# Multi-agent AI: Introduction

Prof. Jun Wang
Computer Science, UCL

# Some Observations on AI Progress

Decision Making

**"human-level"**


DeepMind

Deep Reinforcement Learning since ~2013

Bots Routine and Navigation

Game AI

Collaborative AI

Renewed Interests on Multiagent Reinforcement Learning since ~2017/2018

Perception (Recognition)

**"human-level"**

IMAGENET

SENSETIME 商汤科技

依图

Face++ 旷视

DEEPGLINT 格灵深瞳

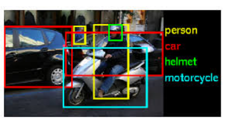
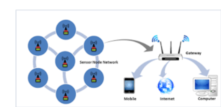



Image Classification (Computer Vision)

Deep Learning since ~2006

Swarm bots



Sensor networks

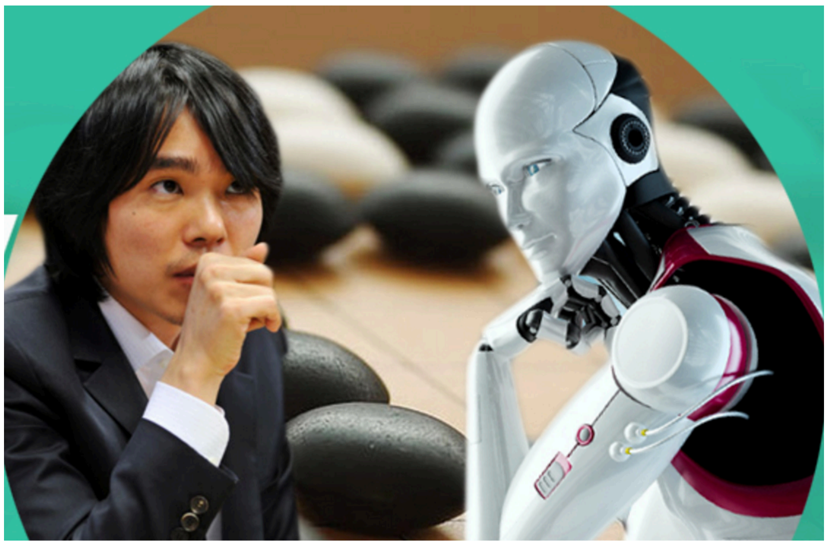
Swarm "Intelligence" since ~1989

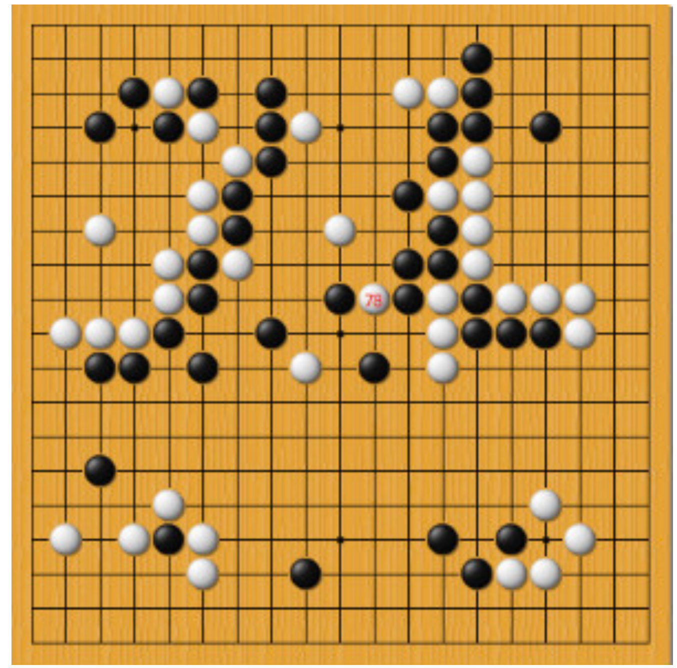
Single-agent AI

Multiagent AI systems

Google DeepMind

UCL PhD/Postdoc Demis Hassabis founded DeepMind and Google acquired it with 400M British Pounds in 2014

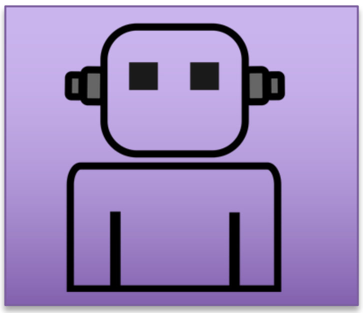




Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y. Mastering the game of Go without human knowledge. Nature. 2017 Oct;550(7676):354.

# Single-agent Reinforcement Learning

Markov decision processes: one decision maker with multiple states



Action $a_t$

Reward $r_{t+1}$

State $S_{t+1}$

An AI Agent

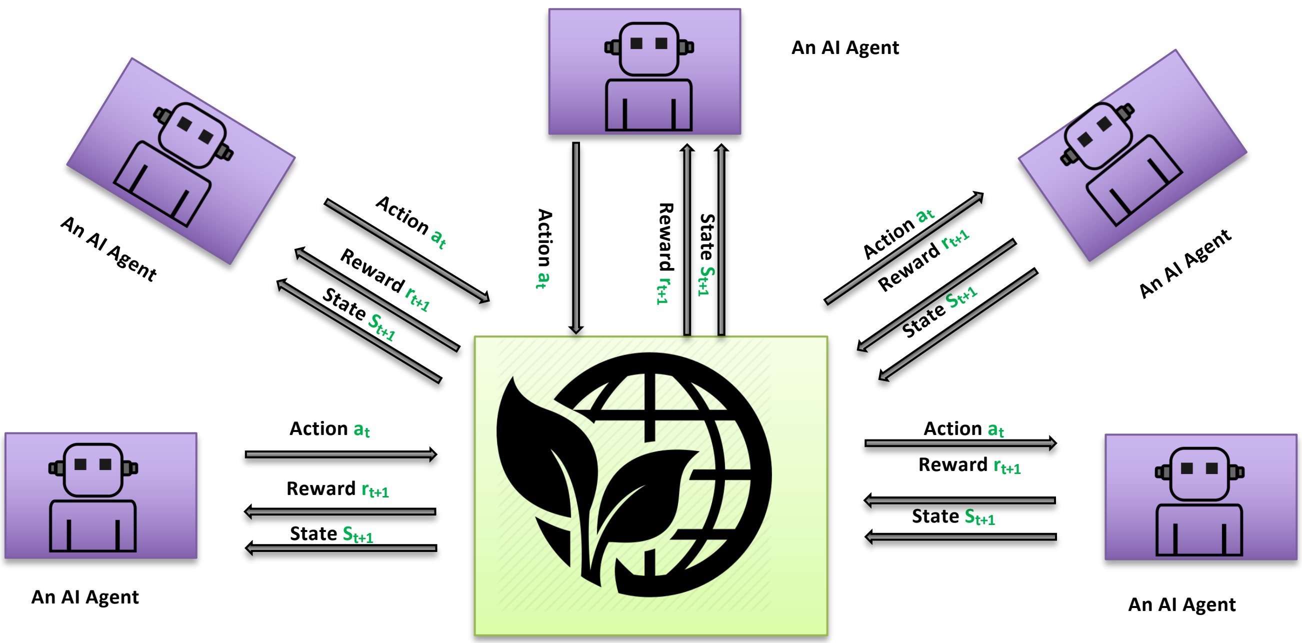
Environment

**Optimal action policy $a^*$ <--- Maximise $r_1 + r_2 + \ldots + r_t + \ldots$**

Source: https://media.giphy.com/media/ZCi0C4OTlVtPa/giphy.gif

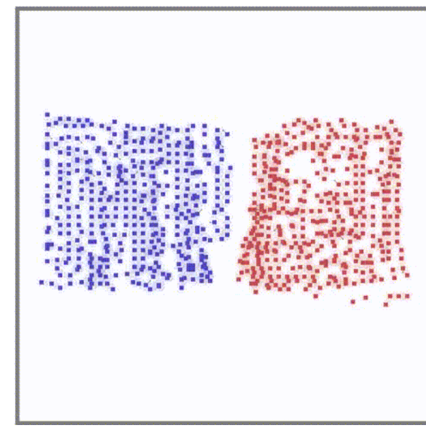
# Multi-agent Reinforcement Learning

Stochastic games (Markov games): multiple decision makers with multiple states (e.g., multiple normal form games)



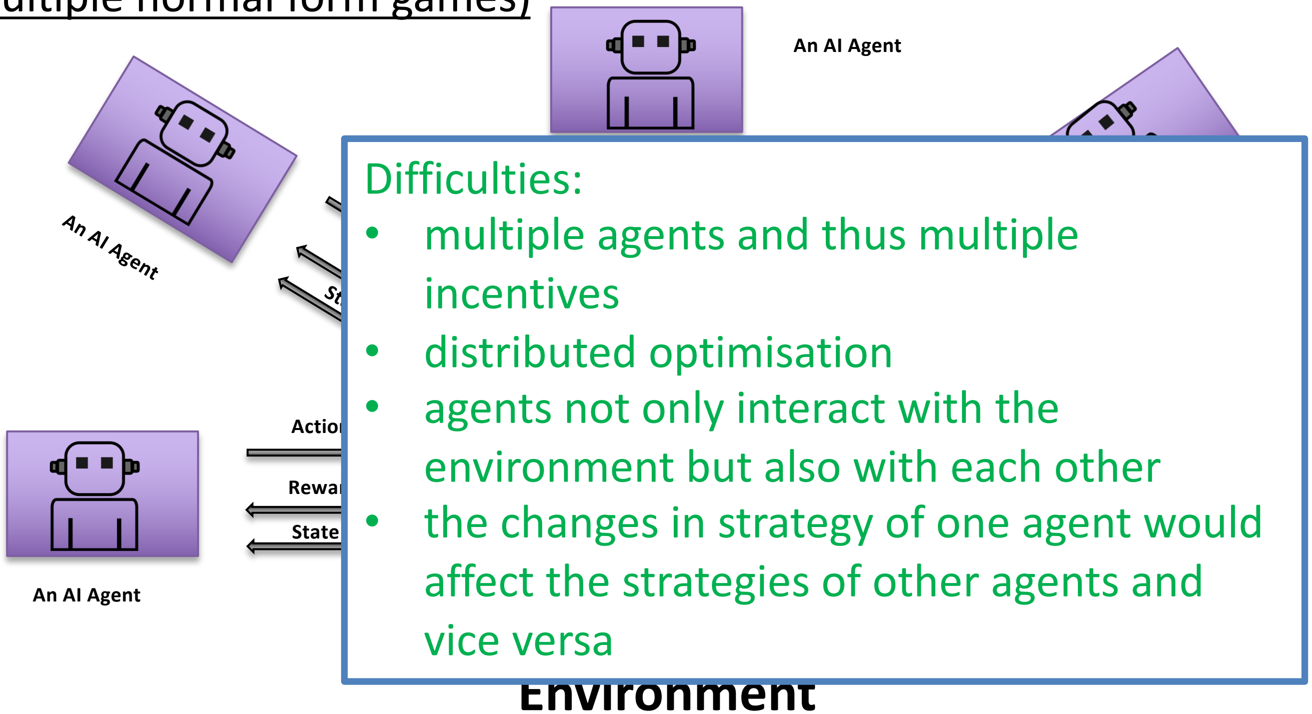

**Environment**

**A set of autonomous agents that share a common environment**

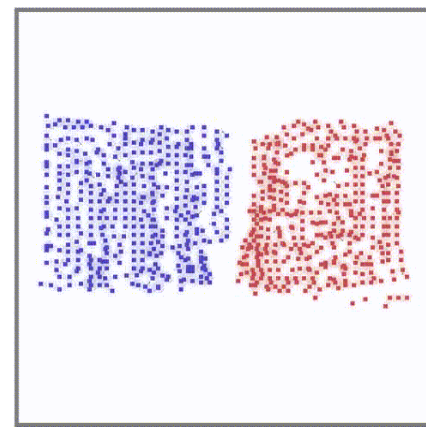
# Multi-agent Reinforcement Learning

Stochastic games (Markov games): multiple decision makers with multiple states (e.g., multiple normal form games)




Difficulties:
- multiple agents and thus multiple incentives
- distributed optimisation
- agents not only interact with the environment but also with each other
- the changes in strategy of one agent would affect the strategies of other agents and vice versa

**A set of autonomous agents that share a common environment**

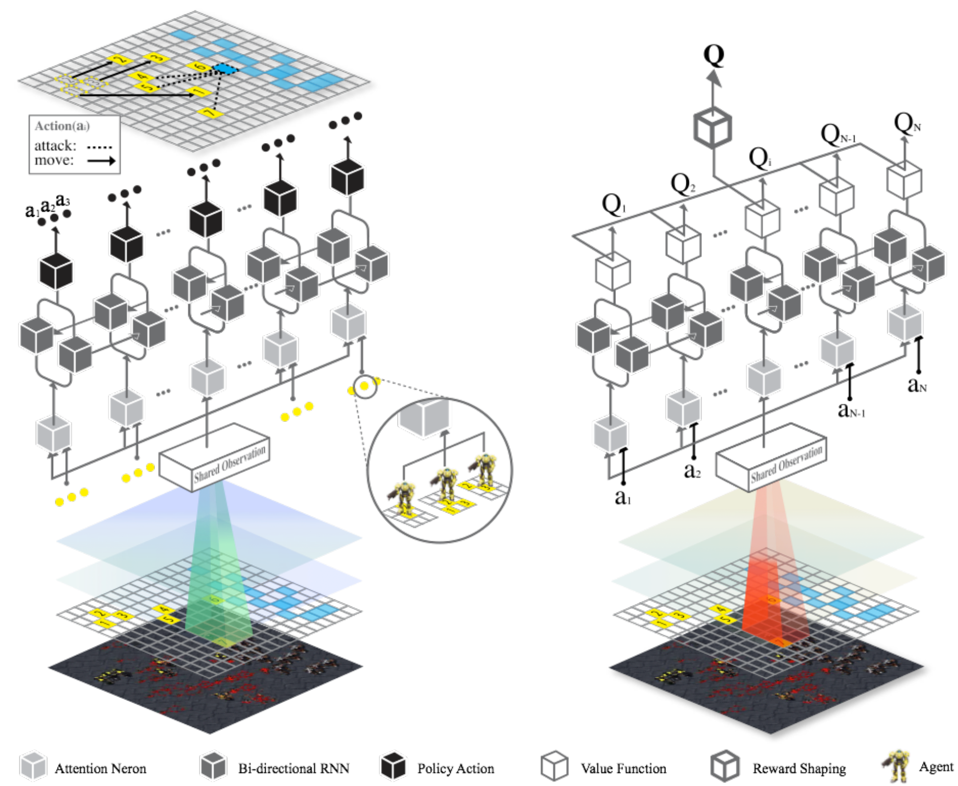
Peng P, Yuan Q, Wen Y, Yang Y, Tang Z, Long H, Wang J. Multiagent Bidirectionally-Coordinated Nets for Learning to Play StarCraft Combat Games. NIPS17 Emergent Communication Workshop. (**StarCraft AI beating Facebook methods**)

# Multiagent AI

- We look at AI systems that interact with each others and human beings, where they have different incentives and economic constraints
- We focus on learning problems in the above settings.

# Multiagent AI






(a) Multiagent policy networks with grouping (b) Multiagent Q networks with reward shaping

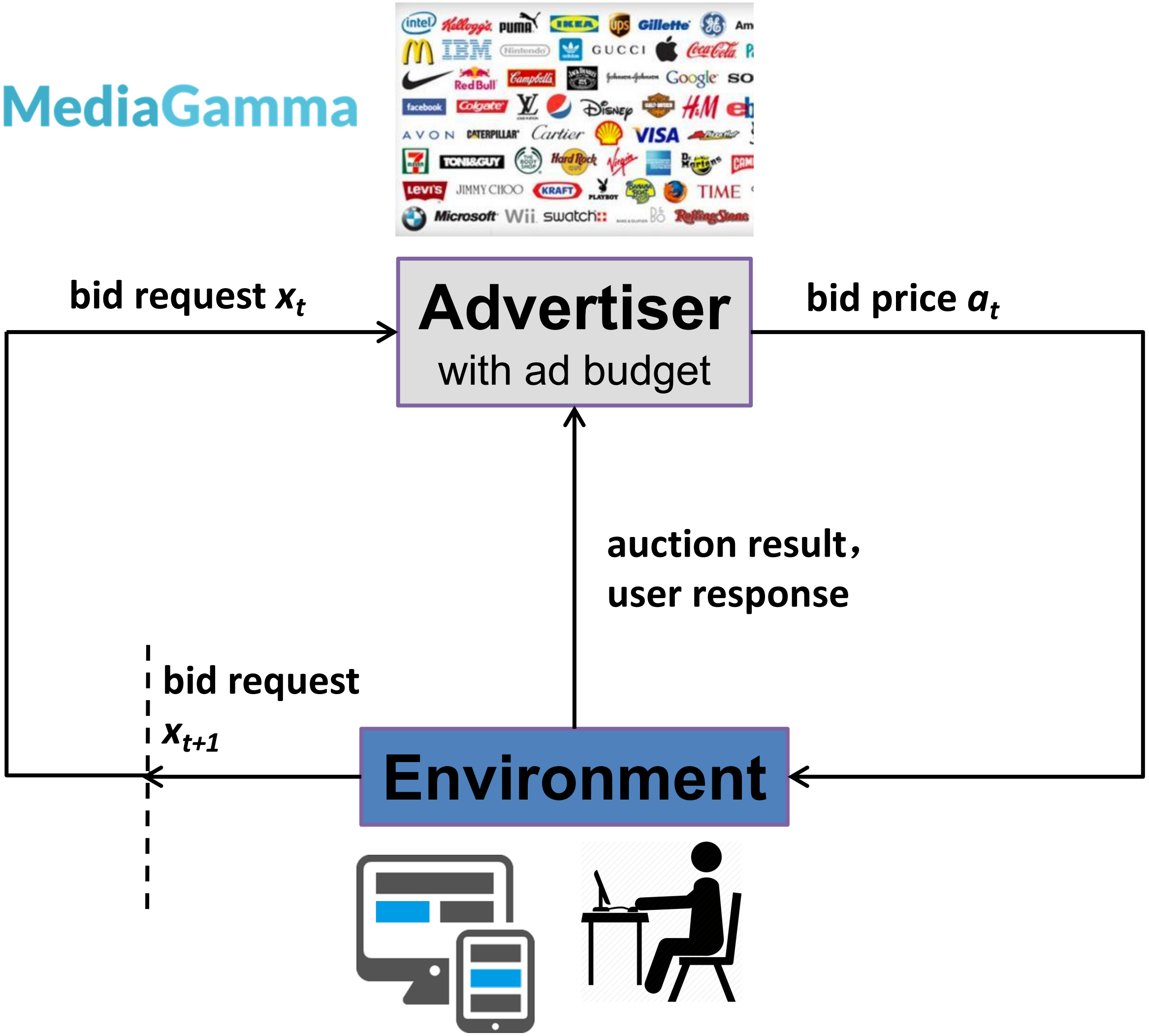
https://www.youtube.com/watch?v=kW2q15MNFug

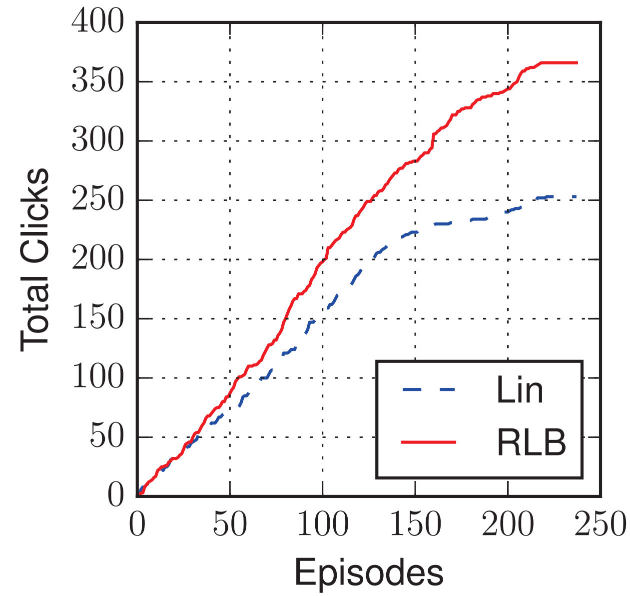
http://v.youku.com/v_show/id_XMjcyMTE0MDkwNA%3D%3D.html

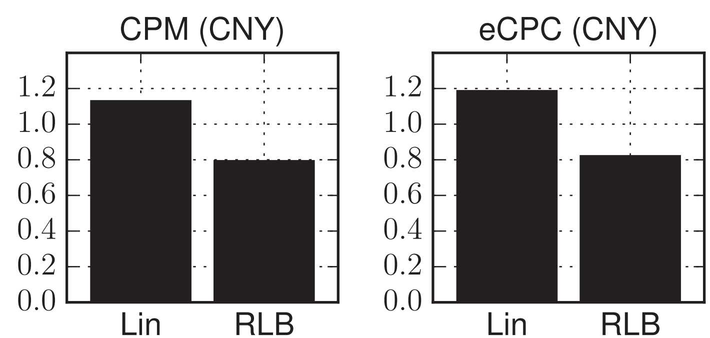
Peng P, Yuan Q, Wen Y, Yang Y, Tang Z, Long H, Wang J. Multiagent Bidirectionally-Coordinated Nets for Learning to Play StarCraft Combat Games. NIPS17 Emergent Communication Workshop. (**StarCraft AI beating Facebook methods**)

# MARL Application: AI in Online Advertising

MediaGamma



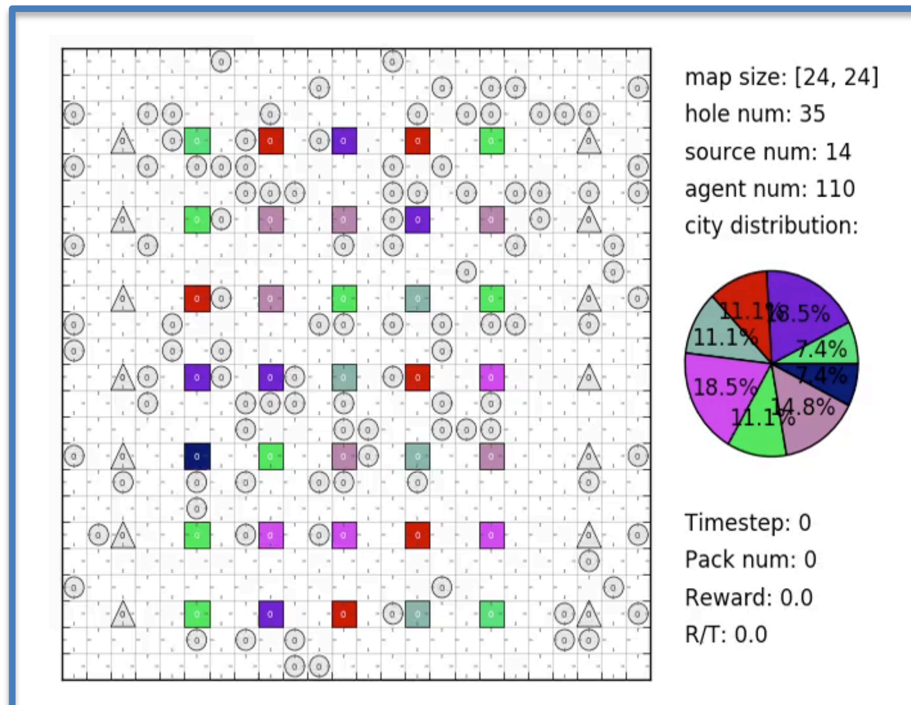

The goal is to maximise the user responses on displayed ads

Cai, H., K. Ren, W. Zhag, K. Malialis, and J. Wang. "Real-Time Bidding by Reinforcement Learning in Display Advertising." In *The Tenth ACM International Conference on Web Search and Data Mining (WSDM)*. ACM, 2017.

# MARL Application: planning

A learnable solution rather than existing rule based planning solutions

- Routing optimisation
- Collision avoidance
- Bot environment designs


Zhang et al, Two-layer Evolution with Fitness Approximation for Warehouse Layout Design, 2018

# Content

- Lecture 1: Multiagent AI and basic game theory
- Lecture 2: Potential games, and extensive form and repeated games
- Lecture 3: Solving (“Learning”) Nash Equilibria
- Lecture 4: Bayesian Games, auction theory and mechanism design
- Lecture 5: Learning and deep neural networks
- Lecture 6: Single-agent Learning (1)
- Lecture 7: Multi-agent Learning (1)
- Lecture 8: Single-agent Learning (2)
- Lecture 9: Multi-agent Learning (2)
- Lecture 10: Multi-agent Learning: advanced topics (3)

**Subject to change**

COMP0124 Multi-agent Artificial Intelligence

# Game Theory

Prof. Jun Wang
Computer Science, UCL

# Why Game Theory?

- The interplay between *economic* and *computational* considerations
  - In many systems, social and economic transactions mediated through *computation*, and,
  - Many systems are also *economic systems* in that they are operated by multiple, often self-interested, parties


Visit our eBay Store


**Game Theory (addressing incentive issues)**

# Computer Science vs. Economics

- ***Computer science***: the study of the representation and processing of information for the purpose of specific calculation tasks
  - understand what types of computation can be carried out efficiently, under time, resource and communication constraints.
- ***Economics***: the study of decision making by multiple parties, each with individual preferences, capabilities, and information, and motivated to act in regard to these own preferences
  - understand what outcomes can be achieved in systems with multiple parties (e.g., resource allocations, production decisions) under incentive constraints.

# Summary

- Game Theory
  - Prisoner's dilemma
  - Nash equilibrium
- Existence of Nash Equilibrium
- ISPs Game

# Game Theory

- Game Theory model situations in which multiple participants interact or affect each other's outcomes
  - In 1713, Waldegrave provides a *minimax mixed strategy solution* to a two-person version of the card game *le Her*
  - Game theory began by John von Neumann in 1928
  - More than 10 game-theorists have won the Nobel Memorial Prize in Economic Sciences

# Example (1/2)

- Suppose you can either study for the *exam* or the *presentation* for a course
  - Want to maximise your average grade (equal weight)
  - Your partner has to make a similar choice
- **Exam**
  - By studying for the exam, you get 92, else 80
- **Presentation**
  - If both prepare, each gets 100; if one prepares 92, but if neither prepares you both get 84.
- Can‘t communicate, how do you decide?
  - Your partner will also need to decide...

# Example (2/2)

0.5*80+0.5*100

Your partner

You

| | Presentation | Exam |
|---|---|---|
| Presentation | 90,90 | 86,92 |
| Exam | 92,86 | 88,88 |

Strategies

Payoff for
row player,column player

0.5*92+0.5*92

0.5*92+0.5*84

- What will happen in this game?

Each player thinks studying for exam is safer

# Reasoning about behavior in a game

- How are the players likely to behave?
- Need a few assumptions
  - The payoff for a player should capture all rewards
  - Players know all possible strategies and payoffs to other players
    - Not always true!
  - Players are *rational*
    - Each players wants to maximise her own payoff
    - Each player succeeds in picking the optimal strategy

# Reasoning about behavior in a game

- Let‘s look at behavior in the example
- Each player has a *strictly dominant strategy*
  - “Exam“ strategy strictly better than all other options
- No matter what your partner does, you should study for the exam
  - We can predict the outcome
- But both could be better off!
  - The 90,90 outcome won‘t happen during rational play

# The prisoner‘s dilemma

- Two suspects arrested and suspected for robbery.
- Interrogated separately
  - If neither confesses, both get 1 year sentence
  - If one confesses, other person gets 10 years
  - If both confess, 4 years in prison each

| | Don‘t confess | Confess |
|---|---|---|
| Don‘t confess | -1, -1 | -10, 0 |
| Confess | 0, -10 | -4, -4 |

# Interpretations

- Confessing is a *strictly dominant strategy*
  - Not confessing is better for both of them
  - But under rational play this is not achieved
- The dilemma arises frequently
  - e.g. performance-enhancing drugs for athletes

| | Don‘t use drugs | Use drugs |
|---|---|---|
| Don‘t use drugs | 3, 3 | 1, 4 |
| Use drugs | 4, 1 | 2, 2 |

  - Known as “arms races“
  - The payoffs must be aligned right

# Best response

- The best choice of one player, given a belief of what the other player will do
- Let‘s make this a bit more formal
  - Player 1 chooses strategy *S*
  - Player 2 chooses strategy *T*
  - Payoff to player i given S,T is $P_i(S,T)$
  - *Def*: *S* is a *best response* to *T* if $P_1(S,T) \geq P_1(S‘,T)$ for all other strategies S‘ of player 1.
  - *S* *strict best response* when $P_1(S,T) > P_1(S‘,T)$

# Dominant strategy

- <u>*Def:*</u> A *dominant strategy* is a best response to *every* strategy of the other player
  - Analogous for a *strictly dominant strategy*.
  - In the prisoner‘s dilemma, both players had strictly dominant strategies
  - Thus easy to predict what would happen
- But what about games without dominant strategies?

# Example: Marketing game (1/2)

- A game in which only one player has a dominant strategy
  - Two firms entering markets. Firm 2 established.
  - 60% of people want low-priced,  40% upscale
  - If they compete directly, Firm 1 gets 80% of sales and Firm 2 gets 20%.
  - Otherwise, each occupies own market segment

60% * 80%

Firm 2

| Firm 1 | Low-Priced | Upscale |
|---|---|---|
| Low-Priced | .48, .12 | .60, .40 |
| Upscale | .40, .60 | .32, .08 |

40% * 80%

# Example: Marketing game (2/2)

- What happens in this game?
  - Firm 1 has a dominant strategy: *Low-priced*, whereas Firm 2 does not
  - Firm 2 can thus assume Firm 1 will play this strategy
  - Our prediction for this game is .60, .40

| Firm 1 \ Firm 2 | Low-Priced | Upscale |
|---|---|---|
| Low-Priced | .48, .12 | .60, .40 |
| Upscale | .40, .60 | .32, .08 |

# Equilibrium concepts

- What if neither player has a dominant strategy?
  - How should we reason about such games?
  - We should expect players to use strategies that are *best responses to one another*
- Example:
  - Three client game.
  - What is the best response to each strategy?

So for player 1, there is no single strategy that is the best response to every strategy of player 2's. Player 2 does not have a dominant strategy either. Why?

| | A | B | C |
|---|---|---|---|
| A | 4, 4 | 0, 2 | 0, 2 |
| B | 0, 0 | 1, 1 | 0, 2 |
| C | 0, 0 | 0, 2 | 1, 1 |

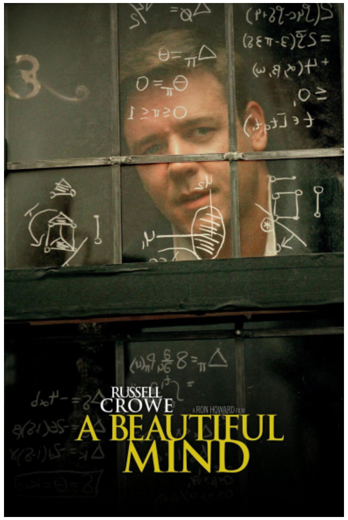
Player 1's A is the best response to player 2's A

Player 1's B is the best response to player 2's B

Player 1's C is the best response to player 2'sC

# Nash equilibrium

- Developed by John Nash in 1950.
  - Made famous in "A Beautiful Mind"
- <u>*Def*</u>:
  - For strategy *S* by player 1 and *T* by player 2, the pair (*S*,*T*) is a *Nash equilibrium* if *S* is a best response to *T*, and *T* is a best response to *S*
- More generally, at Nash equilibrium if *no player wants to unilaterally (<u>done only by one player</u>) deviate to an alternative strategy*



# Example: Nash equilibrium

- Three client game
  - Suppose Firm 1 chooses A and Firm 2 also chooses A
  - These strategies are the best responses to each other - neither player wants to deviate
  - Thus (A,A) is a *Nash equilibrium*.
    - It is also unique – no other pair works

| | A | B | C |
|---|---|---|---|
| A | 4, 4 | 0, 2 | 0, 2 |
| B | 0, 0 | 1, 1 | 0, 2 |
| C | 0, 0 | 0, 2 | 1, 1 |

# Example: Nash equilibrium

- Coordination game
  - Prepare a presentation with a partner
  - But don‘t know what software she will use
    - Incompatibility takes effort

| | PowerPoint | Keynote |
|---|---|---|
| **Powerpoint** | 1, 1 | 0, 0 |
| **Keynote** | 0, 0 | 1, 1 |

  - (PowerPoint,PowerPoint) is a Nash equilibrium
  - (Keynote,Keynote) is *also* a Nash equilibrium.
- Can have multiple Nash equilibria!

# Formualtion: what is a game?

- Many of the motivating examples are in fact from actual games
  - Soccer penalty kick, board games
  - Model of course more widely applicable
- *Def:* A *game* consists of three things.
  - **(1)** A set of *players*
  - **(2)** Each player has set of options how to behave called *strategies*
  - **(3)** For each choice of strategies, each player receives a *payoff* from the game.

# *n*-player game: formulation

- The players are indexed by $i$, where $i \in \{1,\ldots,n\}$
- Let $S_i$ denote the set of strategies available to player *i* and let $s_i \in S_i$ denote an arbitrary member of this set
  - So, $(s_1,\ldots,s_n)$ denotes a combination of strategies
- Let $u_i(s_1,\ldots,s_n)$ be the payoff to player *i* if the players choose the strategies
- Thus, a n-player game is specified by the players' strategy space and utility functions $(s_1,\ldots,s_n)$

$$G = \{S_1,\ldots,S_n; u_1,\ldots,u_n\}$$

# A dominant strategy

- When a strategy is best for a player no matter what strategy the other player uses, that strategy is said to
  - dominate all other strategies and
  - is called a dominant strategy.

$$u_i(s_i, \dots, s_{i-1}, s_i', \dots, s_n) \geq u_i(s_i, \dots, s_{i-1}, s_i'', \dots, s_n)$$

# Nash Equilibrium

In the n-player game $G = \{S_1, \ldots, S_n; u_1, \ldots, u_n\}$ the strategies $(s^*_1, \ldots, s^*_n)$ are a **Nash equilibrium** if, for each *i*, $s^*_i$ is player *i*'s best response to the strategies by others, $(s^*_1, \ldots, s^*_{i-1}, s^*_{i+1}, s^*_n)$

$$u_i(s^*_1, \ldots, s^*_{i-1}, s^*_i, s^*_{i+1}, \ldots, s^*_n)$$
$$> u_i(s^*_1, \ldots, s^*_{i-1}, s_i, s^*_{i+1}, \ldots, s^*_n)$$

for every feasible strategy $s_i \in S_i$; that is

$$s^*_i = \arg\max_{s_i \in S_i} u_i(s^*_1, \ldots, s^*_{i-1}, s_i, s^*_{i+1}, \ldots, s^*_n)$$

# Important games

- Battle of the sexes
  - Which kind of movie to rent?

| | Romance | Thriller |
|---|---|---|
| Romance | 1, 2 | 0, 0 |
| Thriller | 0, 0 | 2, 1 |

  - Two equilibria, but which one will be played?
- Hard to predict the outcome
  - Depends on social conventions

# Important games

- Stag Hunt
  - If hunters work together, they can catch a stag
  - On their own they can each catch a hare (rabbit)
  - If one hunter tries for a stag, he gets nothing

| | Hunt Stag | Hunt Hare |
|---|---|---|
| Hunt Stag | 4, 4 | 0, 3 |
| Hunt Hare | 3, 0 | 3, 3 |

- Two equilibria, but “riskier“ to hunt stag
  - What if other player hunts hare? Get nothing
  - Similar to prisoner‘s dilemma
    - Must trust other person to get best outcome!

# Important games

- Hawk-Dove (or Game of Chicken) refers to a situation where there is a competition for a shared resource
  - Each player either aggressive (H) or passive (D)
  - If both passive, divide resources evenly
  - If both aggressive – war! Disastrous for both
  - Otherwise aggressor wins

| | Dove | Hawk |
|---|---|---|
| **Dove** | 3, 3 | 1, 5 |
| **Hawk** | 5, 1 | 0, 0 |

- Can model the foreign policy of countries
  - each player, in attempting to secure her best outcome, risks the worst.

# Mixed strategies

- Do Nash equilibria always exist?
  - Matching Pennies game
  - Player 1 wins on same outcome, 2 on different

| | Heads | Tails |
|---|---|---|
| **Heads** | -1, +1 | +1, -1 |
| **Tails** | +1, -1 | -1, +1 |

- Example of a *zero-sum* game
  - What one player gains, the other loses
  - E.g. Allied landing in Europe on June 6, 1944
- How would you play this game?

# Mixed strategies

- You are *randomizing* your strategy
  - Instead of choosing H/T directly, choose a *probability* you will choose H.
- Player 1 commits to play H with some probability $p$
  - Similarly, player 2 plays H with probability $q$
- This is called a *mixed strategy*
  - As opposed to a *pure* strategy (e.g. $p$=0)
- What about the payoffs?

# Mixed strategies

| | Heads | Tails |
|---|---|---|
| Heads | -1, +1 | +1, -1 |
| Tails | +1, -1 | -1, +1 |

- Suppose player 1 evaluates pure strategies
  - Player 2 meanwhile chooses strategy $q$
  - If Player 1 chooses H, he gets a payoff of -1 with probability $q$ and +1 with probability $1-q$
  - If Player 1 chooses T, he gets -1 with probability $1-q$ and +1 with probability $q$
- Is H or T more appealing to player 1?
  - Rank the *expected values*
  - Pick H: expect $(-1)(q) + (+1)(1-q) = 1-2q$
  - Pick T: expect $(+1)(q) + (-1)(1-q) = 2q - 1$

| | Heads | Tails |
|---|---|---|
| **Heads** | -1, +1 | +1, -1 |
| **Tails** | +1, -1 | -1, +1 |

# Mixed strategies

- *Def:* Nash equilibrium for mixed strategies
  - A pair of strategies (now probabilities) such that each is a best response to the other.
  - Thm: Nash proved that this *always* exists.
- In Matching Pennies, no Nash equilibrium can use a pure strategy (by examining the table):
  - Player 2 would have a unique best response which is a pure strategy
  - But this is not the best response for player 1...
- What is Player 1's best response to strategy $q$?
  - If $1-2q \neq 2q-1$ (play 1 chooses H or T), then a pure strategy (either H or T) is a unique best response to player 1.
  - This can't be part of a Nash equilibrium by the above
  - So must have $1-2q=2q-1$ in any Nash equilibrium
    - Which gives $q=1/2$. Similarly $p=1/2$ for Player 1.
    - This is a unique Nash equilibrium (check!)

# Mixed strategies

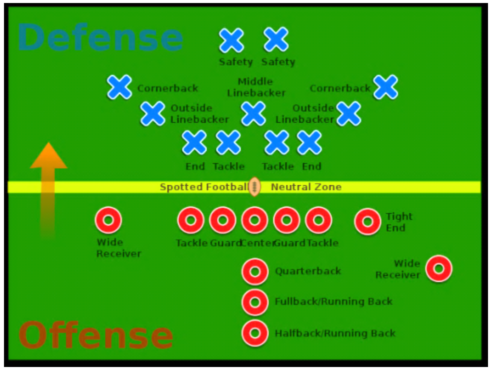
- Intuitively, mixed strategies are used to make it harder for the opponent to predict what will be played
  - By setting $q$=1/2, Player 2 makes Player 1 *indifferent* between playing H or T.
- How do we interpret mixed equilibria?
  - In sports (or real games)
    - Players are indeed randomizing their actions
  - Competition for food among species
    - Individuals are hardwired play certain strategies
    - Mixed strategies are *proportions* within populations
    - Population as a *whole* is a mixed equilibrium
  - Nash equilibrium is an equilibrium in beliefs
    - If you believe other person will play a Nash equilibrium strategy, so will you.
    - It is self-reinforcing – an equilibrium

# Mixed strategies: Examples

- American football
  - Offense can run with the ball, or pass forward

Defense

| Offense | | Defend pass ($q$) | Defend run ($1-q$) |
|---|---|---|---|
| | Pass | 0, 0 | 10, -10 |
| | Run | 5, -5 | 0, 0 |



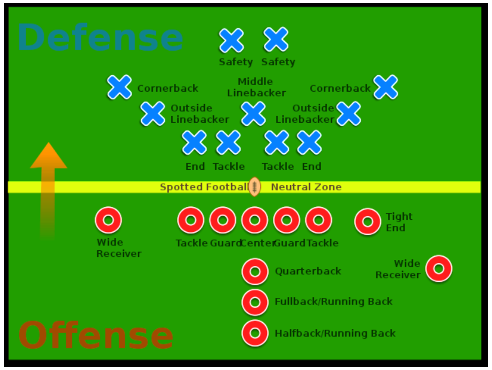

- What happens?
  - Suppose the defense defends against a pass with probability $q$
  - Pass: expect $(0)(q) + (10)(1-q) = 10-10q$
  - Run: expect $(5)(q) + (0)(1-q) = 5q$
  - Offense is indifferent between Pass and Run when $q=2/3$

# Mixed strategies: Examples

- American football
  - Offense can run with the ball, or pass forward

Defense

| Offense | | | Defend pass | Defend run |
|---|---|---|---|---|
| | *p* | **Pass** | 0, 0 | 10, -10 |
| | *1-p* | **Run** | 5, -5 | 0, 0 |




- What happens?
  - Suppose offense passes with probability $p$
  - Similarly, defense is indifferent when $p$=1/3
  - (1/3,2/3) is a Nash equilibrium
    - Expected payoff to offense: 10/3 (yard gain)

# Mixed strategies: Examples

- Penalty-kick game
  - An economist analyzed 1,417 penalty kicks from five years of professional soccer matches among European clubs
  - The success rates of penalty kickers given the decision by both the goalkeeper and the kicker to kick or dive to the left or the right are as follows:

goalkeeper

kicker

| | Defend left | Defend right |
|---|---|---|
| Left | 58% | 95% |
| Right | 93% | 70% |

Left and right are from kicker's perspective

# Mixed strategies: Examples

- Penalty-kick game

goalkeeper

kicker

| | Defend left | Defend right |
|---|---|---|
| Left | 0.58, -0.58 | 0.95, -0.95 |
| Right | 0.93, -0.93 | 0.70, -0.70 |


  – As a kicker, what is your strategy of kicking left or right?
  – As a goalkeeper what is your strategy of defending left or right?

# Mixed strategies: Examples

- Penalty-kick game
  - Soccer penalties have been studied extensively

| | Defend left | Defend right |
|---|---|---|
| Left | 0.58, -0.58 | 0.95, -0.95 |
| Right | 0.93, -0.93 | 0.70, -0.70 |


  - Suppose goalkeeper defends left with probability $q$
  - Kicker indifferent when
    - $(0.58)(q) + (0.95)(1-q) = (0.93)(q) + (0.70)(1-q)$
  - Get $q$ =0.42. Similarly $p$=0.39
  - True values from data? $q$=0.42 , $p$=0.40 !!
    - The theory predicts reality very well

Palacios-Huerta, Ignacio. "Professionals play minimax." *The Review of Economic Studies* 70.2 (2003): 395-415.
http://pricetheory.uchicago.edu/levitt/Papers/ChiapporiGrosecloseLevitt2002.pdf
http://www.mikeshor.com/courses/gametheory/docs/topic4/mixedsoccer.html

# Mixed strategies: Examples

- Penalty-kick game
  - More impressive is that the players' ability to randomize without patterns
    - Neither the player's past kicks nor his opponent's past behavior is useful in predicting the next kick
  - The overall chance of scoring on a penalty kick, in equilibrium is 80%
    - Why?


http://www2.owen.vanderbilt.edu/mike.shor/courses/game-theory/docs/lecture05/MixedSoccer.html

# Nash, 1951

- Proved the existence of a Nash equilibrium in finite (strategic form) games
- Theorem (Nash, 1951)

  *Every finite game has a mixed strategy Nash equilibrium*

  – Without knowing the existence of an equilibrium, it is difficult (perhaps meaningless) to try to understand its properties.
  – Armed with this theorem, we also know that every finite game has an equilibrium, and thus we can simply try to locate the equilibria.

Nash, J., 1951. Non-cooperative games. *Annals of mathematics*, pp.286-295.

# Mixed Strategies

- Let $\Sigma_i$ denote the set of probability measures over the pure strategy (action) set $S_i$ .
  - For example, if there are two actions, $\Sigma_i$ can be thought of simply as a number between 0 and 1, designating the probability that the first action will be played.

- We use $\sigma_i \in \Sigma_i$ to denote the mixed strategy of player i, and $\sigma \in \Sigma = \prod_{i \in I} \Sigma_i$ to denote a mixed strategy profile.
  - Note that this implicitly assumes that players randomize independently

- We similarly define $\sigma_{-i} \in \Sigma_{-i} = \prod_{j \neq i} \Sigma_j$
- Following expected utility theory, we extend the payoff functions $u_i$ from S to $\Sigma$ by

$$u_i(\sigma) = \int_S \; u_i(s) \, d\,\sigma\,(s)$$

# Nash Equilibrium

- A mixed strategy profile $\sigma*$ is a NE if

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i\left(\sigma_i \quad, \sigma_{-i}^*\right), \quad \text{for all } \sigma_i \quad \in \Sigma_i$$

- In other words, $\sigma*$ is a NE if and only if $\sigma_i^* \in B_i^*(\sigma_{-i}^*)$ for all *i*
  - $B_i^*(\sigma_{-i}^*)$ is the best response **correspondence** for player i, given other play's strategies $\sigma_{-i}^*$

$B_i(\sigma_{-i}) = argmax_{\sigma_i} u_i(\sigma_i, \sigma_{-i})$ is a correspondence, not a function as there may be multiple best responses

A correspondence between two sets A and A' is any subset R of Cartesian product AXA'

# Nash Equilibrium

- We define the correspondence B : Σ⇉Σ such that for all $\sigma$ ∈ Σ, we have
- A mixed strategy profile $\sigma*$ is a NE if

$$B(\sigma) = [B_i(\sigma_{-i})]_{i \in I}$$

$$\text{e.g., } B(\sigma_i \ , \sigma_{-i}) = [\sigma'_{-i}, \sigma'_i \ ]$$

- The existence of a Nash equilibrium is then equivalent to the existence of a mixed strategy $\sigma$* such that $\sigma$* ∈ B($\sigma$*), e.g., $B(\sigma_i^*, \sigma_{-i}^*) = [\sigma_{-i}^*, \sigma_i^*]$
  - i.e., existence of a fixed point of the mapping B
- Existence of a Nash equilibrium in finite games -> using a fixed point theorem

# Kakutani's Fixed Point Theorem

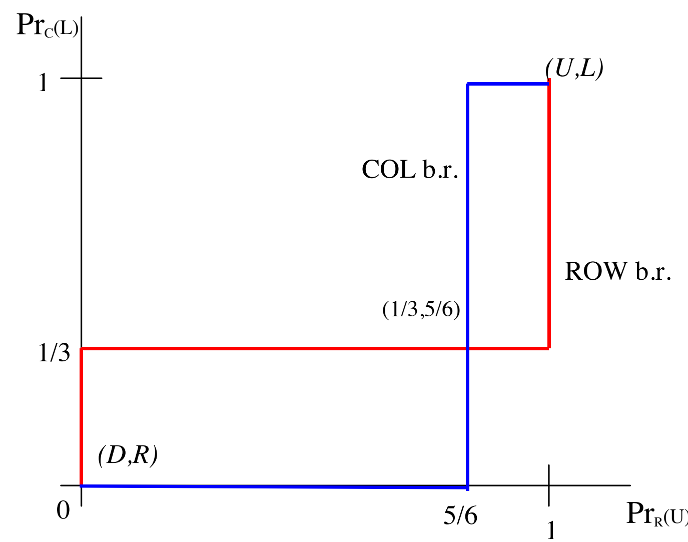
Recall we define a best response **correspondence** B : $\Sigma \rightrightarrows \Sigma$

- the sufficient conditions for a best response correspondence B to have a fixed point:
  1. $\Sigma$ is a compact, convex, nonempty subset of a (finite-dimensional) Euclidean space
  2. B() is nonempty for all $\sigma$
  3. B($\sigma$) is convex for all $\sigma$
  4. B($\sigma$) has a **closed graph**: that is, if $\{x,y\} \rightarrow \{x^*,y^*\}$ with $y \in B(x)$, then $y^* \in B(x^*)$

# An example

- Define $\sigma_i$(s) the probability of player i to play the strategy s.
  - For example $\sigma_R$(U) is the probability ROW plays U
- NE: {(5/6)U+(1/6)D,(1/3)L+2/3R}

| | | $\sigma_C(L)$ | $\sigma_C(R)$ |
|---|---|---|---|
| | | $L$ | $R$ |
| $\sigma_R(U)$ | $U$ | $2,1$ | $0,0$ |
| $\sigma_R(D)$ | $D$ | $0,0$ | $1,5$ |

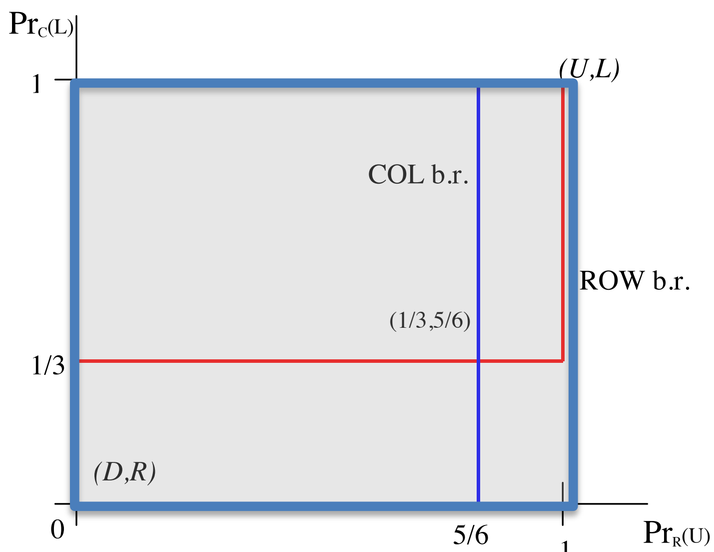



Best response correspondence

# Condition (1)

- A set in a Euclidean space is compact if and only if it is bounded and closed
  - Bounded as Σ is prob. Measure ([0,1])
  - Closed too as Σ contains its boundary
- A set Σ is convex if for any x,y ∈ Σ and any $\lambda$ ∈ [0,1], $\lambda$x + (1 − $\lambda$)y ∈ Σ

| | | $\sigma_C(L)$ | $\sigma_C(R)$ |
|---|---|---|---|
| | | $L$ | $R$ |
| $\sigma_R(U)$ | $U$ | $2,1$ | $0,0$ |
| $\sigma_R(D)$ | $D$ | $0,0$ | $1,5$ |

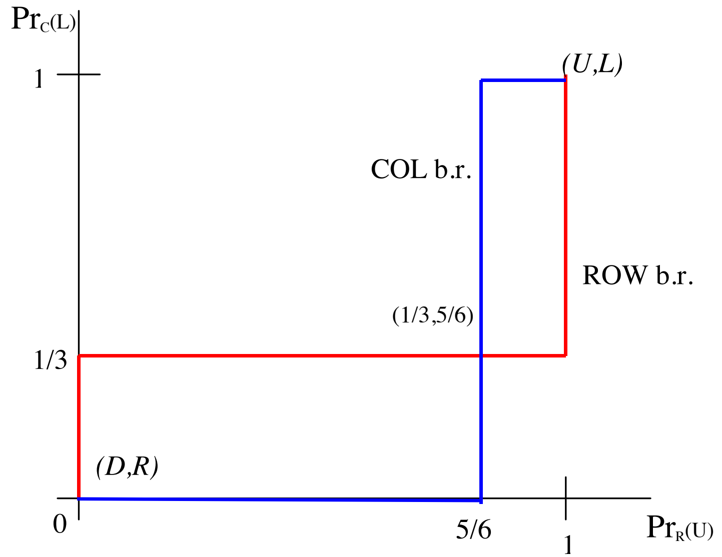


# Condition (2)

- B() is nonempty for all $\sigma$
  - every player has a best response to the other players' strategies—whatever those strategies are

NE: {(5/6)U+(1/6)D,(1/3)L+2/3R}

| | | $\sigma_C(L)$ | $\sigma_C(R)$ |
|---|---|---|---|
| | | $L$ | $R$ |
| $\sigma_R(U)$ | $U$ | $2,1$ | $0,0$ |
| $\sigma_R(D)$ | $D$ | $0,0$ | $1,5$ |



Best response correspondence

# Condition (3)

- $B(\sigma)$ is convex-valued correspondence (aka $B(\sigma)$ is convex set) for all $\sigma$

Equivalently, $B(\sigma) \subset \Sigma$ is convex if and only if $B_i(\sigma_{-i})$ is convex for all i

**Proof:**

- Let $\sigma_i', \sigma_i'' \in B_i(\sigma_{-i})$. we thus have by definition,
  $u_i(\sigma_i', \sigma_{-i}) \geq u_i(\tau_i, \sigma_{-i})$ for all $\tau_i \in \Sigma_i$,
  $u_i(\sigma_i'', \sigma_{-i}) \geq u_i(\tau_i, \sigma_{-i})$ for all $\tau_i \in \Sigma_i$

- Then, for all $\lambda \in [0,1]$, we have
  $\lambda u_i(\sigma_i', \sigma_{-i}) + (1-\lambda) u_i(\sigma_i'', \sigma_{-i}) \geq u_i(\tau_i, \sigma_{-i})$ for all $\tau_i \in \Sigma_i$

- By the linearity of $u_i$
  $u_i(\lambda \sigma_i' + (1-\lambda)\sigma_i'', \sigma_{-i}) \geq u_i(\tau_i, \sigma_{-i})$ for all $\tau_i \in \Sigma_i$

- Therefore, $\lambda \sigma_i' + (1-\lambda)\sigma_i'' \in B_i(\sigma_{-i})$, showing that $B(\sigma)$ is indeed convex-valued

Asu Ozdaglar, Game Theory with Engineering Applications, 2010

Fudenberg, Drew, and Jean Tirole. *Game theory*. MIT press, 1991.

# Condition (4)

- B($\sigma$) has a **closed graph**: that is, if {x,y} → {x*,y*} with y ∈ B(x), then y*∈ B(x*)

- Suppose to obtain a contradiction, that B($\sigma$) does not have a closed graph
- Then, there exists a sequence $(\sigma_n, \ \hat{\sigma}_n) \rightarrow (\sigma, \ \hat{\sigma})$ with $\hat{\sigma}_n \in$ B($\sigma_n$), but $\hat{\sigma} \notin$ B ($\sigma$), i.e., there exists some i such that $\hat{\sigma}_i \notin$ B($\sigma_{-i}$)
- This implies that there exists some $\sigma'_i \in \Sigma_i$ and some $\epsilon > 0$ such that
$$u_i(\sigma'_i, \sigma_{-i}) > u_i(\hat{\sigma}_i, \sigma_{-i}) + 3\epsilon$$
- By the continuity of $u_i$ and the fact that $(\sigma_n)_{-i} \rightarrow \sigma_{-i}$ , we have for sufficiently large n,
$$u_i(\sigma'_i, (\sigma_n)_{-i}) > u_i(\sigma'_i, \sigma_{-i}) \ - \ \epsilon$$
- Combining the two relations, we obtain
$$u_i(\sigma'_i, (\sigma_n)_{-i}) > u_i(\hat{\sigma}_i, \sigma_{-i}) + 2\epsilon \geq ((\hat{\sigma}_n)_i, (\hat{\sigma}_n)_{-i}) + \epsilon$$

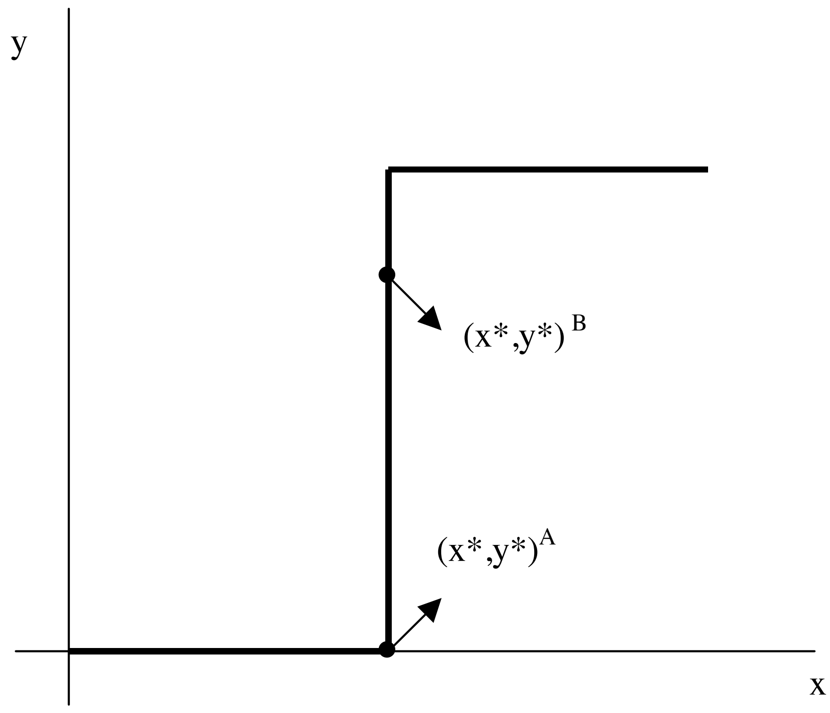
where the second relation follows from the continuity of $u_i$ . This contradicts the assumption that $\hat{\sigma}_n \in$ B($\sigma_n$), and completes the proof

Asu Ozdaglar, Game Theory with Engineering Applications, 2010

Fudenberg, Drew, and Jean Tirole. *Game theory*. MIT press, 1991.

# Condition (4)

- $B(\sigma)$ has a **closed graph**: that is, if $\{x,y\} \rightarrow \{x^*,y^*\}$ with $y \in B(x)$, then $y^* \in B(x^*)$




- First take a sequence $\{x_n\}$ that converges to $x^*$ from the left.
- Second take the values $y_n$ in $B(x_n)$ (which in this case would be $y_n$ = 0 for all $x_n$).
- As can be seen $\{y_n\}$ approaches $y^*$ = 0 such that $y^*$ = 0 belongs to the correspondence of $x^*$ (i.e. $y^*$ in $B(x^*)$) in the point $(x^*;y^*)^A$).
- Hence the figure corresponds to a closed graph

$B(\sigma)$ has a **closed graph**: that is, if $\{x,y\} \rightarrow \{x^*,y^*\}$ with $y \in B(x)$, then $y^* \in B(x^*)$

# Pareto optimality

- Even playing best responses does not always reach a good outcome as a group
  - E.g. prisoner‘s dilemma
- Want to define a *socially good outcome*
- *Def*:
  - A choice of strategies is *Pareto optimal* if no other choice of strategies gives all players a payoff at least as high, and at least one player gets a strictly higher payoff
  - In other words, no player can gain higher payoff without sacrificing others’ payoffs
- Note: *Everyone* must do at least as well

# Social optimality

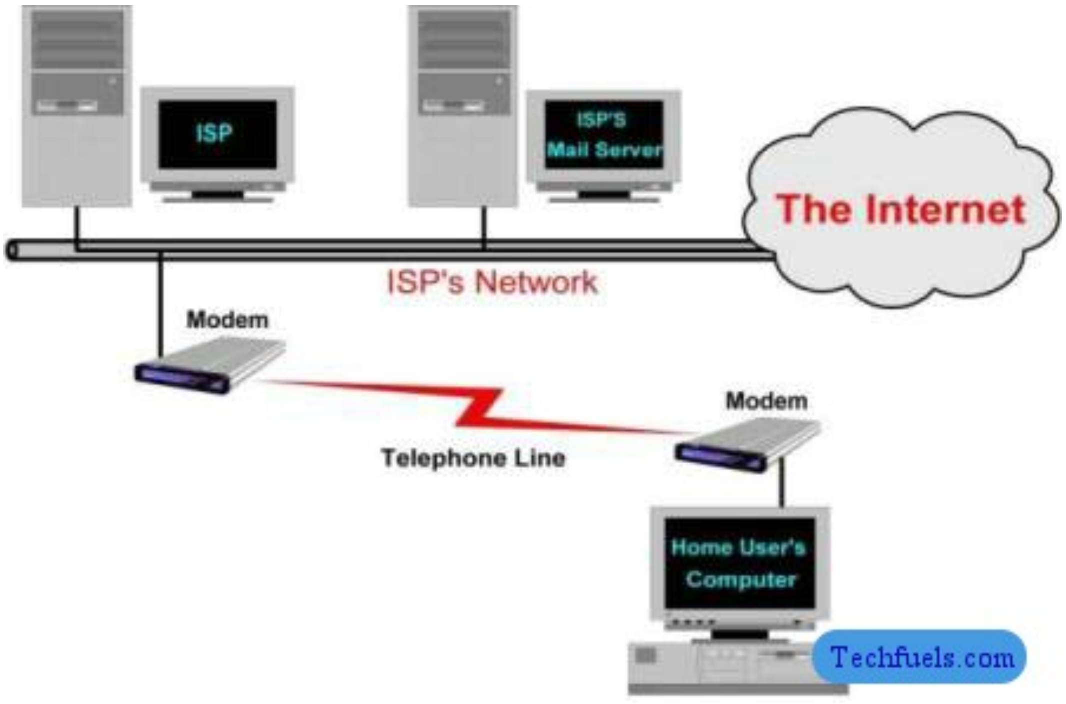
- ***Def:***
  - A choice of strategies is a *social welfare maximizer* (or *socially optimal*) if it maximizes the sum of the players' payoffs.
- Example:
  - The unique Nash equilibrium in this game is socially optimal

| | Presentation | Exam |
|---|---|---|
| **Presentation** | 98,98 | 94,96 |
| **Exam** | 96,94 | 92,92 |

# ISP COMPETITION

# Competition between ISPs

- How does the outcome of competition among two Internet Service Providers (ISPs) in a competitive market?



# Cournot Model of Duopoly

- *n* number of ISPs offer their service (Bandwidth) to the customers
- The ***cost*** to ISP *i* of producing $q_i$ units of the Bandwidth is $C_i(q_i)$, where $C_i$ is an increasing function
- The Bandwidth is sold at a single ***price*** $P(Q)$, where $Q$ is the total output and $P$ is a decreasing function unless it is already zero.
- If the output of each ISP *i* is $q_i$, then the price is $P(q_1 + \ldots + q_n)$, so ***revenue*** of *i* is $q_i P(q_1 + \ldots + q_n)$.

# Cournot Model of Duopoly

- Thus ISP i's profit, equal to its revenue minus its cost, is

  $$F_i(q_1, \ldots, q_n) = q_i P(q_1 + \ldots + q_n) - C_i(q_i)$$

- The ISP Games:
  - Players: the ISPs.
  - Actions: the set of its possible outputs $q_1, \ldots, q_n$
  - Utilities: ISPs preferences are represented by their profits $F_1, \ldots, F_n$

# Cournot Model of Duopoly

- With certain forms of $C_i$ and $P$ a Nash equilibrium can be computed
  - Suppose there are two ISPs, each ISP's cost function is the same, given by

$C_i(q_i) = cq_i$ for all $q_i$ , where $c$ (>=0) is a constant unit cost

  - The inverse demand function is linear where it is positive, given by

$$P(Q) = a - Q \quad if\ Q <= a$$

$$P(Q) = 0 \quad if\ Q > a, \quad where\ a > 0$$

# Cournot Model of Duopoly

- Assume that $c < a$, so that there is some value of total output Q for which the market price P(Q) is greater than the ISPs' common unit cost c
  - If c were to exceed a, there would be no output for the ISPs at which they could make any profit
- ISP *1* s' profit (the same case for ISP *2*):

$$F_1(q_1, q_2) = q_1(P(q_1 + q_2) - c)$$
$$= q_1(a - c - (q_1 + q_2))$$

# Cournot Model of Duopoly

To find the Nash equilibrium, recall for every feasible strategy $s_i \in S_i$ there is

$$s^*_i = \arg\max_{s_i \in S_i} u_i(s^*_1, \ldots, s^*_{i-1}, s_i, s^*_{i+1}, \ldots, s^*_n)$$

Thus $$q_1^* = \arg\max_{0 \le q_1 < \infty} F_1(q_1,\ q_2^*)$$

$$= \arg\max_{0 \le q_1 < \infty}\ q_1(a - c - (q_1\ +\ q_2^*)\ )$$

The first-order condition for the optimization problem yields (assume $q_1 < a-c$):

$$q_1^* = \frac{1}{2}(a - q_2^* - c)$$

The same case for $q_2$: $$q_2^* = \frac{1}{2}(a - q_1^* - c)$$

# Cournot Model of Duopoly

Solving this pair of equations yields:

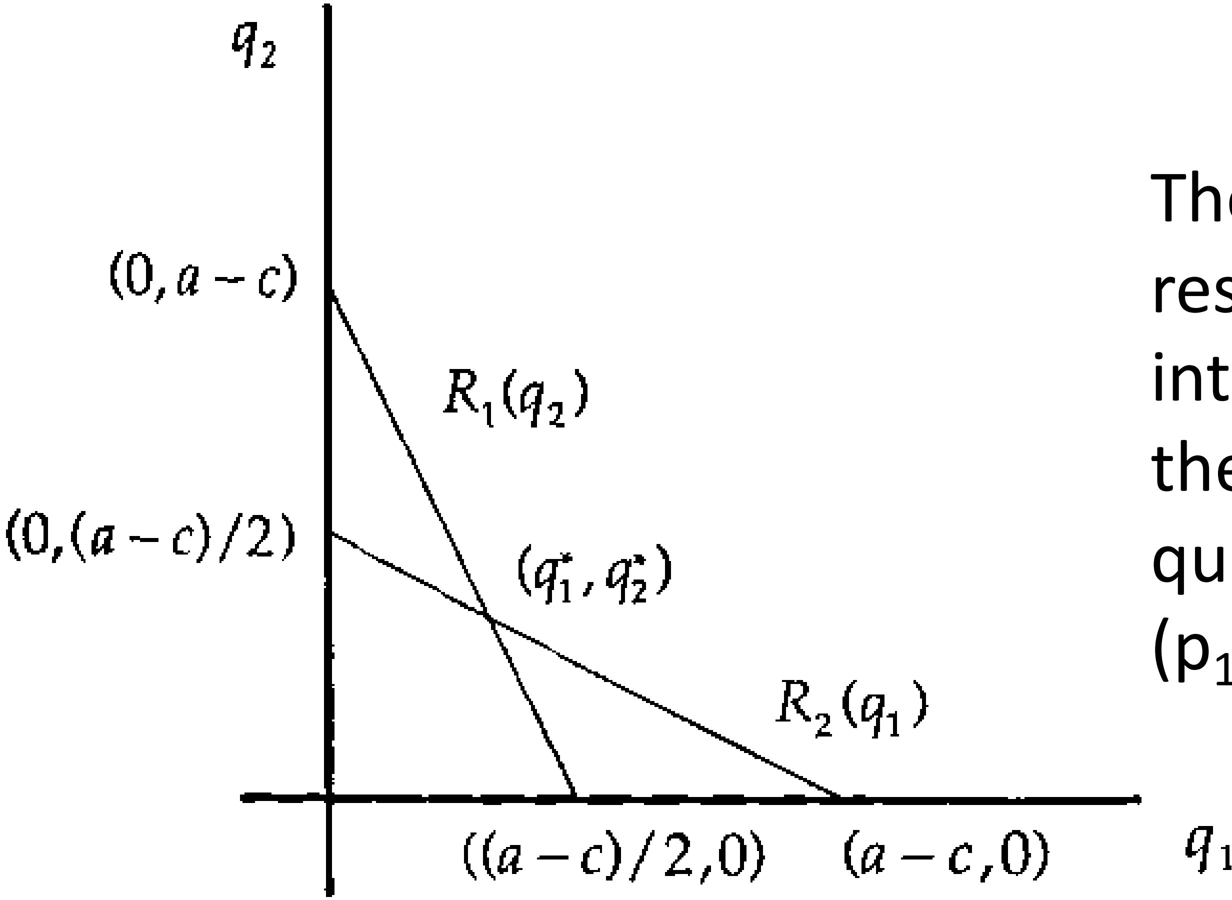
$$q_1{}^* = q_2{}^* = \frac{a-c}{3},$$

which is indeed smaller than *a-c and the profit in the equilibrium is*

$$F_1{}^* = F_2{}^* = \frac{a-c}{3}\left(a - \frac{a-c}{3} - c\right) = \frac{2(a-c)^2}{9},$$

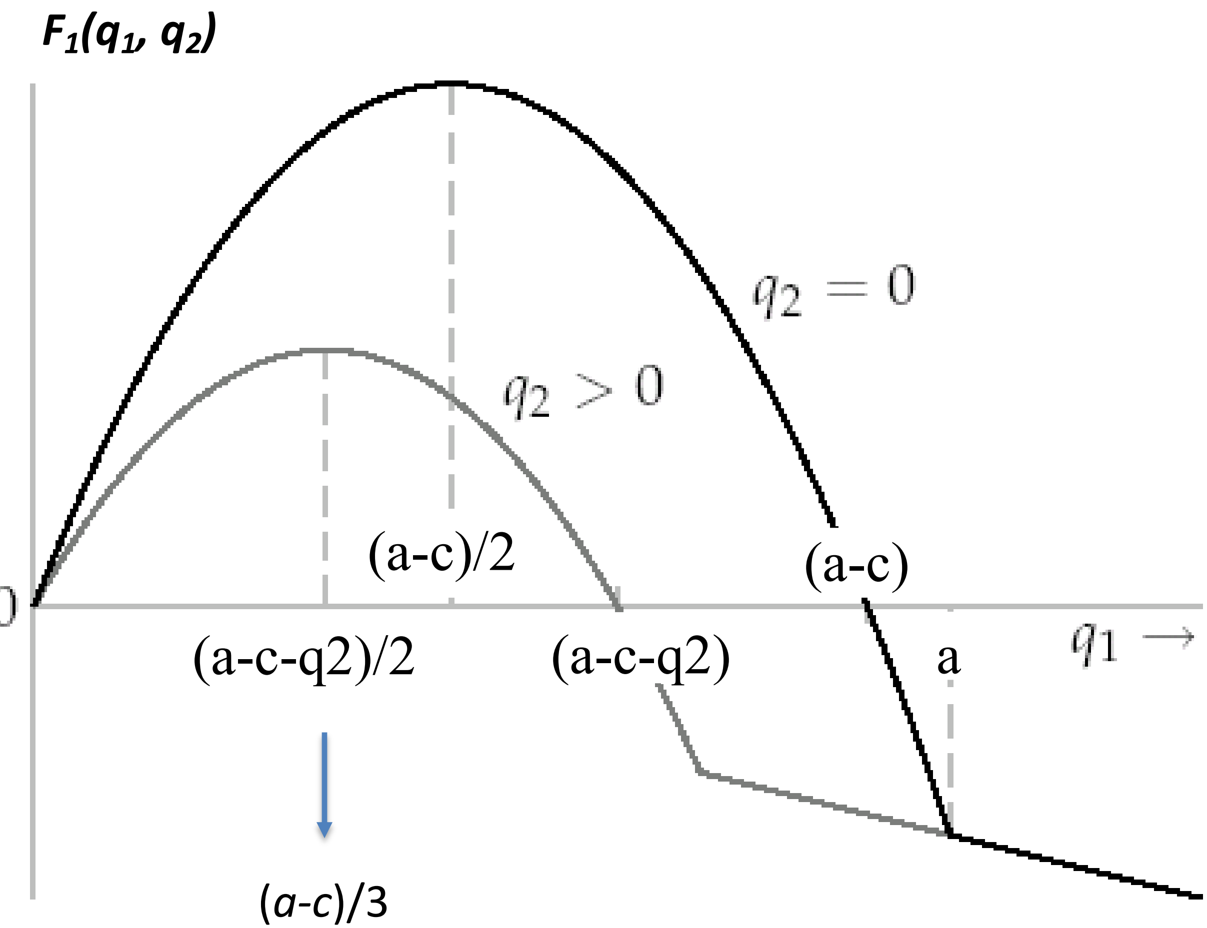
# Solution by Best Response

- ISP 1's best response $R_1(q_2) = \frac{1}{2}(a - q_2 - c).$
- ISP 2's best response $R_2(q_1) = \frac{1}{2}(a - q_1 - c);$




These two best response functions intersect only once at the equilibrium quantity pair $(p_1^*, p_2^*)$

# An analysis of the result



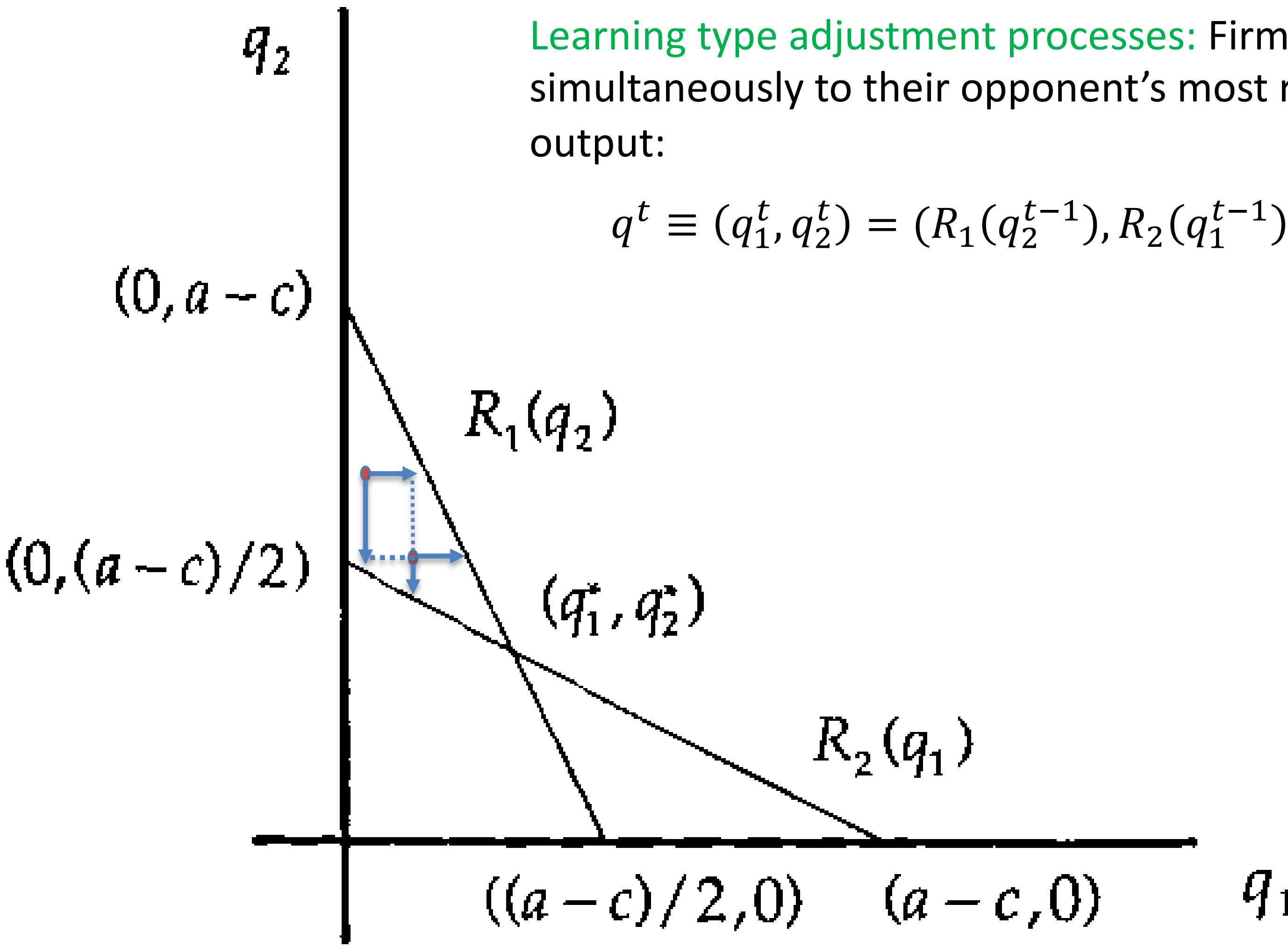

- Each ISP would like to be a monopolist, (the right cure)
- In equilibrium, the output of ISP 1 reduces to $q_1=(a-c)/3$ (the left curve)

# An analysis of the result

- Given two ISPs, the aggregate profits for the *duopoly* would be maximized by setting $q_1 = q_2 = q_m/2 = (a-c)/4$ the half of the monopoly quantity
- However, because $q_1$ or $q_2$ is low, each ISP has an incentive to deviate
  - in order to increase their individual profits
- As the aggregate quantity goes higher, the price is lower; the temptation to increase output is reduced
- In the equilibrium, $q_1 = q_2 = (a-c)/3$

# Solution as a result of learning

- ISP 1's best response $R_1(q_2) = \frac{1}{2}(a - q_2 - c).$
- ISP 2's best response $R_2(q_1) = \frac{1}{2}(a - q_1 - c);$




Learning type adjustment processes: Firms react simultaneously to their opponent's most recent output:

$$q^t \equiv (q_1^t, q_2^t) = (R_1(q_2^{t-1}), R_2(q_1^{t-1})) \equiv f(q^{t-1})$$

# Discussion about adjustment processes

- Assumption: expects that his opponent's output in the future will be the same as it is now
- More plausible based on the average value of their opponent's past play – fictitious play
- Asymptotically stable if converges to a particular steady state (a single NE) for all initial state close to it
  - Global stable if from every starting state
- But not always converge – circling effect
- Also it ignores the way current action will influence their opponent's action in the next period – repeated game (they know they face one another repeatedly)

# References and Further Readings


- Martin J. Osborne, An Introduction to Game Theory
- Networks, Crowds, and Markets. Reasoning about a Highly Connected World. by David Easley and Jon Kleinberg. Cambridge University Press 2010
- Fudenberg, Drew, and Jean Tirole. *Game theory*. MIT press, 1991.
- Ignacio Palacios-Huerta, "Professionals Play Minimax," Review of Economic Studies, vol. 70(2), April 2003, pp. 395-415.
- Asu Ozdaglar, Game Theory with Engineering Applications, 2010
- Mohammad T. Irfan, Explanation of Nash's Theorem and Proof with Examples
- Nash, John. "Non-cooperative games." *Annals of mathematics* (1951): 286-295.