CS 540 Introduction to Artificial Intelligence
Games I

University of Wisconsin-Madison

Spring 2025

Announcements

- **Homework**:
  - HW8 due on **Wednesday 16th at 11:59 PM**
- Class roadmap:

Games – Part I
Games – Part II
Reinforcement Learning

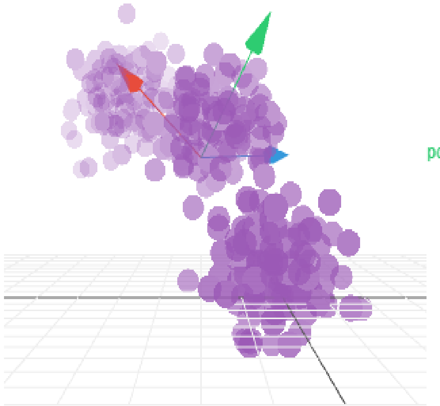
Outline

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

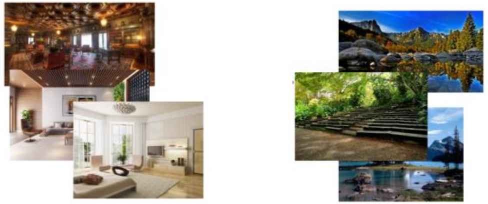
So Far in The Course

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.



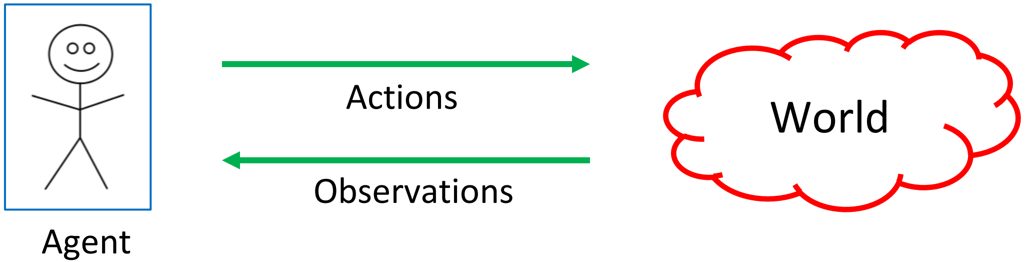
Victor Powell



indoor outdoor

More General Model

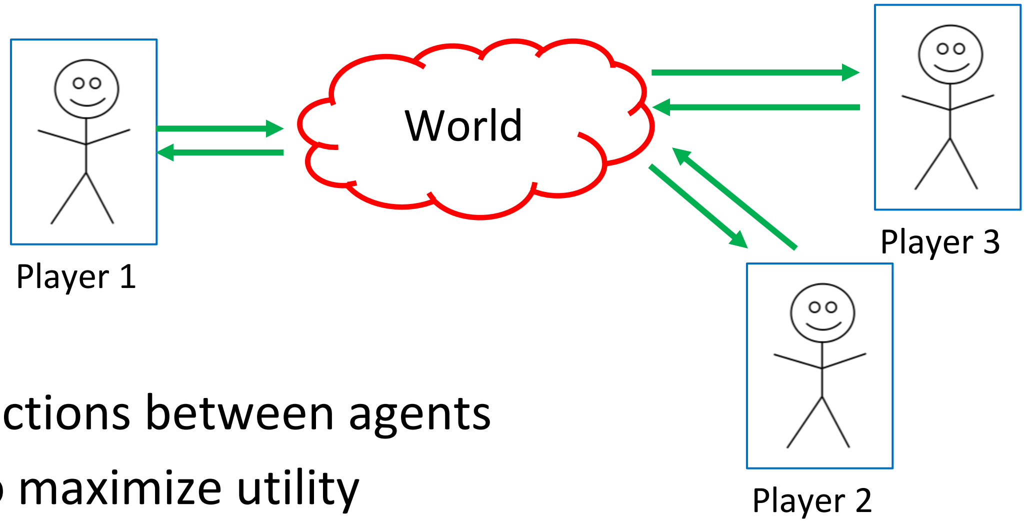
Suppose we have an **agent interacting** with the **world**



- 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

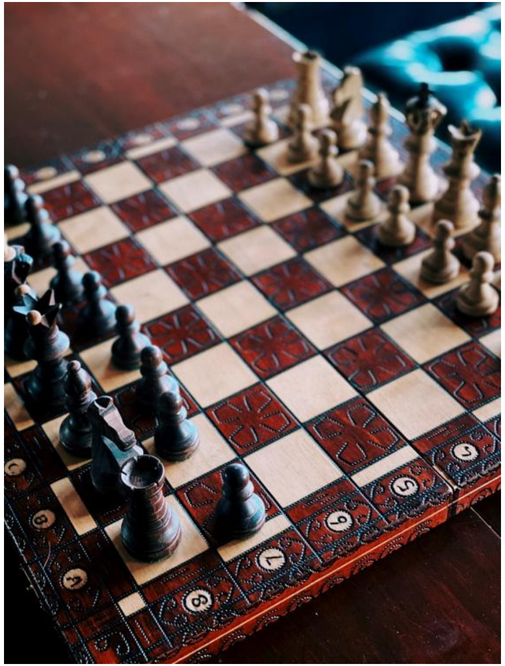
- **Number** of agents/players
- Action space: **finite** or **infinite**
- **Deterministic** or **random**
- **Zero**-sum or **general**-sum
- **Sequential** or **simultaneous moves**

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 stochastic?
- Sum of pay-offs?
- Finite or infinite action-space?
- Sequential or simultaneous?

Normal Form Game

Mathematical description of simultaneous games.

- $\boldsymbol{n}$ players $\{1,2,\ldots,n\}$
- Player i chooses strategy $\boldsymbol{a_i}$ from action space $\boldsymbol{A_i}$.
- Strategy profile: $\boldsymbol{a = (a_1, a_2, \ldots, a_n)}$
- Player i gets rewards $\boldsymbol{u_i(a)}$
  - **Note**: reward depends on other players!

- We consider the simple case where all reward functions are common knowledge.

Example of Normal Form Game

Ex: Prisoner's Dilemma

Player 2 Player 1	*Stay silent*	*Betray*
Stay silent	−1, −1	−3, 0
Betray	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 Player 1	*Stay silent*	*Betray*
Stay silent	−1, −1	−3, 0
Betray	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 / Player 1	*Stay silent*	*Betray*
Stay silent	−1, −1	−3, 0
Betray	0, −3	−2, −2

Dominant Strategy: Absolute Best Responses

Player *i*'s best response to strategy to $a_{-i} BR(a_{-i}) = \underset{b}{\arg\max}\, u_i(b, a_{-i})$

BR(player2=silent) = betray

BR(player2=betray) = betray

Player 2 Player 1	*Stay silent*	*Betray*
Stay silent	−1, −1	−3, 0
Betray	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 Player 1	*L*	*R*
T	2, 1	0, 0
B	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 Player 1	*L*	*R*
T	2, 1	0, 0
B	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:

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

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

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

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**)

- 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 Player 1	*L*	*R*
T	<u>2</u>, 1	0, 0
B	0, 0	<u>1</u>, 2

Finding (pure) Nash Equilibria by hand

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

Player 2 Player 1	*L*	*R*
T	2, 1	0, 0
B	0, 0	1, 2

Finding (pure) Nash Equilibria by hand

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

Player 2 / Player 1	*L*	*R*
T	$\overline{\underline{2, 1}}$	0, 0
B	0, 0	$\overline{\underline{1, 2}}$

Pure Nash Equilibrium may not exist

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

Player 2 / Player 1	*rock*	paper	*scissors*
rock	0, 0	-1, 1	1, -1
paper	1, -1	0, 0	-1, 1
scissors	-1, 1	1, -1	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 $x^* = (x_1^*, \ldots, x_n^*)$

- This is a **Nash equilibrium** if

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

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 Player 1	*rock*	paper	*scissors*
rock	0, 0	-1, 1	1, -1
paper	1, -1	0, 0	-1, 1
scissors	-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 / even	*f1*	f2
f1	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:

even \ odd	*f1*	f2
f1	2, -2	-3, 3
f2	-3, 3	4, -4

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

Average is no greater than components
→ $u_1(p,q) = u_1(f_1,q) = u_1(f_2,q)$

		q	1-q
	odd / even	*f1*	f2
p	*f1*	2, -2	-3, 3
1-p	f2	-3, 3	4, -4

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

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

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

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

Similarly, $u_2(p, f_1) = u_2(p, f_2)$

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

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

		q	1-q
	odd / even	*f1*	f2
p	*f1*	2, -2	-3, 3
1-p	f2	-3, 3	4, -4

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 2.1: 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 2.1: 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) (i) is indeed false: easy to check that there's no deterministic dominant strategy; (ii) is true: there is a mixed strategy Nash equilibrium
- C. (ii) but not (i)
- D. Both

Break & Quiz

Q 2.2: 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
- B. (i) but not (ii)
- C. (ii) but not (i)
- D. Both

Break & Quiz

Q 2.2: 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
- B. (i) but not (ii)
- C. (ii) but not (i)
- **D. Both**

Break & Quiz

Q 2.2: Which of the following is **true?**

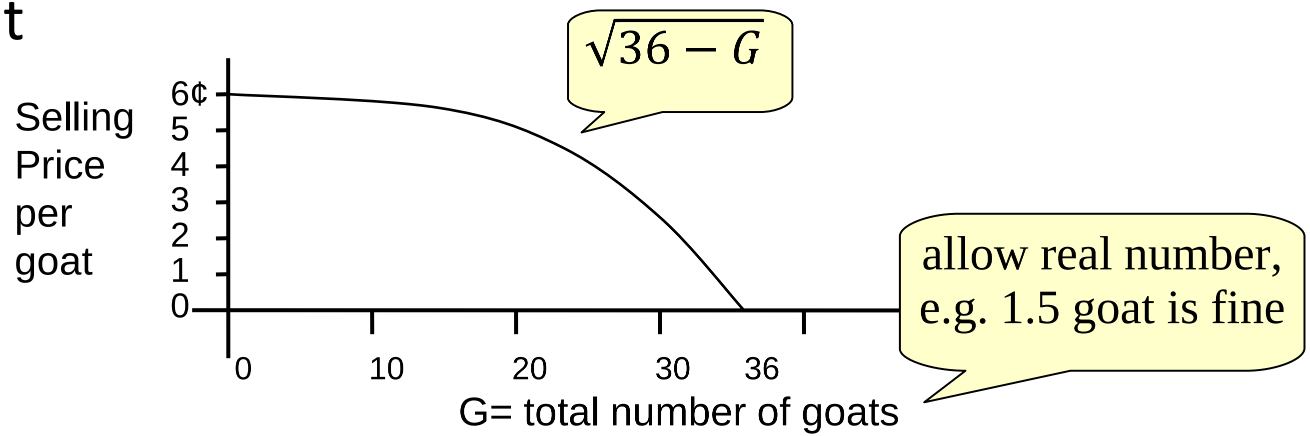
(i) Nash equilibria require each player to know other players' possible strategies

(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**

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 $(g_1, g_2, \ldots, g_n)$ is

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

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

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*,..., g*)

$$g^* = argmax_{h \in [0,36]} u_i(g^*, \dots, h, \dots, g^*)$$

i-th argument

$$= argmax_h h\sqrt{36 - (n-1)g^* - h}$$

Finding g*

$$g^* = argmax_h h\sqrt{36 - (n-1)g^* - h}$$

- Taking derivative w.r.t. *h* of the RHS, setting to 0:

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

$$g^* = \frac{72}{2n+1}$$

So what?

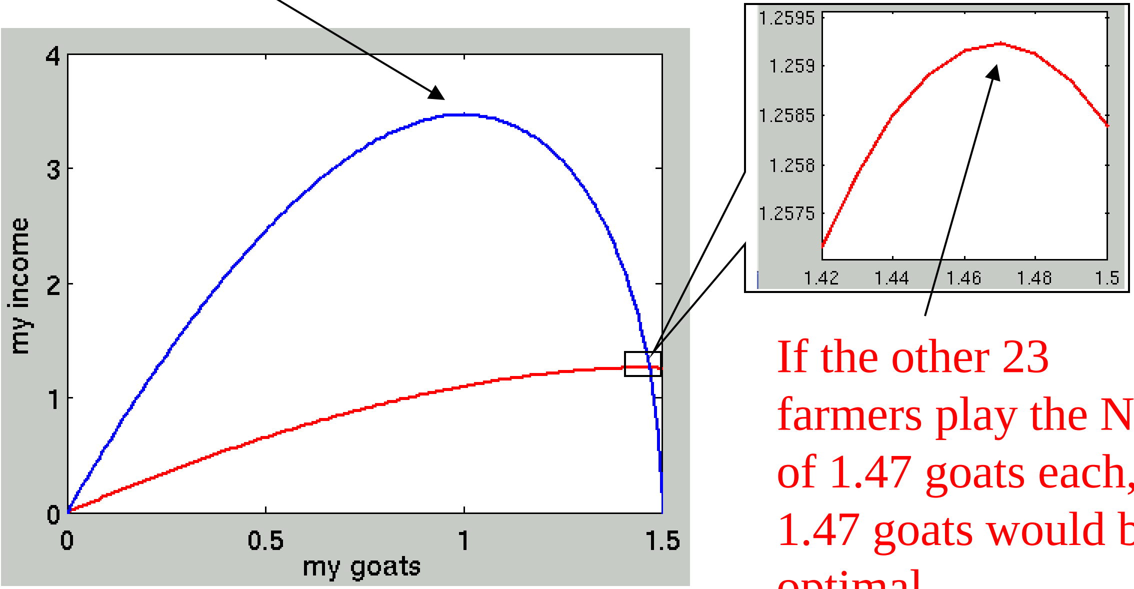
The tragedy of the Commons

- 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

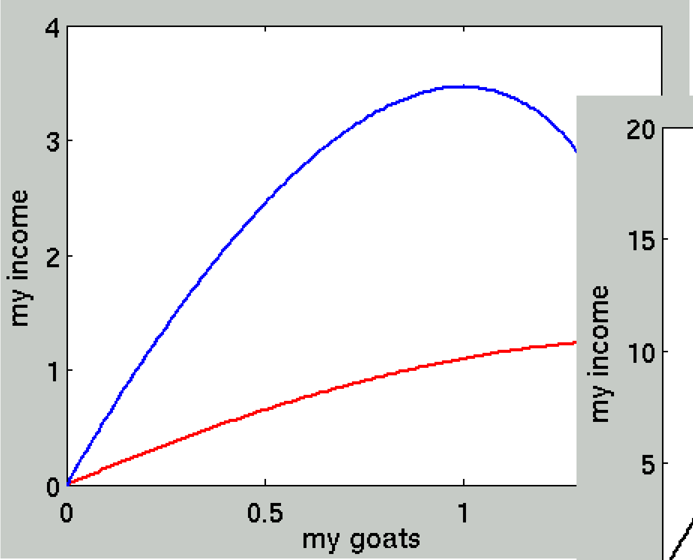


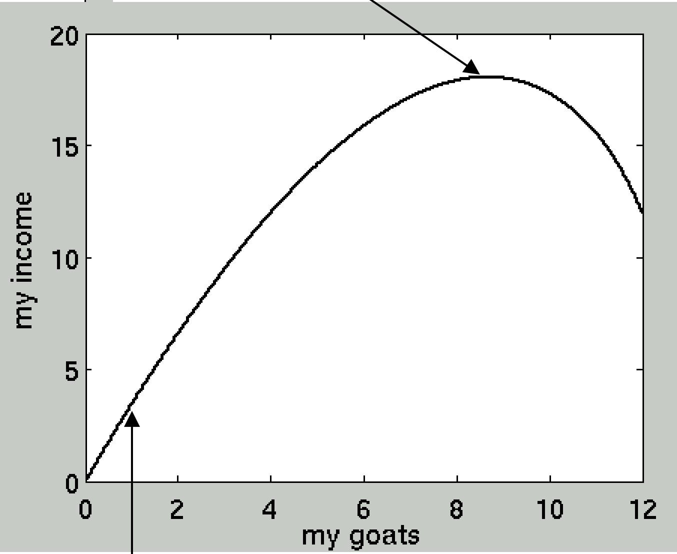
If the other 23 farmers play the N.E. of 1.47 goats each, 1.47 goats would be optimal

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  : the use of the atmosphere and the oceans for dumping of pollutants.

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