COMP0124 Multi-agent Artificial Intelligence

Repeated Games

Prof. Jun Wang
Computer Science, UCL

Recap: Game Theory

- One-shot game: the prisoner‘s dilemma

	Cooperation	Defection
Cooperation	-1, -1	-10, 0
Defection	0, -10	-4, -4

- At *Nash equilibrium* if no player wants to deviate to an alternative strategy

Exercise 1

- Find all pure strategy Nash equilibria in the game below

Player B

		L	R
Player A	U	$1,2$	$3,2$
	D	$2,4$	$0,2$

  - The rows correspond to player A's *strategies* and the columns correspond to player B's *strategies*.
  - The first entry in each box is player A's *payoff* and the second entry is player B's *payoff*

Exercise 2

- Consider the two-player game with players, strategies and payoffs described in the following game matrix

Player B

Player A	L	M	R
t	$0, 3$	$6, 2$	$1, 1$
m	$2, 3$	$0, 1$	$7, 0$
b	$5, 3$	$4, 2$	$3, 1$

(a) Does either player have a dominant strategy? Explain briefly (1-3 sentences).

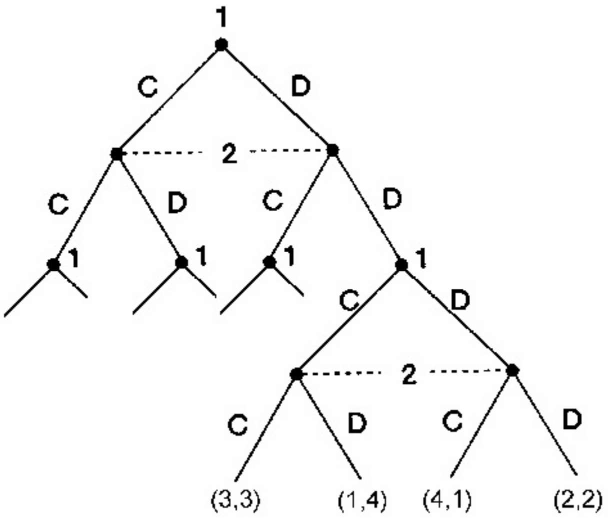
(b) Find all pure strategy Nash equilibria for this game.

Repeated Games: Motivation

- Many strategic interactions in which we are involved are ongoing:
  - we *repeatedly* interact with the same people
  - In many such interactions we have the opportunity to “take advantage” of our co-players, but do not
- For instance,
  - We look after our neighbors’ house while they’re away, even if it is time-consuming for us to do so;
  - We may give money to friends who are temporarily in need
  - In P2P networks, we share our bandwidth with someone we do not know - *bittorrent*
  - We answer questions by others in the online forums/Q&A systems, e.g., *Yahoo! Answers, Stack Overflow*
- The theory of **Repeated Games** provides a framework that we can use to study such behavior

Repeated Games

- In a Repeated Game, a (simultaneous- move) normal-form game is played over and again by the same players.
  - For example, the players may play a Prisoner's Dilemma game for 10 time periods
  - The game that is repeated in each period is referred to as the *stage game*



The “threat” of “punishment”

- The equilibria of a Repeated Game can differ from those of the associated stage game.
- There are many new strategies,
  - Players are able to condition their play on the history of play in previous rounds
- The basic idea, however, is that
  - a player may be deterred from exploiting her short-term advantage by *the “threat” of “punishment”* that reduces her long-term payoff

Repeated games: The Prisoner's Dilemma

- Suppose that two people are involved repeatedly in an interaction

  the short-term incentives are captured by the *Prisoner's Dilemma*, with payoffs as

	C	D
C	$2, 2$	$0, 3$
D	$3, 0$	$1, 1$

where C as "cooperation" (don't confess) and D as "defection" (confess)

An analysis

	C	D
C	$2,2$	$0,3$
D	$3,0$	$1,1$

- Recall the *Prisoner's Dilemma* has a unique Nash equilibrium, in which each player chooses *D*
- Now suppose that a player adopts the following long-term strategy (called *grim trigger strategy*):
  - choose *C* so long as the other player chooses *C*;
  - if in any period the other player chooses *D*, then choose *D* in *every subsequent period*
- What should the other player do?
  - she is better off choosing *C* in every period, as long as the value she attaches to future payoffs is not too small compared with the current payoffs

An analysis

	C	D
C	$2, 2$	$0, 3$
D	$3, 0$	$1, 1$

- If a player is *sufficiently patient*, the strategy that chooses *C* after every history is a *best response* to the strategy that <u>starts off choosing *C* and "punishes" any defection by switching to *D*</u>
- Why?
  - If she chooses *C* in every period then the outcome is (*C*, *C*) in every period and she obtains a payoff of 2 every period
  - If she switches to *D* in some period then she obtains a payoff of 3 in that period and a payoff of 1 in every subsequent period
  - However, as long as her future payoffs is not too small compared with her current payoff, the stream of payoffs (3,1,1,...) is worse for her than the stream (2,2,2,...)
    - so that she is better off choosing *C* in every period

Yet another Nash equilibrium

	C	D
C	$2,2$	$0,3$
D	$3,0$	$1,1$

- However, this strategy pair is not the only Nash equilibrium
- Another Nash equilibrium is the strategy pair in which each player chooses *D* after every history
- Why?
  - if one player adopts this strategy then the other player can do no better than to adopt the strategy herself, regardless of how she values the future,
  - since whatever she does has no effect on the other player's behavior.

Some open questions

- The outcome in which (C, C) occurs in every period is supported as a Nash equilibrium if the players are sufficiently patient
  – Exactly how patient do they have to be?
- We have seen also that the outcome in which (D, D) occurs in every period is supported as a Nash equilibrium
  – What other outcomes are supported?

Discounting

- We need discount future payoffs because:
  - The game might stop in any round with some probability $1 - \delta$
    - Based on this, the next time period will be reached with probability δ (delta), the one after that with probability δ^2, etc
  - Payoffs in the future are worth less due to *inflation* or because gains received now can be invested
  - In an Infinitely Repeated Game, cannot just define the utility as the sum over all periods as this would be unbounded in general

Discounting

- Each player assigns a higher weight to utility in the current period than in the future
  - Given a discount factor $\delta \in (0, 1)$, then if player *i* obtains utility u^t_i in time period *t*, then she values

    the utility obtained in period 0 as u^0_i ,

    the utility from period 1 as δu^1_i ,

    the utility from period 2 as $\delta^2 u^2_i$, and so on.
- For an Infinitely Repeated Game, player *i*'s total utility for history action $h = (a^0, a^1, a^2, \ldots)$ is

$$u_i(h) = \sum_{k=0}^{\infty} \delta^k u_i(a^k).$$

where $a^k = (a^k_1, a^k_2)$ is the set of actions for all the two players at time *k*

Automaton strategies

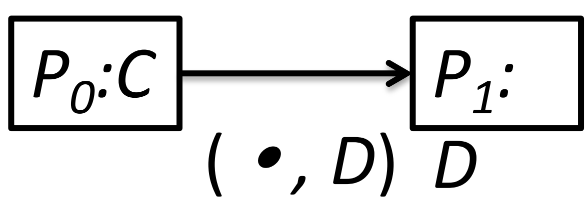
- A player's strategy specifies her action when it is her turn to move, given the history actions
- Recall a *grim trigger strategy* is a mode of behavior in which a defection by the other player triggers *relentless ("grim") punishment*: $s_i(\varnothing) = C$

and
$$s_i(a^1, \ldots, a^t) = \begin{cases} C & \text{if } a_j^\tau = C \text{ for } \tau = 1, \ldots, t \\ D & \text{otherwise.} \end{cases}$$

  - Player *i* chooses *C* at the start of the game (after the initial history Ø) and,
  - Player *j* also *chooses C* when every previous action of player *j* was *C*; she chooses *D* after every other history

Grim trigger strategy

- Need a compact way to represent a strategy
- Consider the strategy as having two *states*:
  - one, call it P_0, in which C is chosen, and
  - another, call it P_1, in which D is chosen
- Initially the state is P_0; if the other player chooses D in any period then the state changes to P_1, where it stays forever



Grim trigger strategy

- Suppose Player 1 adopts the grim trigger strategy
- If player 2 does so then the outcome is (*C*, *C*) in every period and she obtains the stream of payoffs (2, 2, . . .)
- The utility of player 2 is

$$u_2(h) = \sum_{t=0}^{\infty} \delta^t u(a_2^t) = \sum_{t=0}^{\infty} \delta^t 2 = 2/(1-\delta)$$

Grim trigger strategy

- If player 2 adopts a different strategy, there is at least one period in which she chooses *D*
- Player 2's choice of *D* triggers the punishment; in all subsequent periods player 1 chooses *D*
- As such player 2 also chooses *D* in every subsequent period. Why?
  - since *D* is her unique best response to *D*

action of player 2: C..D,D,... payoff 2,...,3,1,...

(Starting *D* earlier would have a better utility)

$$u_2(h)' = \sum_{t=0}^{\infty} \delta^t u(a_2^t) = 3 + \delta \ + \delta^2 \ldots = 2 + \frac{1}{1-\delta} = \frac{3-2\delta}{1-\delta}$$

Condition that Grim trigger strategy is a Nash equilibrium

- Thus player 2 cannot increase her payoff by deviating if and only if

$$u_2(h) \geq u_2(h)'$$

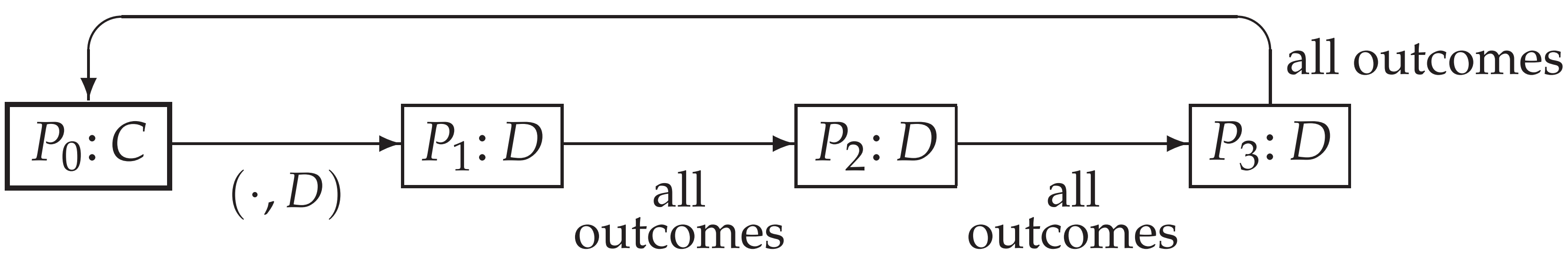
- We thus have $\dfrac{2}{1-\delta} \geq \dfrac{3-2\delta}{1-\delta} \Rightarrow \delta \geq \dfrac{1}{2}$

- So if $\delta \geq 1/2$ then the *grim trigger strategy* is a Nash equilibrium of the Infinitely Repeated *Prisoner's Dilemma*

	C	D
C	$2,2$	$0,3$
D	$3,0$	$1,1$

limited punishment strategy

- This strategy punishes deviations for only three periods:
  - it responds to a deviation by choosing the action *D* for three periods, then reverting to *C*, no matter how the other player behaved during her punishment



A general case for limited punishment

	C	D
C	$2, 2$	$0, 3$
D	$3, 0$	$1, 1$

- Rather than 3 periods, a player who chooses *D* is punished for *k* periods
- If one player adopts this strategy, is it optimal for the other to do so?
- To understand, suppose player 1 adopts the limited punishment and let us see
  - whether player 2 wants to deviate and,
  - If yes, in what condition

A general case: formulation

	C	D
C	$2,2$	$0,3$
D	$3,0$	$1,1$

- Suppose player 2 chooses deviate, say *D* in the first period
- Then player 1 chooses *D* in each of the next *k* periods, regardless of player 2's choices
  - player 2 also should choose *D* in these periods, why?

  Player 1: C D D … D

  Player 2: D D D … D
- In the (k + 1)st period after the deviation player 1 switches back to *C* (regardless of player 2's behavior in the previous period)

  Player 1: C D D … D C

  Player 2: D D D … D
- Again Player 2 faces precisely the same situation that she faced at the beginning of the game. Let us say she is back to be 'nice' again

  Player 1: C D D … D C C …

  Player 2: D D D … D C C …
- So if her deviation increases her payoff, it increases her payoff during the first k + 1 periods, which is

$$3+\delta + \delta^2 + \ldots + \delta^k = 2 + \frac{1-\delta^{k+1}}{1-\delta}$$

Why the equality hold?

A general case: formulation

	C	D
C	$2,2$	$0,3$
D	$3,0$	$1,1$

- If Player 2 does not deviate during the first $k + 1$ periods,

  Player 1: C C C … C

  Player 2: C C C … C
- her payoff (utility) is

$$2+2\delta \;+2\delta^{2}+\ldots+2\delta^{k}=\frac{2(1-\delta^{k+1})}{1-\delta}$$

- Thus she cannot increase her payoff by deviating *if and only if*

$$\frac{2(1-\delta^{k+1})}{1-\delta}\geq 2+\frac{1-\delta^{k+1}}{1-\delta}$$

- Which leads to

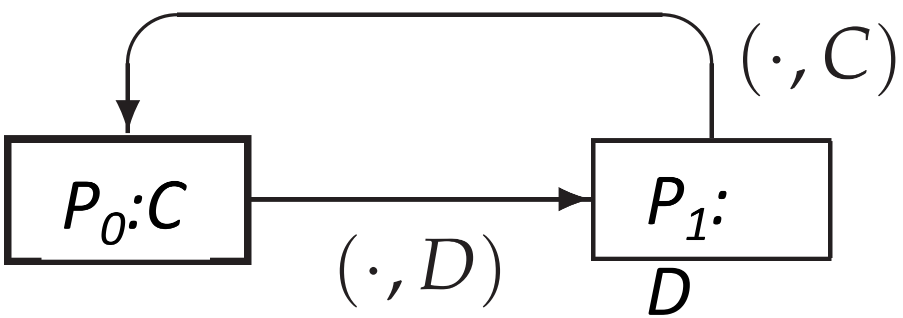
$$\delta^{k+1}-2\delta+1\leq 0$$

A general case: formulation

- No deviation condition $\delta^{k+1} - 2\delta + 1 \leq 0$ suggests that
  - If *k = 1* then no value of δ less than 1 satisfies the inequality:
    - one period of punishment is not severe enough to discourage a deviation, not matter how patient the players are.
  - If $k = 2$ then the inequality is satisfied for $\delta \geq 0.62$, and
  - if $k = 3$ it is satisfied for $\delta \geq 0.55$.
  - As k increases the lower bound on δ approaches 1/2 , the lower bound for the grim strategy.

Tit-for-tat strategy

- In strategy *tit-for-tat*, the length of the punishment depends on the behavior of the player being punished
  - If she continues to choose *D* then *tit-for-tat* continues to do so;
  - if she reverts to *C* then *tit-for-tat* reverts to *C* also.



Condition that Tit-for-tat strategy is a Nash equilibrium

- Suppose that player 1 adheres to this strategy starting with C
- Then, if player 2 can gain by deviating then she can gain by choosing *D* in the first period. Why?

  Player 1: C

  Player 2: D
- If she does so, player 1 chooses *D* in the second period, and continues to choose *D* until player 2 reverts to *C*

  Player 1: C D

  Player 2: D
- Thus player 2 has two options:
  - revert to *C*, in which case in the next period she faces the same situation as she did at the start of the game, or
  - continue to choose *D*, in which case player 1 will continue to do so too
- If she alternates between *D* and *C* then her stream of payoffs is (3, 0, 3, 0, . . .) with $3/((1-\delta)(1+\delta))$
- if she chooses *D* in every period her stream of payoffs is (3, 1, 1, . . .), with utility $3-2\delta/1-\delta$

Tit-for-tat strategy

- Since the player 2 utility of adhering to the strategy *tit-for-tat* is $2/(1-\delta)$,
- we conclude that *tit-for-tat* is a best response to *tit-for-tat* *if and only if*

$$2/(1-\delta) \geq 3-2\delta/1-\delta \text{ and } 2/(1\text{-}\delta) \geq 3/((1-\delta)(1+\delta))$$

- Both of these conditions are equivalent to

$$\delta \geq ½$$

the Iterated Prisoners' Dilemma (IPD) tournament

- Robert Axelrod in *The Evolution of Cooperation (1984)*, reported a tournament with
  - the N-step prisoners' dilemma (with N fixed).
- Researchers all over the world are invited to devise computer strategies to compete
  - programs that were entered varied widely
- When with many players, each with different strategies, *greedy strategies* tended to do very poorly while more *altruistic* strategies did better
- The winning deterministic strategy was *tit for tat*

References and Further Readings

- Osborne, Martin J. An introduction to game theory. Vol. 3, no. 3. New York, NY: Oxford University Press, 2004.

COMP0124 Multi-agent Artificial Intelligence

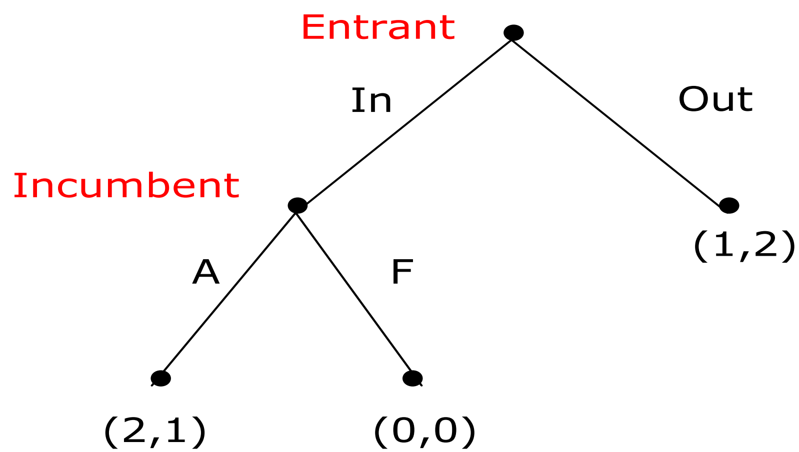
Extensive Form Games

Prof. Jun Wang
Computer Science, UCL

Extensive Form Games

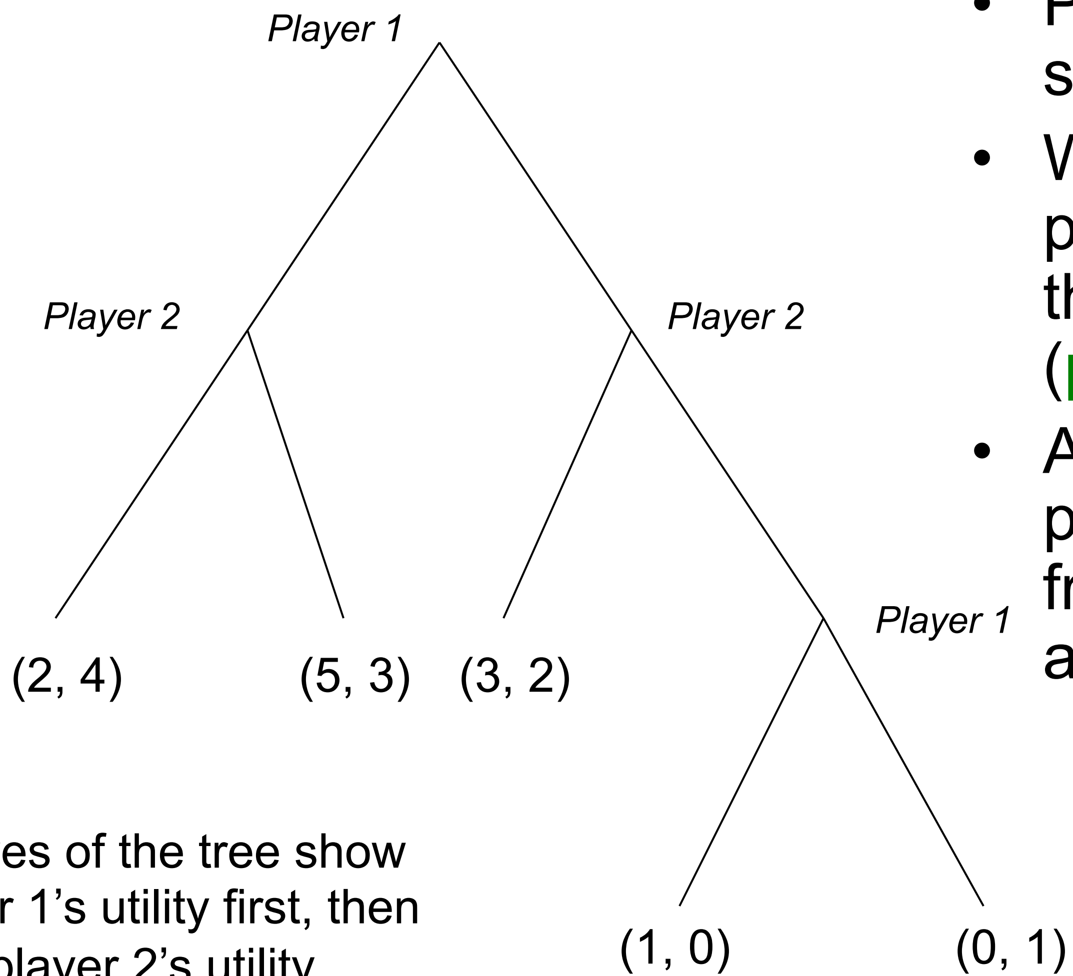
- Strategic form games -> model one-shot games in which each player chooses his action once and for all simultaneously.
- Extensive form games -> multi-agent sequential decision making
- Our focus will be on multi-stage games with observed actions where:
  - All previous actions are observed, i.e., each player is perfectly informed of all previous events.
  - Some players may move simultaneously at some stage k.
- Extensive form games can be conveniently represented by game trees.
- Additional component of the model, histories (i.e., sequences of action profiles).

Example: Entry Game



- There are two players.
  - Player 1, the entrant, can choose to enter the market or stay out.
  - Player 2, the incumbent, after observing the action of the entrant, chooses to accommodate him or fight with him.
- The payoffs for each of the action profiles (or histories) are given by the pair (x,y) at the leaves of the game tree:
  - x denotes the payoff of player 1 (the entrant) and y denotes the payoff of player 2 (the incumbent)

Extensive form games with perfect information



Leaves of the tree show player 1's utility first, then player 2's utility

- Players do not move simultaneously
- When moving, each player is aware of all the previous moves (perfect information)
- A (pure) strategy for player i is a mapping from player i's nodes to actions

Extensive Form Game Model

- A set of players, $N = \{1,\ldots,I\}$.
- Histories: A set H of sequences which can be finite or infinite.

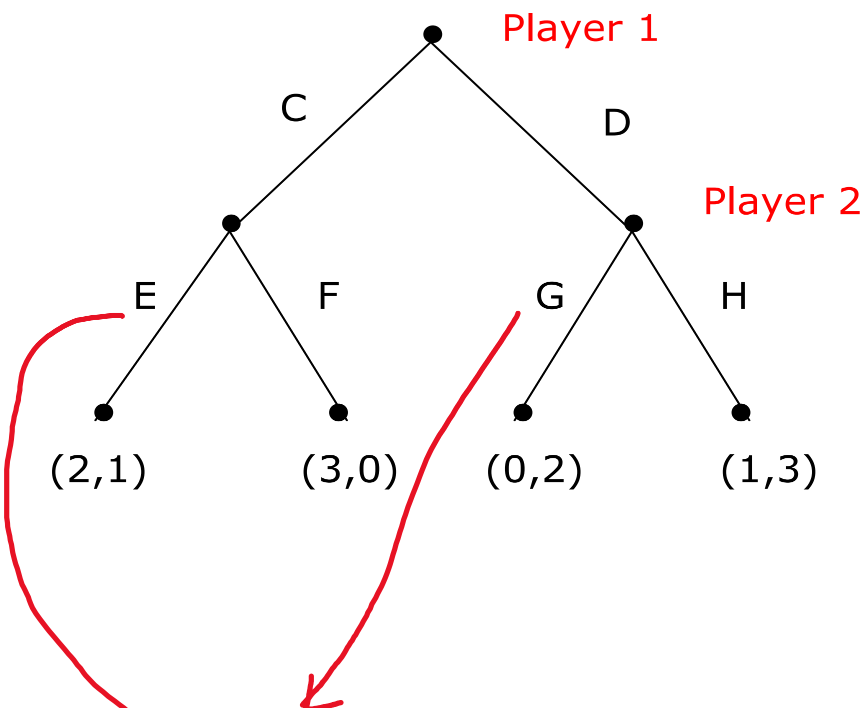
$h^0 = \emptyset$	initial history
$a^0 = (a_1^0, \ldots, a_I^0)$	stage 0 action profile
$h^1 = a^0$	history after stage 0
… …	
$h^{k+1} = (a^0,a^1,\ldots,a^k)$	history after stage k

- If the game has a finite number (K + 1) of stages, then it is a finite horizon game.
- Let $H^k = \{h^k\}$ be the set of all possible stage k histories.
- Then H^{K+1} is the set of all possible terminal histories, and
- $H = \bigcup_{k=0}^{K+1} H^k$ is the set of all possible histories.

Extensive Form Game Model

- Pure strategies for player i is defined as a **contingency plan** for every possible history h^k.
  - Let $S_i(H^k) = \cup_{h^k \in H^k} S_i(h^k)$ be the set of actions available to player i at stage k.
  - Let s_i^k: $H^k \rightarrow S_i(H^k)$ such that $s_i(h^k) \in S_i(h^k)$.
  - Then the pure strategy of player i is the set of sequences $s_i = \{s_i^k\}_{k=0}^{K}$, i.e., a pure strategy of a player is a collection of maps from all possible histories into available actions.
  - Observe that the path of strategy profile s will be $a^0 = s^0(h^0)$, $a^1 = s^1(a^0)$, $a^2 = s2(a^0, a^1)$, and so on.
- Preferences/payoffs are defined on the outcome of the game H^{K+1}. We can represent the preferences of player i by a utility function $u_i : H^{K+1} \rightarrow R$. As the strategy profile s determines the path $a^0, \ldots, a^k$ and hence h^{K+1}, we will denote $u^i(s)$ as the payoff to player i under strategy profile s.

Strategies in Extensive Form Games



EG <- (E or F) and (G or H)

- Player 1's strategies:

$s_1 : H^0 = \emptyset \rightarrow S_1 = \{C,D\}$; two possible strategies: C,D

- Player 2's strategies:

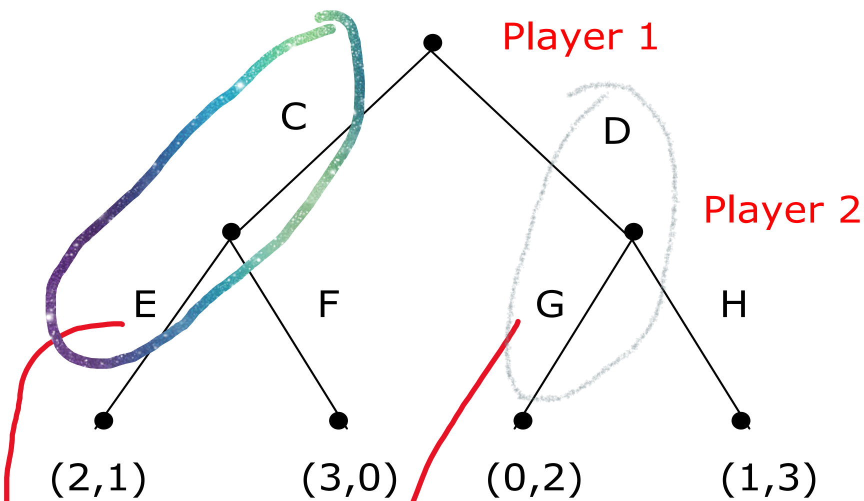
$s_2 : H^1 = \{\{C\},\{D\}\} \rightarrow S_2$; four possible strategies: which we can represent as EG, EH, FG and FH.

- If s = (C,EG), then the outcome will be {C,E}. On the other hand, if the strategy is s = (D,EG), the outcome will be {D,G}.

A player's strategy specifies the action the player chooses for **every** history (e.g., either C or D) after which it is her turn to move.

Osborne, Martin J. An introduction to game theory, 2004.

Strategies in Extensive Form Games



- Player 1's strategies:

$s_1 : H^0 = \emptyset \rightarrow S_1 = \{C,D\}$; two possible strategies: C,D

- Player 2's strategies:

$s_2 : H^1 = \{\{C\},\{D\}\} \rightarrow S_2$; four possible strategies: which we can represent as EG, EH, FG and FH.

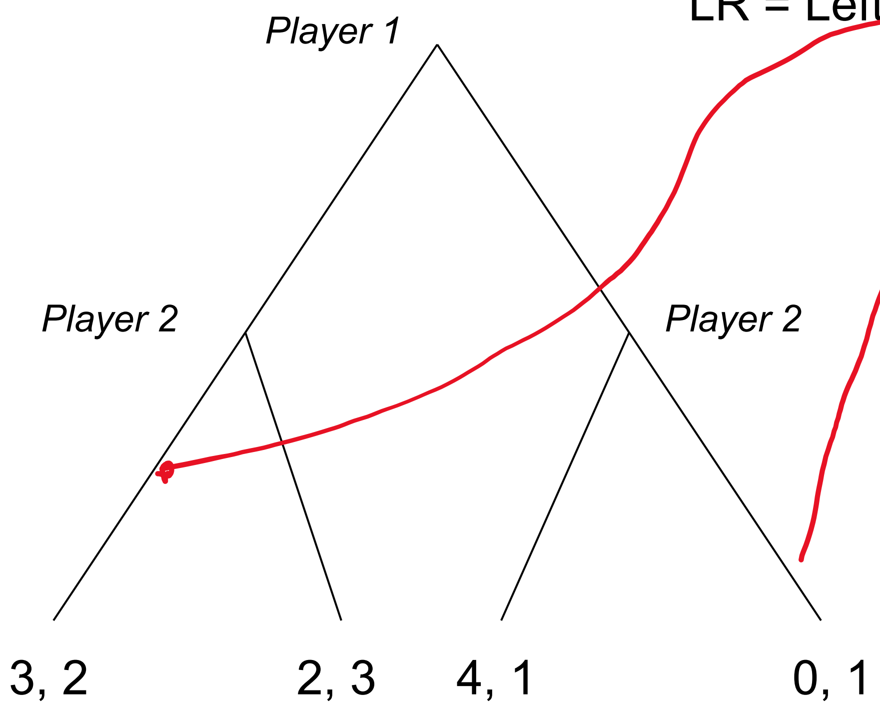
- If s = (C,EG), then the outcome will be {C,E}. On the other hand, if the strategy is s = (D,EG), the outcome will be {D,G}.

EG <- (E or F) and (G or H)

Player 2

Player 1		EG	EH	FG	FH
	C	**(2,1)**	**(2,1)**	**(3,0)**	**(3,0)**
	D	**(0,2)**	**(1,3)**	**(0,2)**	**(1,3)**

Conversion from extensive to normal form

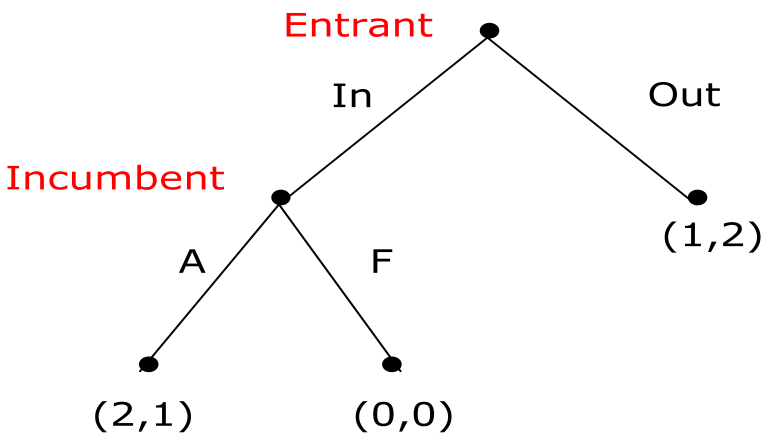


LR = Left if 1 moves Left, Right if 1 moves Right; etc.

	LL	LR	RL	RR
L	3, 2	3, 2	2, 3	2, 3
R	4, 1	0, 1	4, 1	0, 1

- Nash equilibria of this normal-form game include (R, LL), (R, RL), (L, RR) + infinitely many mixed-strategy equilibria
- In general, normal form can have exponentially many strategies

But Are These Equilibria Reasonable?



- Equivalent strategic form representation:

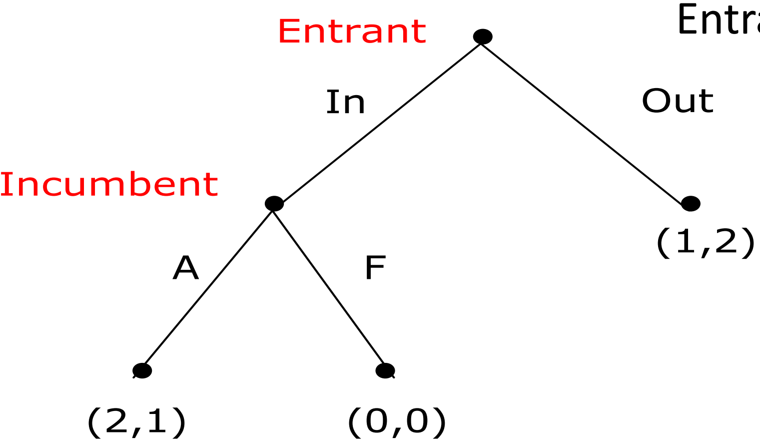
Incumbent

Entrant

	Accommodation	Fight
In	**(2,1)**	**(0,0)**
Out	**(1,2)**	**(1,2)**

- Two pure Nash equilibria: (In,A) and (Out,F).

But Are These Equilibria Reasonable?

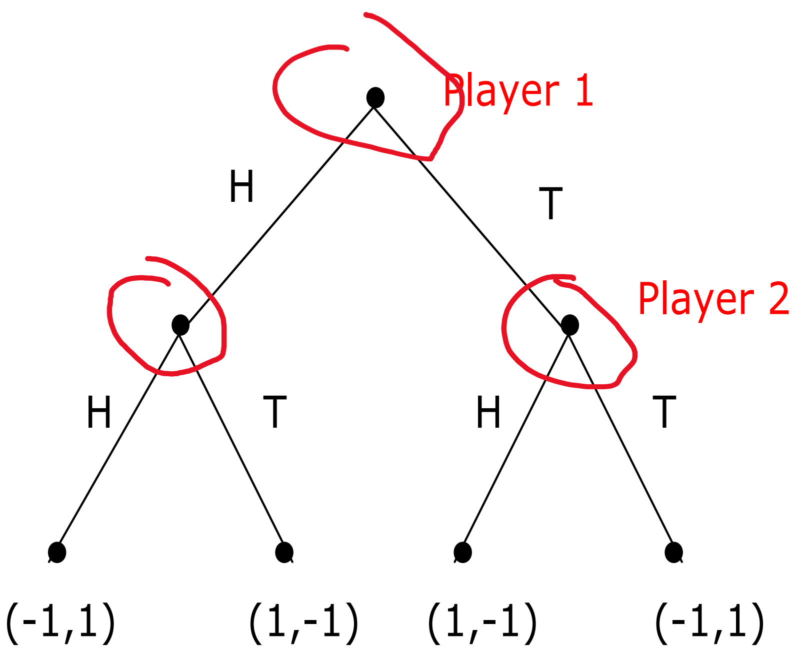


Entrant \Incumbent

	Accommodatio n	Fight
In	**(2,1)**	**(0,0)**
Out	**(1,2)**	**(1,2)**

- The equilibrium (Out,F) is sustained by a **noncredible** threat of the monopolist (Incumbent)
  - Entrant choosing Out is due to the treat of F from Incumbent, however, F is not credible as Incumbent won't choose it as A is better off (2>0)
- Equilibrium notion for extensive form games: **Subgame Perfect (Nash) Equilibrium**.
  - It requires each player's strategy to be "optimal" not only at the start of the game, but also after every history.
  - For finite horizon games, found by **backward induction**.

Subgames: Examples

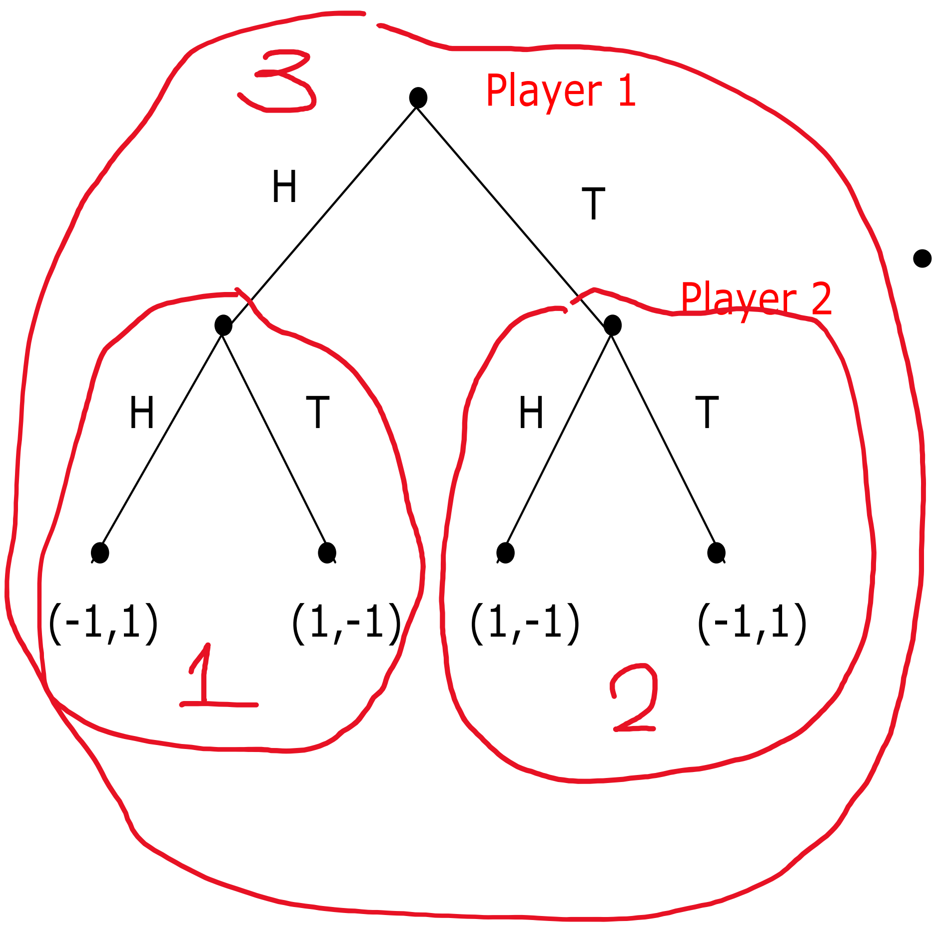


- Recall the two-stage extensive-form version of the matching pennies game
- In this game, there are two proper subgames and the game itself which is also a subgame, and thus a total of three subgames.

Each node in a (perfect-information) game tree, together with the remainder of the game after that node is reached, is called a subgame

i.e., each non-terminal history h has a subgame

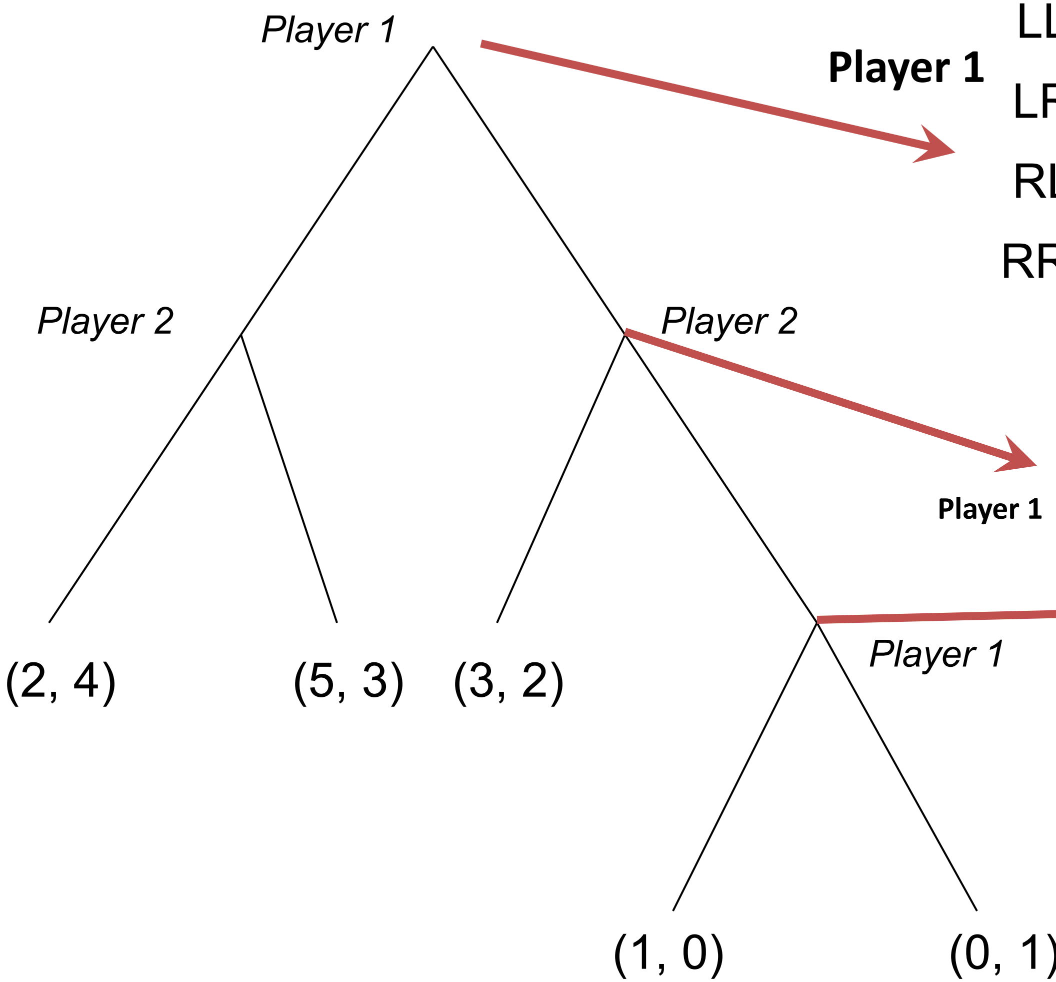
Subgames: Examples



- Recall the two-stage extensive-form version of the matching pennies game
- In this game, there are two proper subgames and the game itself which is also a subgame, and thus a total of three subgames.

Subgame perfect equilibrium

- Each node in a (perfect-information) game tree, together with the remainder of the game after that node is reached, is called a subgame
- A strategy profile is a subgame perfect equilibrium if it is an equilibrium for every subgame



Player 1 \ Player 2	LL	LR	RL	RR
LL	2, 4	2, 4	5, 3	5, 3
LR	2, 4	2, 4	5, 3	5, 3
RL	3, 2	1,0	3,2	1, 0
RR	3, 2	0,1	3,2	0, 1

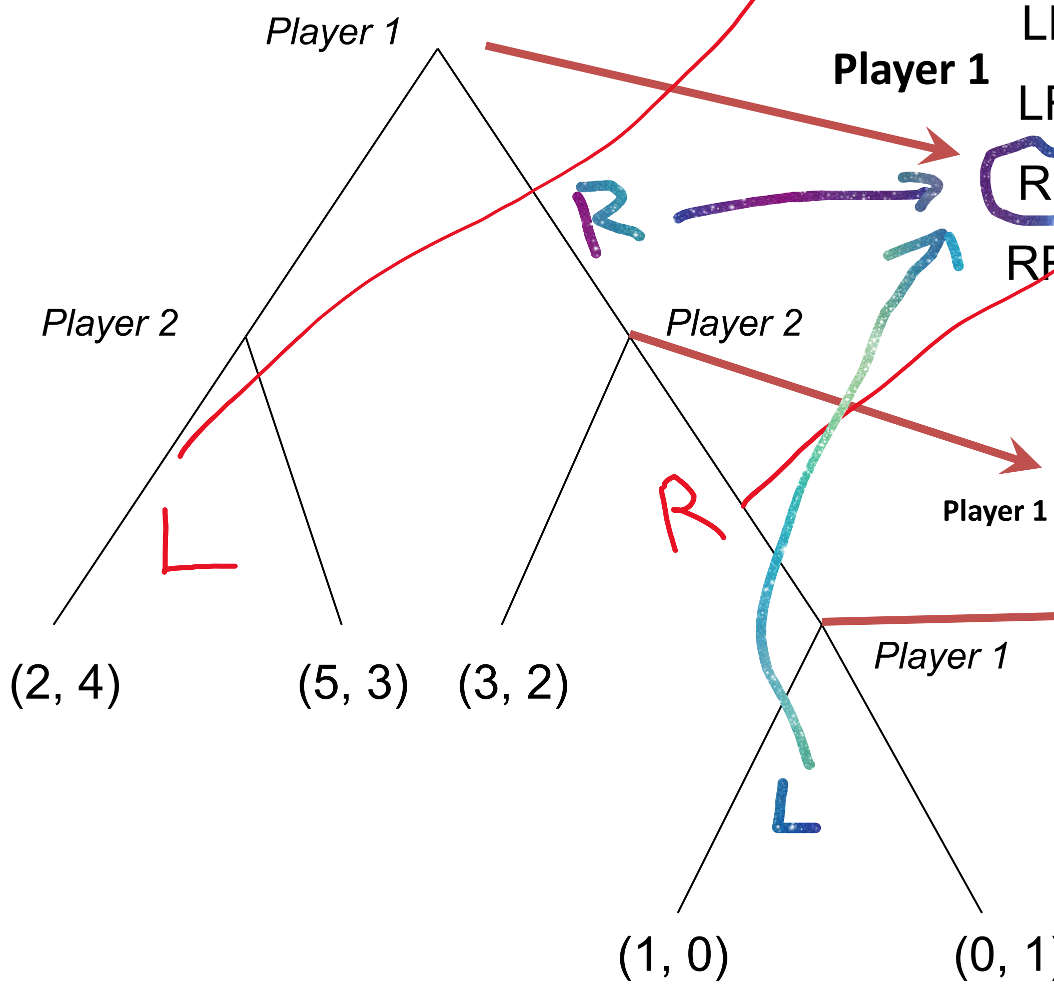
Player 1 \ Player 2	*L	*R
*L	3, 2	1, 0
*R	3, 2	0, 1

Player 1 \ Player 2	**
*L	1, 0
*R	0, 1

- (RR, LL) and (LR, LR) are not subgame perfect equilibria because (*R, **) is not an equilibrium
- (LL, LR) is not subgame perfect because (*L, *R) is not an equilibrium
  - *R is not a credible threat

Subgame perfect equilibrium

- Each node in a (perfect-information) game tree, together with the remainder of the game after that node is reached, is called a subgame
- A strategy profile is a subgame perfect equilibrium if it is an equilibrium for every subgame



Player 1 \ Player 2	LL	LR	RL	RR
LL	2, 4	2, 4	5, 3	5, 3
LR	2, 4	2, 4	5, 3	5, 3
RL	3, 2	1,0	3,2	1, 0
RR	3, 2	0,1	3,2	0, 1

Player 1 \ Player 2	*L	*R
*L	3, 2	1, 0
*R	3, 2	0, 1

Player 1 \ Player 2	**
*L	1, 0
*R	0, 1

- (RR, LL) and (LR, LR) are not subgame perfect equilibria because (*R, **) is not an equilibrium
- (LL, LR) is not subgame perfect because (*L, *R) is not an equilibrium
  - *R is not a credible threat

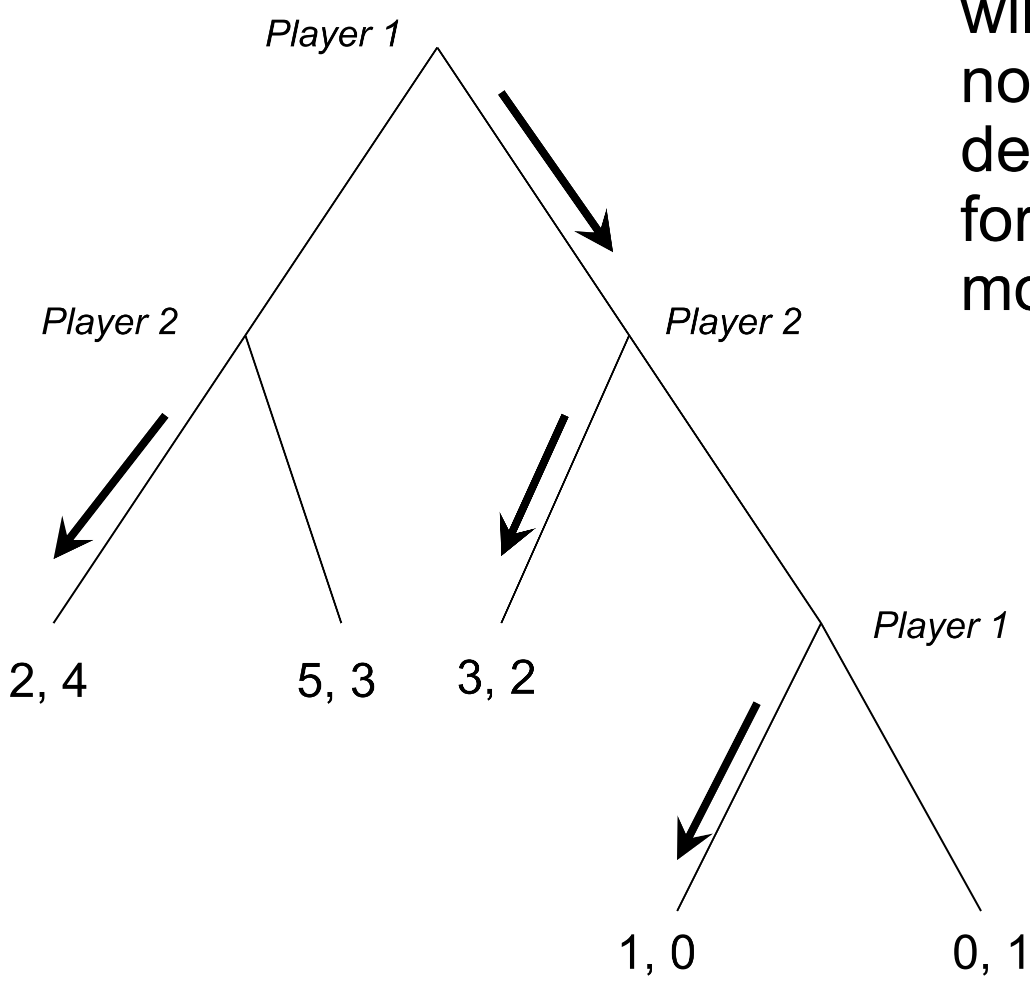
Subgame Perfect Equilibrium

- Loosely speaking, **subgame perfection** will remove noncredible threats, since these will not be Nash equilibria in the appropriate subgames.
- In the entry game, following entry, F is not a best response, and thus not a Nash equilibrium of the corresponding subgame. Therefore, (Out,F) is not a SPE.
- How to find SPE? One could find all of the Nash equilibria, for example as in the entry game, then eliminate those that are not subgame perfect.
- But there are more economical ways of doing it.

Backward Induction

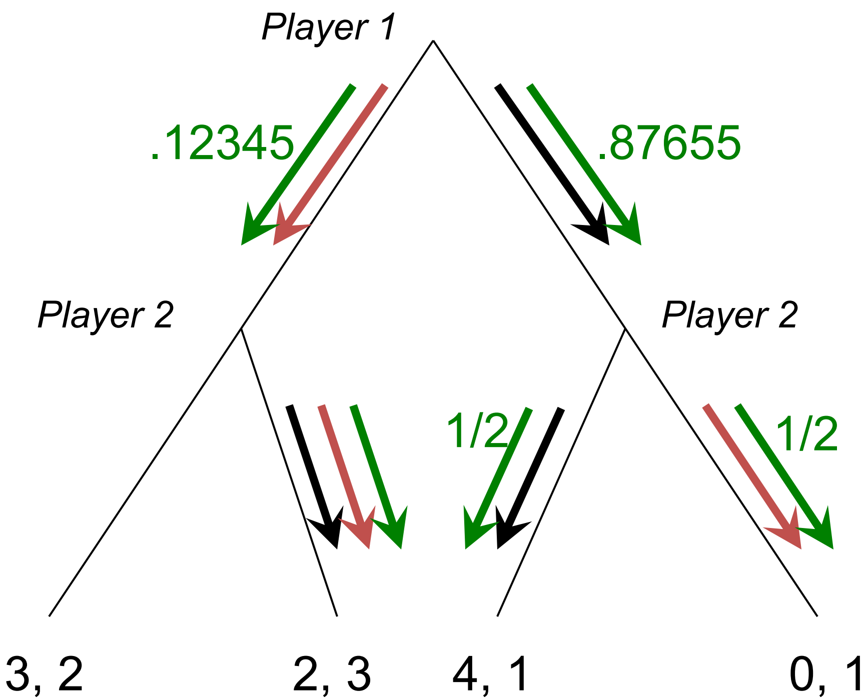
- **Backward induction** refers to starting from the last subgames of a finite game, then finding the best response strategy profiles or the Nash equilibria in the subgames, then assigning these strategies profiles and the associated payoffs to be subgames, and moving successively towards the beginning of the game.

Backward induction



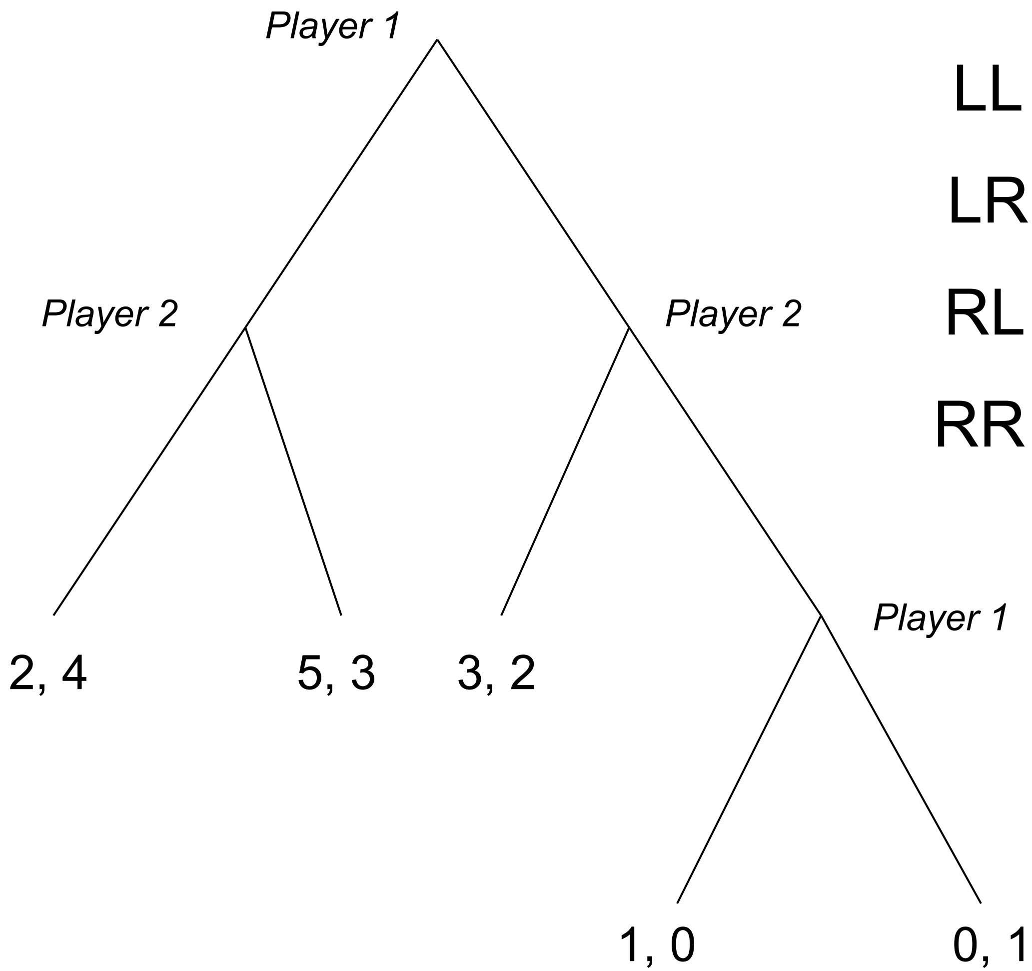
- When we know what will happen at each of a node’s children, we can decide the best action for the player who is moving at that node

A limitation of backward induction



- If there are ties, then how they are broken affects what happens higher up in the tree
- Multiple equilibria…

Conversion from extensive to normal form



	LL	LR	RL	RR
LL	2, 4	2, 4	5, 3	5, 3
LR	2, 4	2, 4	5, 3	5, 3
RL	3, 2	1, 0	3, 2	1, 0
RR	3, 2	0, 1	3, 2	0, 1

- Pure-strategy Nash equilibria of this game are (LL, LR), (LR, LR), (RL, LL), (RR, LL)
- Normal form fails to capture some of the structure of the extensive form
- The only **backward induction solution** is (RL, LL)

References and Further Readings

- Osborne, Martin J. An introduction to game theory. Vol. 3, no. 3. New York, NY: Oxford University Press, 2004.
- Vincent Conitzer, Game Theory
- Asu Ozdaglar, Game Theory with Engineering Applications, 2010
- Multiagent Systems: Algorithmic, Game-theoretic and Logical Foundations chapter 5

COMP0124 Multi-agent Artificial Intelligence

Potential Games

Prof. Jun Wang
Computer Science, UCL

Revisit: Cournot Competition

- I number of firms choose quantity $q_i \in (0, \infty)$
- The payoff function for player i given by

  $u_i(q_i, q_{-i}) = q_i (P(Q) - c)$.
- We define a function

  $$\Phi(q_1, \dots, q_I) = \left(\prod_{i=1}^{I} q_i\right)(P(Q) - c) \qquad (1)$$
- Note that for all i firms and all $q_{-i} > 0$,

$u_i(q_i,q_{-i}) - u_i(q'_i,q_{-i}) > 0$ iff $\Phi(q_i,q_{-i}) - \Phi(q'_i,q_{-i}) > 0$, $\forall\, q_i,q'_i > 0$.

Revisit: Cournot Competition

- A function Φ satisfying (1) is called an ordinal potential, and a game that possesses an ordinal potential is called an ordinal potential game.
- Clearly, the pure-strategy equilibrium set of the Cournot game coincides with the pure-strategy equilibrium set of the game in which every firm's profit is given by Φ, i.e.,

$q=(q_i,q_{-i})$ is an equilibrium point if and only if for every i,

$\Phi(q_i,q_{-i}) - \Phi(q'_i,q_{-i}) > 0$, for any $q'_i > 0$

Potential Games

- A strategic form game is a potential game (such as ordinal potential game or exact potential game) if there exists a function $\Phi : S \to R$ such that $\Phi(s_i, s_{-i})$ gives information about $u_i(s_i, s_{-i})$ for each $i \in I$.
- If so, Φ is referred to as the potential function.
- The potential function has a natural analogy to "energy" in physical systems. It will be useful both for locating pure strategy Nash equilibria and also for the analysis of "myopic" - short sighted- dynamics.

Why called potential function?

- In physics, Φ is a potential function for $(u_1, u_2, \ldots, u_n)$ if

$$\frac{\partial u_i(q)}{\partial q_i} = \frac{\partial \Phi(q)}{\partial q_i} \quad \text{for all i} \in I.$$

Potential Functions and Games

- Let $G =< I, (S_i), (u_i) >$ be a strategic (matrix) form game

Definition

A function $\Phi : S \to R$ is called an ordinal potential function for the game G if for each $i \in I$ and all $s_{-i} \in S_{-i}$,

$$u_i(x,s_{-i}) - u_i(z,s_{-i}) > 0 \text{ iff } \Phi(x,s_{-i}) - \Phi(z,s_{-i}) > 0, \text{ for all } x,z \in S_i$$

G is called an ordinal potential game if it admits an ordinal potential.

Potential Functions and Games

- Let $G =< I, (S_i), (u_i) >$ be a strategic (matrix) form game

Definition

A function Φ : S → R is called an **exact** potential function for the game G if for each i ∈ I and all $s_{-i} \in S_{-i}$,

$u_i(x,s_{-i}) - u_i(z,s_{-i}) = \Phi(x,s_{-i}) - \Phi(z,_{s-i})$, for all x,z ∈ Si

G is called an **exact** potential game if it admits an ordinal potential.

Example

- A potential function assigns a real value for every s ∈ S.
- Thus, when we represent the game payoffs with a matrix (in finite games), e.g.,

$$G = \begin{pmatrix} (1,1) & (9,0) \\ (0,9) & (6,6) \end{pmatrix},$$

"Prisoner's dilemma"

- and we can also represent the potential function as a matrix, each entry corresponding to the vector of strategies from the payoff matrix.

$$P = \begin{pmatrix} 4 & 3 \\ 3 & 0 \end{pmatrix}.$$

Pure Strategy Nash Equilibria in Ordinal Potential Games

Theorem

Every finite ordinal potential game has at least one pure strategy Nash equilibrium.

- **Proof:** The global maximum of an ordinal potential function is a pure strategy Nash equilibrium. To see this, suppose that s∗ corresponds to the global maximum. Then, for any i ∈ I, we have, by definition, $\Phi(s_i^*, s_{-i}^*) - \Phi(s_i, s_{-i}^*) \geq 0$ for all $s_i \in S_i$. But since Φ is a potential function, for all i and all $s_i \in S_i$,

$$u_i(s_i^*, s_{-i}^*) - u_i\left(s_i, s_{-i}^*\right) \geq 0 \; iff \; \Phi(s_i^*, s_{-i}^*) - \Phi\left(s_i, s_{-i}^*\right) \geq 0$$

- Therefore, $u_i(s_i^*, s_{-i}^*) - u_i(s_i, s_{-i}^*) \geq 0$ for all $s_i \in S_i$ and for all i ∈ I. Hence s∗ is a pure strategy Nash equilibrium.
- Note, however, that there may also be other pure strategy Nash equilibria corresponding to local maxima.

Example: Cournot competition again

- Suppose now that P(Q) = a – bQ and costs $c_i(q_i)$ are arbitrary.
- We define a function

$$\Phi^*(q_1, \ldots, q_I) = a\sum_{i=1}^{I} q_i - b\sum_{i=1}^{I} q_i^2 - b\sum_{1\le i<m\le I}^{I} q_i q_m - \sum_{i=1}^{I} c_i(q_i)$$

- It can be shown that for all i and all q_{-i},

$u_i(q_i,q_{-i})-u_i(q'_i,q_{-i}) = \Phi^*(q_i,q_{-i})-\Phi^*(q'_i,q_{-i})$, for all $q_i,q'_i > 0$.

Simple Dynamics in Finite Ordinal Potential Games

Definition

A path in strategy space S is a sequence of strategy vectors $(s_0, s_1, \cdots)$ such that every two consecutive strategies differ in one coordinate (i.e., exactly in one player's strategy).

An improvement path is a path $(s_0, s_1, \cdots)$ such that,

- $u_{i^k}(s^k) < u_{i^k}(s^{k+1})$ where s^k and s^{k+1} differ in the i^kth coordinate. In other words, the payoff improves for the player who changes his strategy.

An improvement path can be thought of as generated dynamically by "myopic players".

Simple Dynamics in Finite Ordinal Potential Games

Proposition

In every finite ordinal potential game, every improvement path is finite.

Proof: Suppose $(s_0, s_1, \cdots)$ is an improvement path. Therefore we have,

$$\Phi(s^0) < \Phi(s^1) < \cdots ,$$

where Φ is the ordinal potential. Since the game is finite, i.e., it has a finite strategy space, the potential function takes on finitely many values and the above sequence must end in finitely many steps.

- This implies that in finite ordinal potential games, every "maximal" improvement path must terminate in an equilibrium point.
- That is, the simple myopic learning process converges to the equilibrium set.

Characterization of Finite Exact Potential Games

- For a finite path $\gamma = (s_0,\ldots,s_N)$, let

$$I(\gamma) = \sum_{i=1}^{N} u^{m_i}(s^i) - u^{m_i}(s^{i-1})$$

where m_i denotes the player changing its strategy in the i^{th} step of the path.

- The path $\gamma = (s^0,\ldots,s^N)$ is closed if $s^0 = s^N$. It is a simple closed path if in addition $s^l <> s^k$ for every $0 \le l <> k \le N-1$.

Theorem

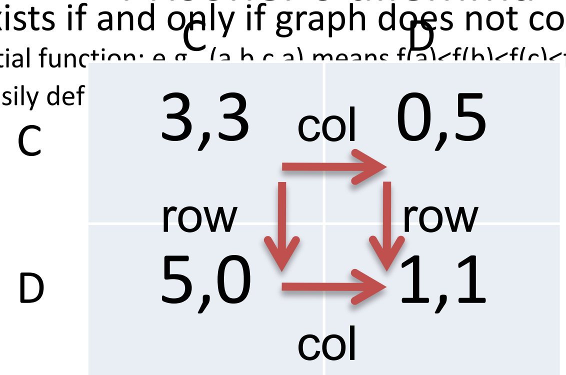
A game G is an exact potential game if and only if for all finite simple closed paths, γ, $I(\gamma) = 0$. Moreover, it is sufficient to check simple closed paths of length 4.

Monderer and Shapley, "Potential Games," Games and Economic Behavior, vol. 14, pp. 124-143, 1996.

Characterization of Finite Exact Potential Games

- Lemma: A game is a potential game if and only if local improvements always terminate
- proof:
  - Define a directed graph with a node for each possible pure strategy profile
  - Directed edge (u,v) means v (which differs from u only in the strategy of a single player, i) is a (strictly) better action for i, given the strategies of the other players

Prisoner's dilemma



  - A potential function exists if and only if graph does not contain cycles
    - If cycle exists, no potential function: e.g., (a,b,c,a) means $f(a)<f(b)<f(c)<f(a)$
    - If no cycles exist, can easily def

Examples

$\longrightarrow$ direction of local improvement

Prisoner's dilemma

$$
\begin{tikzcd}
 & C & D \\
C & 3,3 \arrow[r, "col"] \arrow[d, "row"'] & 0,5 \arrow[d, "row"] \\
D & 5,0 \arrow[r, "col"'] & 1,1
\end{tikzcd}
$$

Battle of the sexes

$$
\begin{tikzcd}
2,1 & 0,0 \arrow[l] \arrow[d] \\
0,0 \arrow[u] \arrow[r] & 1,2
\end{tikzcd}
$$

Matching pennies

$$
\begin{tikzcd}
-1,1 \arrow[d] & 1,-1 \arrow[l] \\
1,-1 \arrow[r] & -1,1 \arrow[u]
\end{tikzcd}
$$

Which are potential games?
Exact potential games?
Are the potential functions unique?

Coordination game

$$
\begin{tikzcd}
2,2 & 0,0 \arrow[l] \arrow[d] \\
0,0 \arrow[u] \arrow[r] & 3,3
\end{tikzcd}
$$

Infinite Potential Games

Proposition

Let G be a continuous potential game with compact strategy sets. Then G has at least one pure strategy Nash equilibrium.

Proposition

Let G be a game such that $S_i \subseteq R$ and the pay off functions $u_i : S \to R$ are continuously differentiable.

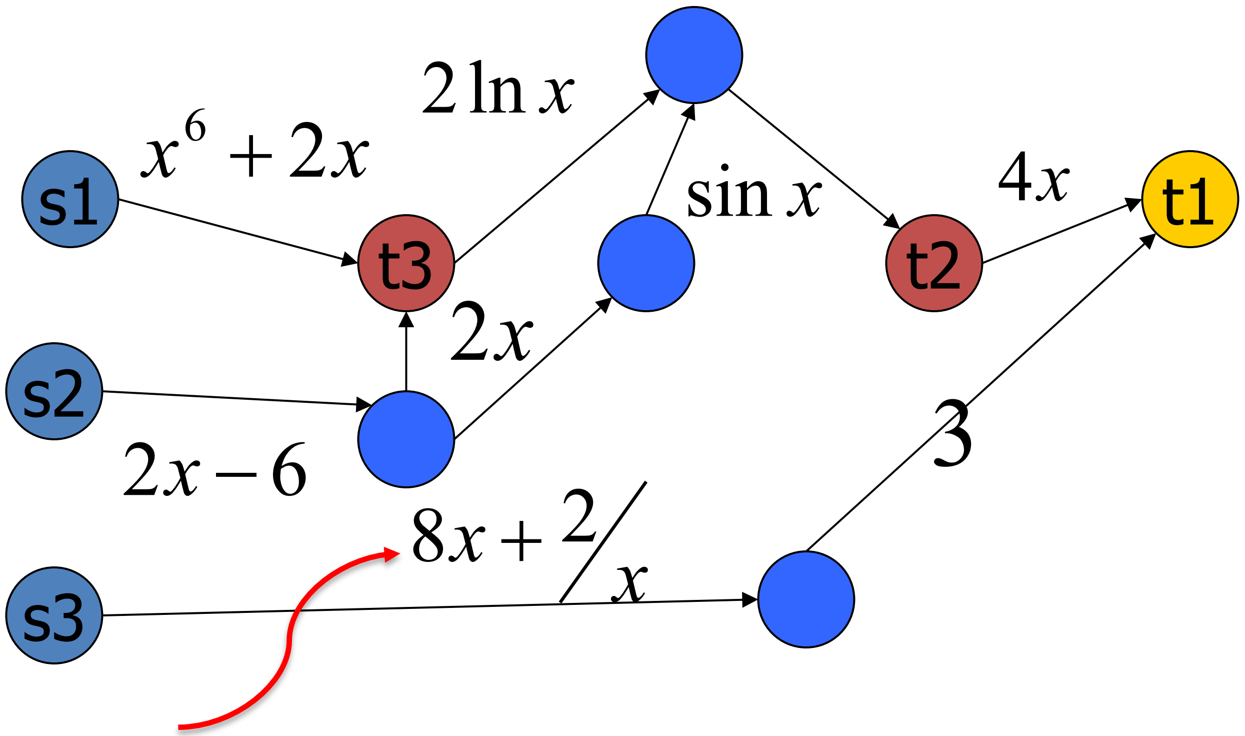
Let $\Phi : S \to R$ be a function. Then, Φ is a potential for G if and only if Φ is continuously differentiable and

$$\frac{\partial u_i(s)}{\partial s_i} = \frac{\partial \Phi(s)}{\partial s_i} \quad \text{for all } i \in I \text{ and all } s \in S.$$

Network Congestion Games

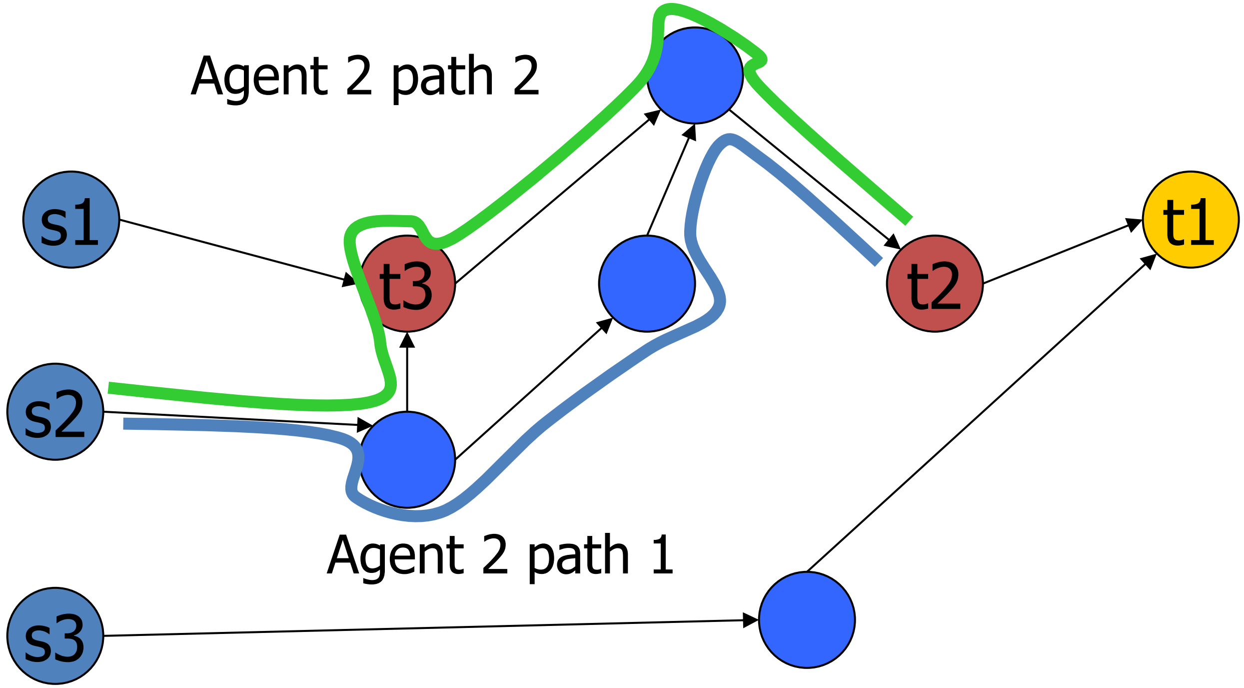
- A directed graph G=(V,E) with n users,
- Each edge e in E(G) has a delay function f_e,
- Strategy of user i is to choose a path A_j from a source s_i to a destination t_i,
- The delay of a path is the sum of delays of edges on the path,
- Each user wants to minimize his own delay by choosing the best path.

Example: Network Congestion Game

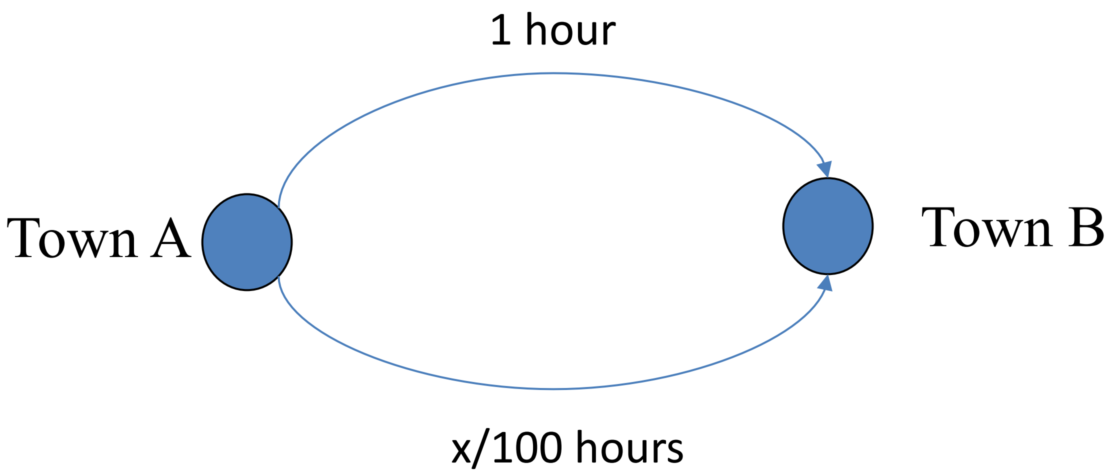


Delay function depending on the traffic,
e.g., the number of agents

Example: Network Congestion Game



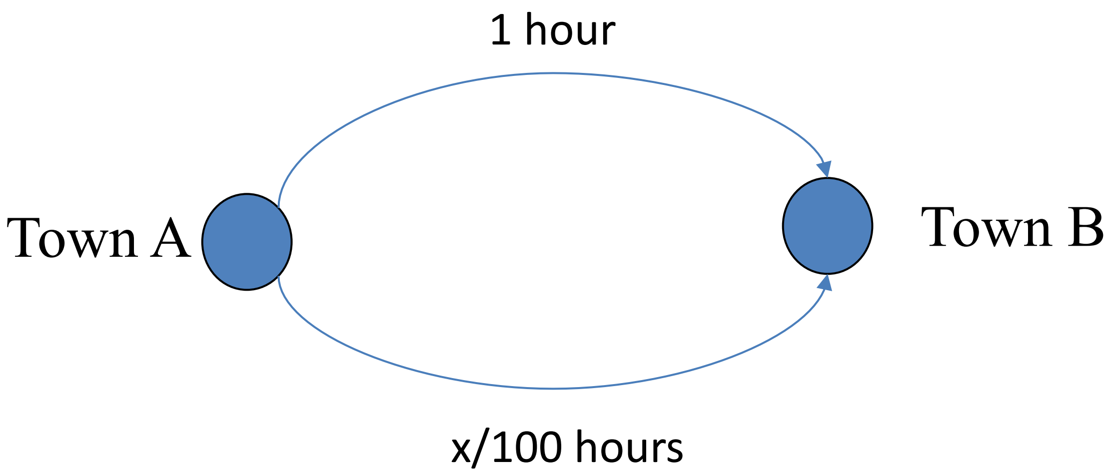
Traffic Routing



- Suppose 100 drivers leave from town A towards town B.
- Every driver wants to minimize her own travel time.
- What is the traffic on the network?

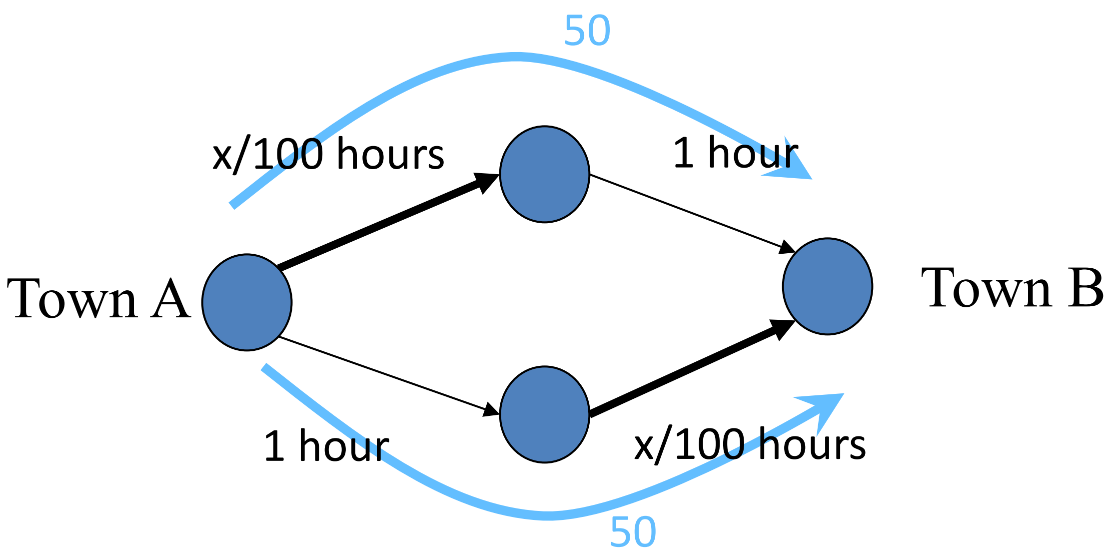
  - In any unbalanced traffic pattern, all drivers on the most loaded path have incentive to switch their path.

Traffic Routing



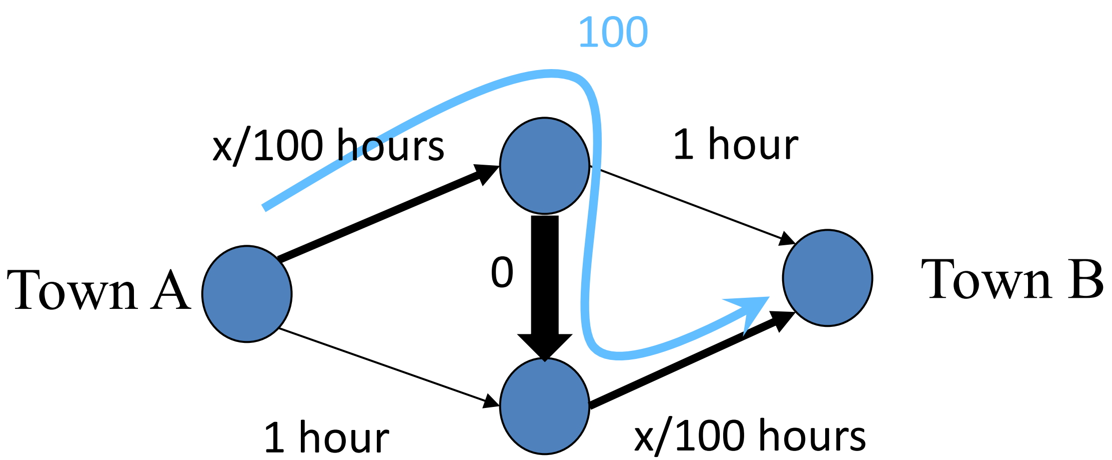
- If both paths have 50, average delay is 0.75 hours.
- In a NE, every one goes bottom. Average delay is 1 hour.
- NE leads to slower travel times !

Traffic Routing



Delay is 1.5 hours for everybody at the unique Nash equilibrium

Traffic Routing

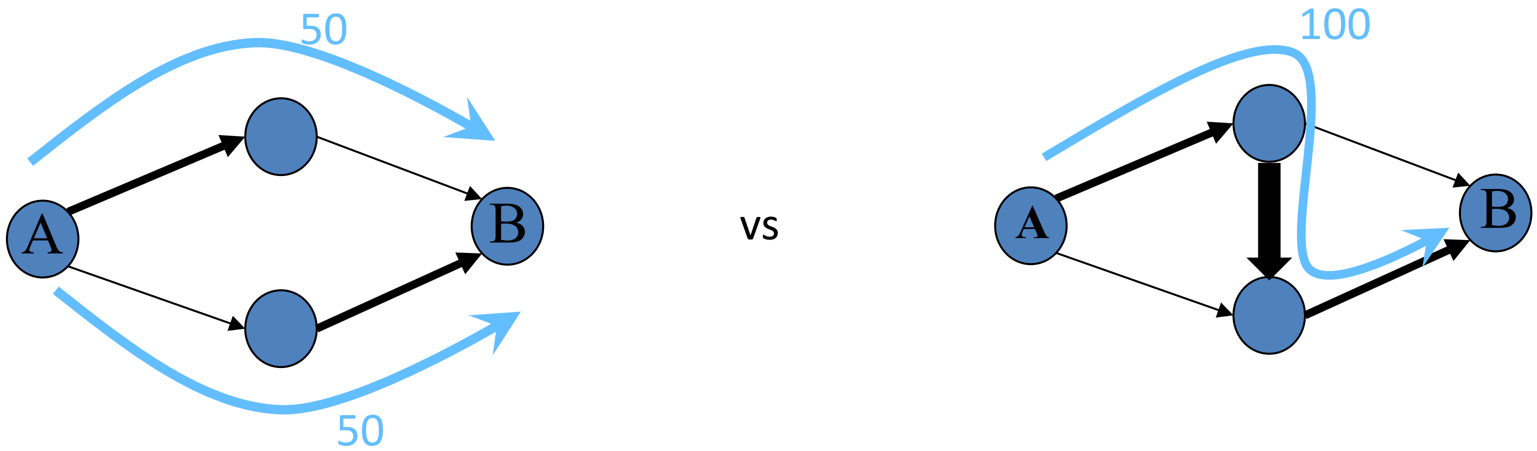


- A benevolent mayor builds a superhighway connecting the fast highways of the network.
- What is now the traffic on the network?

No matter what the other drivers are doing it is always better for me to follow the zig-zag path.

Delay is 2 hours for everybody at the unique Nash equilibrium.

Traffic Routing: Braess's paradox



- Adding a fast road on a road-network is not always a good idea!

- In the network left there exists a traffic pattern where all players have delay 1.5 hours.

- Question: How well can a Nash Equilibrium perform, compared to the optimal solution?

Congestion Games

Congestion Model: $C = \langle I, M, (S_i)_{i \in I}, (c_j)_{j \in M} \rangle$ where:

- $I = \{1, 2, \cdots, I\}$ is the set of players.
- $M = \{1, 2, \cdots, m\}$ is the set of resources.
- S_i is the set of resource combinations (e.g., links or common resources) that player i can take/use. A strategy for player i is $s_i \in S_i$, corresponding to the subset of resources that this player is using.
- $c^j(k)$ is the benefit (the negative of the cost) to each user who uses resource j if k users are using it.
- Define congestion game $\langle I, (S_i), (u_i) \rangle$ with utilities

$$u_i(s_i, s_{-i}) = \sum_{j \in s_i} c^j(k_j)$$

- where k_j is the number of users of resource j under strategy s.

Every congestion game is a potential game

Theorem (Rosenthal (73))

Every congestion game is a potential game and thus has a pure strategy Nash equilibrium.

- **Proof**: For each j define $\bar{k}_j^i$ as the usage of resource j excluding player i, i.e.,

$$\bar{k}_j^i = \sum_{i' \neq i} \mathbf{I}\left[j \in s_{i'}\right],$$

  where $\mathbf{I}\left[j \in s_{i'}\right]$ is the indicator for the event that $j \in s_{i'}$.
- With this notation, the utility difference of player i from two strategies s_i and s_i' (when others are using the strategy profile s_{-i}) is

$$u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}) = \sum_{j \in s_i} c^j(\bar{k}_j^i + 1) - \sum_{j \in s_i'} c^j(\bar{k}_j^i + 1).$$

Every congestion game is a potential game

- Now consider the function

j is from all the sets of resources $s_{i'}$ used by all players i'

$$\Phi(s) = \sum_{j \in \bigcup_{i' \in \mathcal{I}} s_{i'}} \left[\sum_{k=1}^{k_j} c^j(k) \right].$$

- We can also write

$$\Phi(s_i, s_{-i}) = \sum_{j \in \bigcup\limits_{i' \neq i} s_{i'}} \left[\sum_{k=1}^{\bar{k}_j^i} c^j(k) \right] + \sum_{j \in s_i} c^j(\bar{k}_j^i + 1).$$

Resources that are not used by i

Resources that are used by i

Every congestion game is a potential game

- Therefore:

$$
\begin{aligned}
\Phi(s_i, s_{-i}) - \Phi(s_i', s_{-i}) &= \sum_{j \in \bigcup_{i' \neq i} s_{i'}} \left[\sum_{k=1}^{\bar{k}_j^i} c^j(k) \right] + \sum_{j \in s_i} c^j(\bar{k}_j^i + 1) \\
&\quad - \sum_{j \in \bigcup_{i' \neq i} s_{i'}} \left[\sum_{k=1}^{\bar{k}_j^i} c^j(k) \right] + \sum_{j \in s_i'} c^j(\bar{k}_j^i + 1) \\
&= \sum_{j \in s_i} c^j(\bar{k}_j^i + 1) - \sum_{j \in s_i'} c^j(\bar{k}_j^i + 1) \\
&= u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}).
\end{aligned}
$$

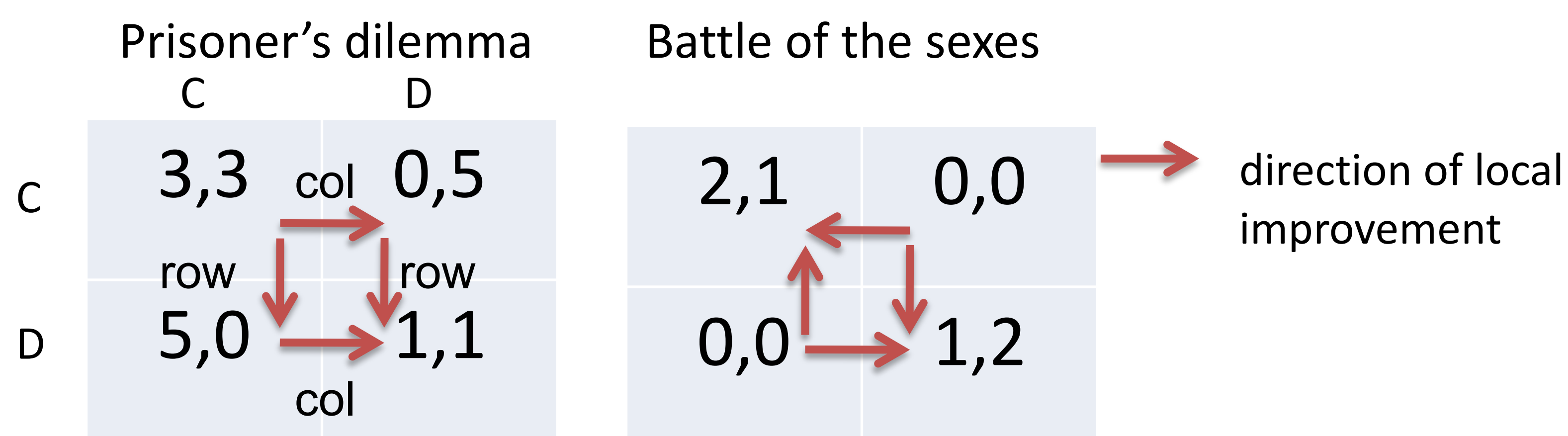
Thus every finite congestion game has a pure strategy (deterministic) equilibrium due to the property of potential game

The existence of a pure strategy Nash equilibrium

- Do a pure strategy NE always exist in traffic routing games?
- Given others' paths, the driver will choose a best path to minimize travel time.

  (best response dynamics)
- Aim to find a pure strategy Nash equilibrium (PSNE): start at some circumstance and perform best response dynamics iteratively.
- Will this process stop?

Theorem:
In a Congestion Game, any iterative best response process will terminate and eventually converge to a PSNE.



References and Further Readings

- Asu Ozdaglar, Game Theory with Engineering Applications, 2010
- Monderer and Shapley, “Potential Games,” Games and Economic Behavior, vol. 14, pp. 124-143, 1996.
- Michal Feldman, Inefficiency of equilibria, and potential games, 2008