	<i>Stay silent</i>	-1, -1	-3, 0
	<i>Betray</i>	0, -3	-2, -2

Dominant Strategy Equilibrium

a^* is a (strictly) dominant strategy equilibrium (DSE), if every player i has a strictly dominant strategy a_i^*

- Rational players will play at DSE, if one exists.

		Player 2	
		<i>Stay silent</i>	<i>Betray</i>
Player 1	<i>Stay silent</i>	-1, -1	-3, 0
	<i>Betray</i>	0, -3	-2, -2

Dominant Strategy: Absolute Best Responses

Player i 's best response to strategy to a_{-i}

$$BR(a_{-i}) = \operatorname{argmax}_b u_i(b, a_{-i})$$

$BR(\text{player2=silent}) = \text{betray}$

$BR(\text{player2=betray}) = \text{betray}$

	Player 2	
	<i>Stay silent</i>	<i>Betray</i>
Player 1		
<i>Stay silent</i>	-1, -1	-3, 0
<i>Betray</i>	0, -3	-2, -2

a_i^* is the dominant strategy for player i , if

$$a_i^* = BR(a_{-i}), \forall a_{-i}$$

Dominant Strategy Equilibrium

Dominant Strategy Equilibrium does not always exist.

		Player 2	
		<i>L</i>	<i>R</i>
Player 1	<i>L</i>	2, 1	0, 0
	<i>R</i>	0, 0	1, 2

Nash Equilibrium

a^* is a Nash equilibrium if no player has an incentive to **unilaterally deviate**

$$u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*) \quad \forall a_i \in A_i$$

		Player 2	
		L	R
Player 1	L	2, 1	0, 0
	R	0, 0	1, 2

Nash Equilibrium: Best Response to Each Other

a^* is a Nash equilibrium:

$$\forall i, \forall b \in A_i: u_i(a_i^*, a_{-i}^*) \geq u_i(b, a_{-i}^*)$$

(no player has an incentive to **unilaterally deviate**)

Equivalently, for each player i :

$$a_i^* \in BR(a_{-i}^*) = \operatorname{argmax}_b u_i(b, a_{-i}^*)$$

Nash Equilibrium: Best Response to Each Other

- Compared to DSE: a **DSE is a NE**, the other direction is generally not true:

$$a_i^* = BR(a_{-i}), \forall a_{-i}$$

- Pure Nash equilibrium:
 - A **pure strategy** is a deterministic choice (no randomness).
 - Later: we will consider **mixed** strategies
 - In pure Nash equilibrium, players can only play pure strategies.

Finding (pure) Nash Equilibria by hand

- As player 1: For each column, find the best response, underscore it.

		Player 2	
		<i>L</i>	<i>R</i>
Player 1	<i>L</i>	<u>2, 1</u>	0, 0
	<i>R</i>	0, 0	<u>1, 2</u>

Finding (pure) Nash Equilibria by hand

- As player 2: For each row, find the best response, upper-score it.

		Player 2	
		<i>L</i>	<i>R</i>
Player 1	<i>L</i>	<u>2, 1</u>	0, 0
	<i>R</i>	0, 0	<u>1, 2</u>

Finding (pure) Nash Equilibria by hand

- Entries with both lower and upper bars are pure NEs.

		Player 2	
		<i>L</i>	<i>R</i>
Player 1	<i>L</i>	<u>2, 1</u>	0, 0
	<i>R</i>	0, 0	<u>1, 2</u>

Break & Quiz

Q 1.1: Which of the following is **true**?

- (i) Nash equilibria require each player to know other players' possible strategies
- (ii) Nash equilibria require rational play

- A. Neither
- B. (i) but not (ii)
- C. (ii) but not (i)
- D. Both

Break & Quiz

Q 1.1: Which of the following is **true**?

- (i) Nash equilibria require each player to know other players' possible strategies
- (ii) Nash equilibria require rational play

- A. Neither
- B. (i) but not (ii)
- C. (ii) but not (i)
- **D. Both**

Break & Quiz

Q 1.2: Which of the following is **true**?

- (i) Nash equilibria require each player to know other players' possible strategies
- (ii) Nash equilibria require rational play

- A. Neither (See below)
- B. (i) but not (ii) (Rational play required: i.e., what if prisoners desire longer jail sentences?)
- C. (ii) but not (i) (The basic assumption of Nash equilibria is knowing all of the strategies involved)
- D. **Both**

Break & Quiz

Q1.2: Two firms, A and B, are deciding whether to launch a new product. Each firm can either launch or not launch. Their profits depend on their choices, and the payoff matrix is as follows:

	<i>B: Launch</i>	<i>B: Not Launch</i>
<i>A: Launch</i>	(20, 20)	(40, 10)
<i>A: Not Launch</i>	(10, 40)	(30, 30)

What is the strictly dominant strategy for each firm?

- A. A's dominant strategy is to launch, and B's dominant strategy is not to launch.
- B. A's dominant strategy is to launch, and B's dominant strategy is to launch.
- C. A's dominant strategy is not to launch, and B's dominant strategy is to launch.
- D. A's dominant strategy is not to launch, and B's dominant strategy is not to launch.

Break & Quiz

Q1.2: Two firms, A and B, are deciding whether to launch a new product. Each firm can either launch or not launch. Their profits depend on their choices, and the payoff matrix is as follows:

	<i>B: Launch</i>	<i>B: Not Launch</i>
<i>A: Launch</i>	(20, 20)	(40, 10)
<i>A: Not Launch</i>	(10, 40)	(30, 30)

What is the strictly dominant strategy for each firm?

- A. A's dominant strategy is to launch, and B's dominant strategy is not to launch.
- B. A's dominant strategy is to launch, and B's dominant strategy is to launch.**
- C. A's dominant strategy is not to launch, and B's dominant strategy is to launch.
- D. A's dominant strategy is not to launch, and B's dominant strategy is not to launch.

Break & Quiz

Q1.2: Two firms, A and B, are deciding whether to launch a new product. Each firm can either launch or not launch. Their profits depend on their choices, and the payoff matrix is as follows:

	<i>B: Launch</i>	<i>B: Not Launch</i>
<i>A: Launch</i>	(20, 20)	(40, 10)
<i>A: Not Launch</i>	(10, 40)	(30, 30)

If B launches, then A is getting the maximum payoff from launching, which is 20. If B does not launch, then A is getting maximum payoff again if they launch which is 40. If A launches, then for B is also better to launch because the payoff will be 20. If A does not launch, B is getting maximum payoff if B launches, which is 40. Thus, for both players the dominant strategy is to launch.

What is the strictly dominant strategy for each firm?

- A. A's dominant strategy is to launch, and B's dominant strategy is not to launch.
- B. A's dominant strategy is to launch, and B's dominant strategy is to launch.**
- C. A's dominant strategy is not to launch, and B's dominant strategy is to launch.
- D. A's dominant strategy is not to launch, and B's dominant strategy is not to launch.

Pure Nash Equilibrium may not exist

So far, pure strategy: each player picks a deterministic strategy. But:

		Player 2		
		<i>rock</i>	<i>paper</i>	<i>scissors</i>
Player 1	<i>rock</i>	0, 0	<u>-1, 1</u>	<u>1, -1</u>
	<i>paper</i>	<u>1, -1</u>	0, 0	-1, 1
	<i>scissors</i>	<u>-1, 1</u>	<u>1, -1</u>	0, 0

Mixed Strategies

Can also randomize actions: “**mixed**”

- Player i assigns probabilities x_i to each action

$$x_i(a_i), \text{ where } \sum_{a_i \in A_i} x_i(a_i) = 1, x_i(a_i) \geq 0$$

- Now consider **expected rewards**

$$u_i(x_i, x_{-i}) = E_{a_i \sim x_i, a_{-i} \sim x_{-i}} u_i(a_i, a_{-i}) = \sum_{a_i} \sum_{a_{-i}} x_i(a_i) x_{-i}(a_{-i}) u_i(a_i, a_{-i})$$

Mixed Strategy Nash Equilibrium

Consider the mixed strategy $x^* = (x_1^*, \dots, x_n^*)$

- This is a **Nash equilibrium** if

$$u_i(x_i^*, x_{-i}^*) \geq u_i(x_i, x_{-i}^*) \quad \forall x_i \in \Delta_{A_i}, \forall i \in \{1, \dots, n\}$$



Better than doing
anything else,
“**best response**”



Space of probability
distributions over
strategies.

- Intuition: nobody can **increase expected reward** by changing only their own strategy.

Mixed Strategy Nash Equilibrium

Example: $x_1^*(\cdot) = x_2^*(\cdot) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$

Player 2	<i>rock</i>	<i>paper</i>	<i>scissors</i>
Player 1			
<i>rock</i>	0, 0	-1, 1	1, -1
<i>paper</i>	1, -1	0, 0	-1, 1
<i>scissors</i>	-1, 1	1, -1	0, 0

Finding Mixed NE in 2-Player 2-action Zero-Sum Game

Example: Two Finger Morra. Show 1 or 2 fingers. The “even player” wins if the sum is even, and vice versa.

	odd		
		<i>f1</i>	f2
even			
	<i>f1</i>	2, -2	-3, 3
	f2	-3, 3	4, -4

Finding Mixed NE in 2-Player 2-action Zero-Sum Game

Two Finger Morra. Two-player zero-sum game. No pure NE:

		odd	
		$f1$	$f2$
even	$f1$	<u>2, -2</u>	<u>-3, 3</u>
	$f2$	<u>-3, 3</u>	<u>4, -4</u>

Finding Mixed NE in 2-Player 2-action Zero-Sum Game

Suppose odd's mixed strategy at NE is $(q, 1-q)$, and even's $(p, 1-p)$

By definition, p is best response to q : $u_1(p, q) \geq u_1(p', q) \forall p'$.

Note $u_1(p, q) = pu_1(f_1, q) + (1 - p)u_1(f_2, q)$

- Players only mix strategies if the expected payoffs are equal.
- If one strategy was better, they would never mix — they'd just pick the best!

$$\rightarrow u_1(f_1, q) = u_1(f_2, q)$$

- Average is no greater than components

$$\rightarrow u_1(p, q) = u_1(f_1, q) = u_1(f_2, q)$$

		q	1-q
odd		f1	f2
even			
p	f1	<u>2, -2</u>	<u>-3, 3</u>
1-p	f2	<u>-3, 3</u>	<u>4, -4</u>

Finding Mixed NE in 2-Player 2-action Zero-Sum Game

$$u_1(f_1, q) = u_1(f_2, q)$$

$$2q + (-3)(1 - q) = (-3)q + 4(1 - q)$$

$$q = \frac{7}{12}$$

$$\text{Similarly, } u_2(p, f_1) = u_2(p, f_2)$$

$$p = \frac{7}{12}$$

At this NE, even gets $-1/12$, odd gets $1/12$.

		q	1-q
		<i>f1</i>	<i>f2</i>
p	even		
	<i>f1</i>	<u>2, -2</u>	<u>-3, 3</u>
1-p	<i>f2</i>	<u>-3, 3</u>	<u>4, -4</u>

Properties of Nash Equilibrium

Major result: (John Nash '51)

- Every **finite** (players, actions) game has at least one Nash equilibrium
 - But not necessarily **pure** (i.e., deterministic strategy)
- Could be more than one
- Searching for Nash equilibria: computationally **hard**.
 - Exception: two-player zero-sum games (can be found with linear programming).

Break & Quiz

Q 2.1: Which of the following is **false**?

- (i) Rock/paper/scissors has a dominant pure strategy
- (ii) There is a Nash equilibrium for rock/paper/scissors

- A. Neither
- B. (i) but not (ii)
- C. (ii) but not (i)
- D. Both

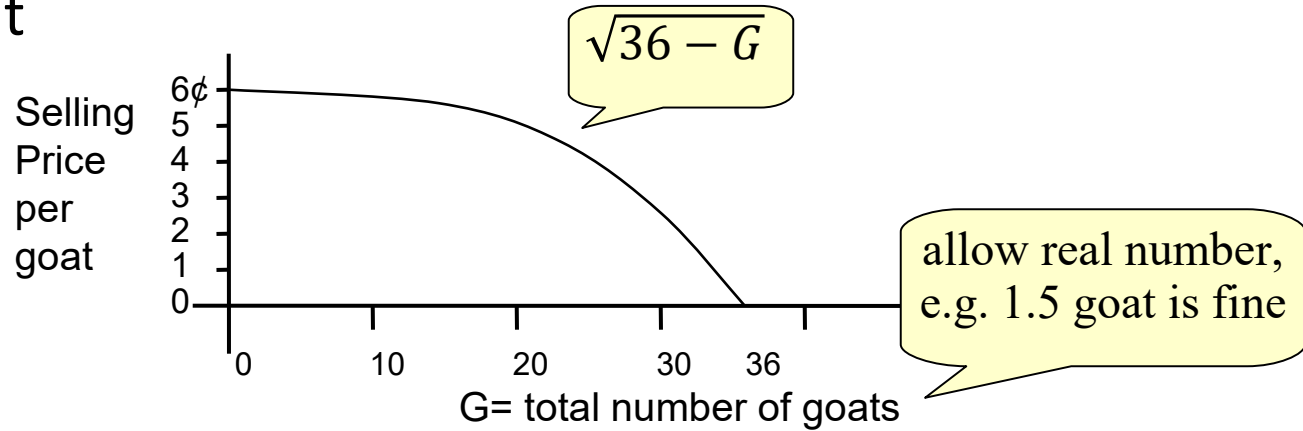
Break & Quiz

Q 2.2: Which of the following statements is true about Nash equilibrium?

- A. Mixed Strategy Nash Equilibrium all players choose their actions deterministically.
- B. Every finite (players, actions) game has at least one pure Nash equilibrium.
- C. A pure Nash Equilibrium is also always a Dominant Strategy Equilibrium.
- D. A Dominant Strategy Equilibrium is also always a pure Nash Equilibrium.

Pure NE in an Infinite game: The tragedy of the Commons

- Price per goat



- How many goats should one (out of n) rational farmer graze?
- How much would the farmer earn?

Continuous Action Game

- Each farmer has infinite number of strategies $g_i \in [0, 36]$
- The value for farmer i , when the n farmers play at (g_1, g_2, \dots, g_n) is

$$u_i(g_1, g_2, \dots, g_n) = g_i \sqrt{36 - \sum_{j \in [n]} g_j}$$

- **Assume** a pure Nash equilibrium exists.
- **Assume** (by apparent symmetry) the NE is (g^*, g^*, \dots, g^*) .

Finding g^*

- $u_i(g_1, g_2, \dots, g_n) = g_i \sqrt{36 - \sum_j g_j}$
- g^* is the best response to others (g^*, \dots, g^*)

$$\begin{aligned} g^* &= \operatorname{argmax}_{g_i \in [0, 36]} u_i(g^*, \dots, g_i, \dots, g^*) \\ &= \operatorname{argmax}_{g_i} g_i \sqrt{36 - (n-1)g^* - g_i} \end{aligned}$$

i-th argument

Finding g^*

$$g^* = \operatorname{argmax}_{g_i} g_i \sqrt{36 - (n-1)g^* - g_i}$$

- Taking derivative w.r.t. g_i , setting to 0:

$$g^* = \frac{72 - 2(n-1)g^*}{3}$$

$$g^* = \frac{72}{2n+1} \quad \text{So what?}$$

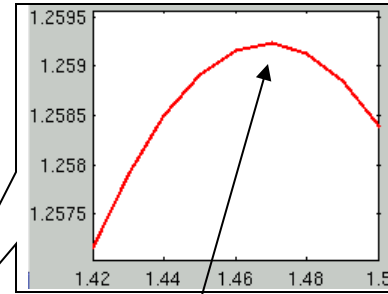
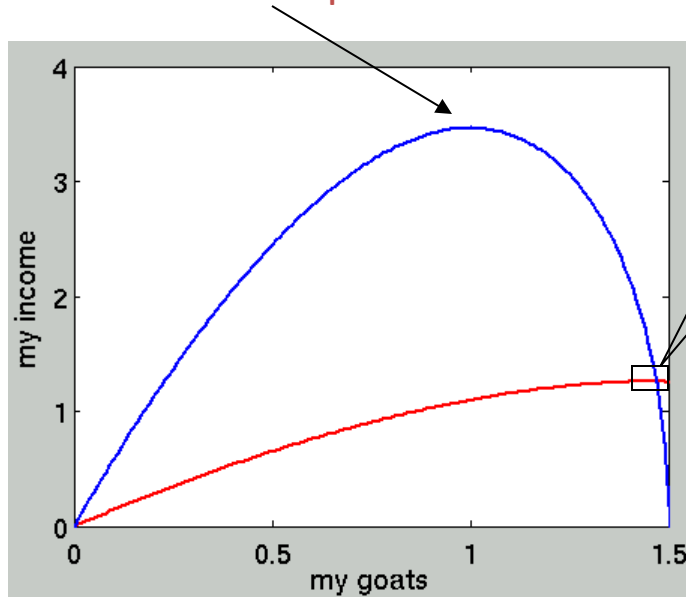
The tragedy of the Commons

- Say there are $n=24$ farmers.
Each would **rationally** graze $g_i^* = 72/(2*24+1) = 1.47$ goats
- Each would get **1.25¢**

- But if they cooperate and each grazes only 1 goat
- Each would get **3.46¢**

The tragedy of the Commons

If all 24 farmers agree on the same number of goats to raise, 1 goat per farmer would be optimal



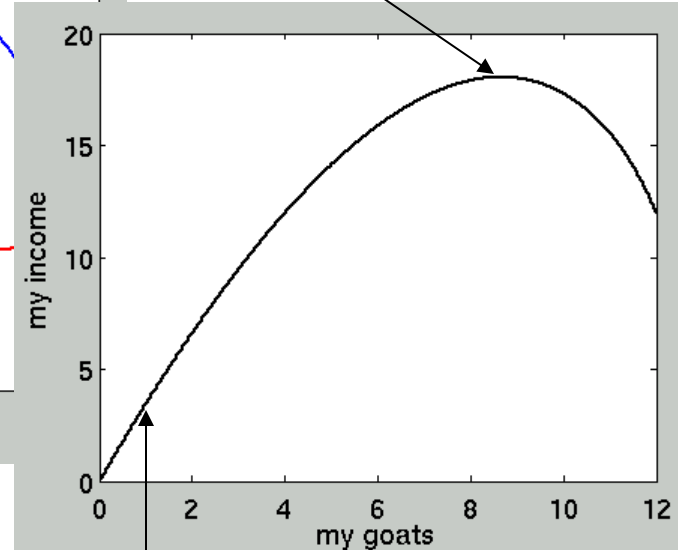
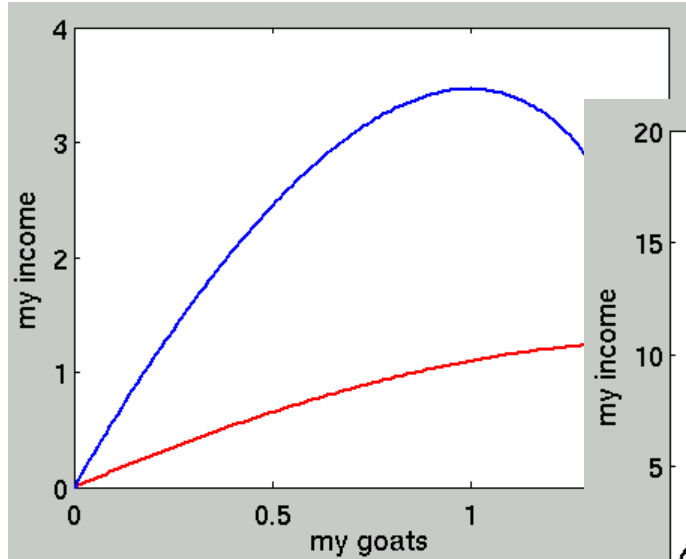
If the other 23 farmers play the N.E. of 1.47 goats each, 1.47 goats would be optimal

The tragedy of the Commons

If all 24 farmers agree on the same number of goats to raise, 1 goat per farmer would be optimal



But this is not a N.E.! A farmer can benefit from cheating (other 23 play at 1):



'by rule'

The tragedy

- Rational behaviors lead to sub-optimal solutions.
- Maximizing individual welfare not necessarily maximizes social welfare
- What went wrong?

Shouldn't have allowed **free** grazing?

It's not just the goats: Common problem for shared resources.

Mechanism design: designing the rules of a game

Summary

- Intro to game theory
 - Characterize games by various properties
- Mathematical formulation for simultaneous games
 - Normal form, dominance, Nash equilibria, mixed vs pure

Suggested Readings

Textbook: Artificial Intelligence: A Modern Approach (4th edition).

Stuart Russell and Peter Norvig. Pearson, 2020.

- Chapter 18