AI-MAS Group

Multi-Agent Systems

University Politehnica of Bucarest

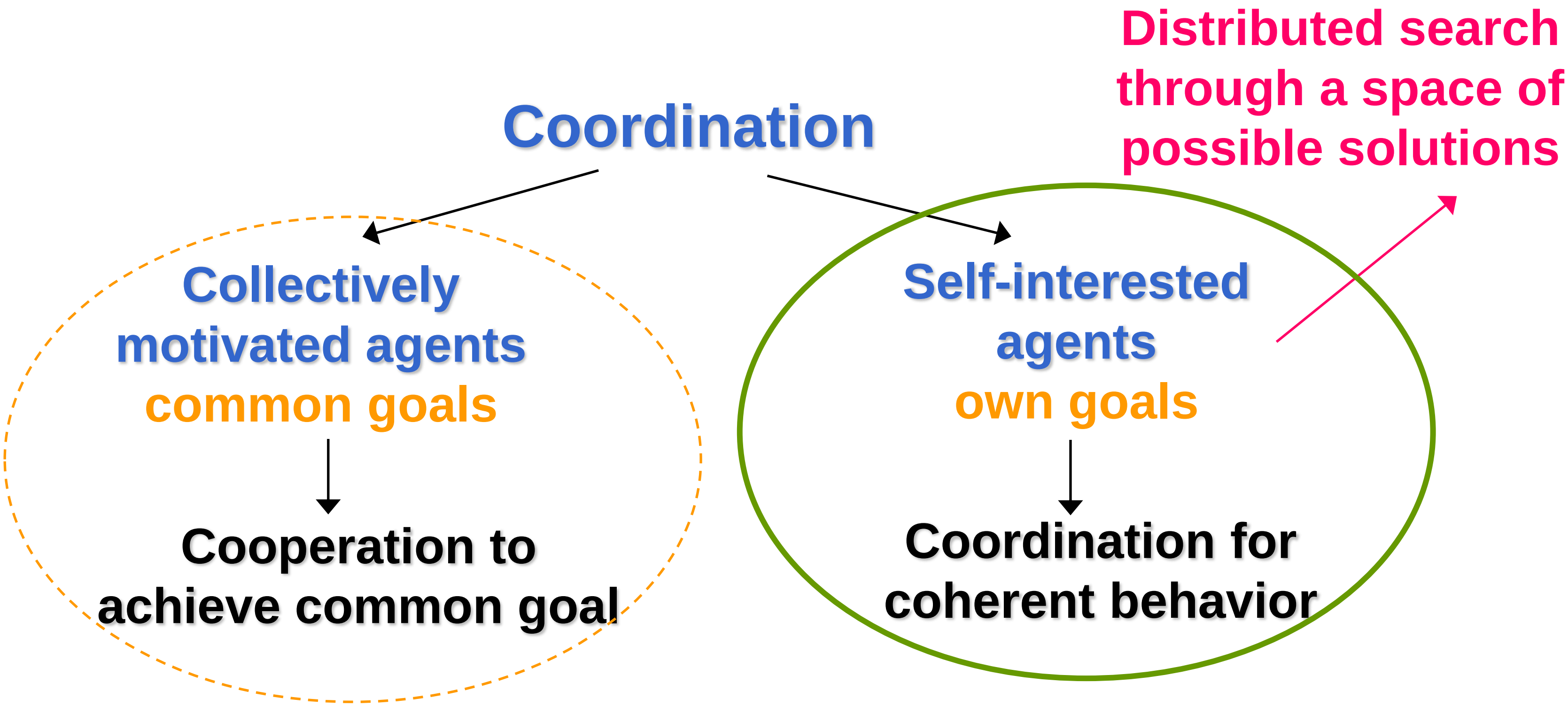
Spring 2024

Adina Magda Florea

Lecture 7: Noncooperative Game Theory I

1 MAS Coordination
2 MAS Negotiation
3 Noncooperative Game Theory
4 Strategies in normal form games

1. Multi-Agent coordination



Agent coordination overview

- **Tightly coupled interactions** - distributed search
- **Complex agents** - distributed planning, task sharing, resource sharing
- **Heterogeneous agents** - interaction protocols: Contract Net, KQML conversations, FIPA protocols
- **Dynamic interactions** - commitments and conventions
- **Complex interactions** - organizational structures to reduce complexity
- **Unpredictable interactions** - social laws

Cooperative

Neutral or competitive

- **Own interests -** negotiation based on game theory, voting, auctions, market mechanisms, Contract Net, coalition formation

2 Negotiation

- **Negotiation = interaction among agents for the purpose of coming to an agreement.**
- **MAS Negotiation techniques**
  - **Game theoretic negotiation**
  - **Protocols for Social choice (voting)**
  - **Protocols for Resource allocation**
    - Task allocation through negotiation
    - Auctions, General equilibrium market mechanisms
  - **Heuristic negotiation, Argumentation based negotiation**
  - **Coalitional Game theory**

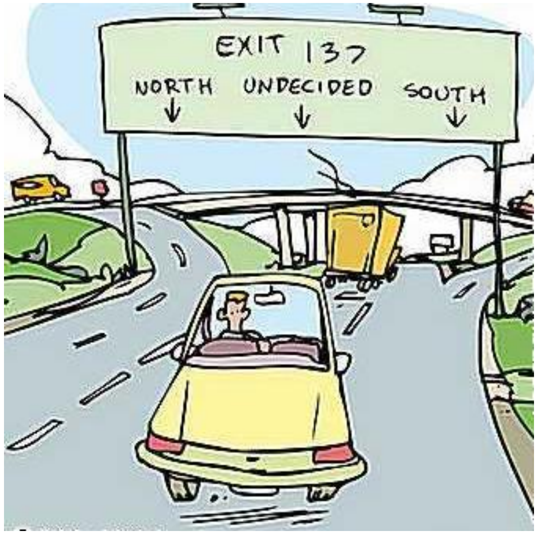
- **Negotiation includes:**
  - a communication language
  - a negotiation protocol
  - a decision process by which an agent decides upon its position, concessions, criteria for agreement, etc.
- Single party **or** multi-party negotiation: one to many (e.g., auctions) or many to many (e.g., eBay, Okazii)
- May include a single shot message by each party or conversation with several messages going back and forth

3 Noncooperative Game Theory

- **Game theory** = the mathematical study of interaction among independent, self-interested agents
- Applied to economics, political sciences, biology, psychology, linguistics and computer science
- **Non-cooperative game theory** = **games in normal form** – main branch of game theory
- Applied in case of self-interested agents

Utility theory

- Utility theory - reference guide for decisions when the payoff is uncertain
- **Utility theory** = every state of the world has a degree of usefulness – **utility** or **payoff**, to an agent, and that agent will prefer states with higher utility
- **Decision theory** = an agent is rational if and only if it chooses the actions that yields the highest expected utility, averaged over all possible outcomes of actions



Maximum Expected Utility

- **The probability** that the result of an action (***a***) executed in *e* to be the new state ***e'***

$$\sum_{e' \in env(e,a)} prob(ex(a,e) = e') = 1$$

- **The** **expected utility** of a action ***a*** in a state ***e***, from the point of view of the agent

$$U(a,e) = \sum_{e' \in env(e,a)} prob(ex(a,e) = e') * utility(e')$$

- **The principle of maximum expected utility (MEU) says that a rational agent should choose an action that maximizes U(a,e)**

Maximum Expected Utility

- The justification for MEU is rooted in the **von Neumann-Morgenstern (VNM) utility theory** (Von Neumann and Morgenstern, 2007)
- This theory essentially proves that an agent is VNM-rational if and only if there exists a real-valued utility function such that every preference of the agent is characterised by maximising the single expected reward.
- The VNM utility theorem is the basis for the well-known **expected utility theory (Schoemaker, 2013),** which essentially states that rationality can be modelled as maximising an expected value
- Specifically, the **VNM utility theorem** provides both necessary and sufficient conditions under which the **expected utility hypothesis holds**
- In other words, rationality is equivalent to VNM-rationality, and it is safe to assume an intelligent entity will always choose the action with the highest expected utility in any complex scenarios.

Stochastic game

- A multiplayer extension of an MDP and a generalization of a repeated game in Non coop game theory
- MDP S, A, P:S x A $\rightarrow$*P*(S),   R:S x A $\rightarrow$*R*

Stochastic game

- n – number of agents
- S – set of states shared by all agents
- **A** = A_1 x … x A_n where A_i is a finite set of actions of agent ***i***.
- P:S x **A** $\rightarrow$ P(S) – transition probability from s to s' given agent joint actions
- **R**=(r_1,..,r_n) where $r_i(s,a_1,..,a_n)$:S x **A** $\rightarrow$*R* is the reward (payoff) function for agent ***i***.

Example of utility function

- To show how utility theory functions can be used as a basis for making decisions

			Bob	Carol
Alice:	**Club (c)**	100	-90	1.5 factor
	Movie (m)	50	-40	1.5 factor
	Home (h)	50		

Bob prefers - Club 60% of time, 40% Movie

Carol prefers – Club 25% , Movie 75%

Which is Alice best course of action?

We list Alice's utilities for each possible state of the world

A=c (club)	B=c (60%)	B=m (40%)
C=c (25%)	15	150
C=m (75%)	10	100

A=m (movie)	B=c (60%)	B=m (40%)
C=c (25%)	50	10
C=m (75%)	75	15

			Bob (B)	Carol (C)
Alice (A)	**Club (c)**	100	-90	1.5 factor
	Movie (m)	50	-40	1.5 factor
	Home (h)	50		

We list Alice's utilities for each possible state of the world

A=c **(club)**	**B**=c (60%)	**B**=m (40%)
C=c (25%)	15	150
C=m (75%)	10	100

A=m **(movie)**	**B**=c (60%)	**B**=m (40%)
C=c (25%)	50	10
C=m (75%)	75	15

			Bob (B)	**Carol (C)**
Alice (A)	**Club (c)**	100	-90	1.5 factor
	Movie (m)	50	-40	1.5 factor
	Home (h)	50		

U(A,c) = 0.25*(0.6*15+0.4*150)+0.75*(0.6*10+0.4*100)=51.75

U(A,m) = 0.25*(0.6*50+0.4*10)+0.75*(0.6*75+0.4*15)=46.75

U(A,h) = 50

Preferences and utility

- The idea of utility can be grounded in a more basic concept = preferences
- **O** – a finite set of outcomes
- For any pair **o1** and **o2** from **O**
- o1 $\geq$ o2 – the agent prefers o1 to o2
- o1 $\sim$ o2 – the agent is indifferent between o1 and o2
- o1>o2 – the agent strictly prefers o1 to o2
- We need only o1 $\geq$ o2 as we can define o1 $\sim$ o2 and o1>o2 based on that

Lottery

- A random selection of one of a set of outcomes according to specified probabilities
- Formaly, a **lottery** is a probability distribution over outcomes written

$$[p_1{:}o_1, \ldots, p_k{:}o_k]$$

where each $o_i \in O$

each $p_i \geq 0$ and

$$\sum_{i=1,k} p_i = 1$$

Games in normal form

- Agents will have utility functions they want to maximize
- As long as the outcomes and their probabilities are known to the agent, it can decide how to act optimally
- Agents need to choose the course of action that *maximize the expected utility*

$$U(a,e) = \sum_{e' \in env(e,a)} prob(ex(a,e) = e') * utility(e')$$

- But when there are **2 or more agents** whose actions can affect each other's utility we turn to <u>game theory</u>

TCP user's game

- You and one of your colleagues are the only one using Internet
- Use of TCP
- You have 2 possible strategies: use Correct implementation or use a Defective implementation
- If both you and your colleague adopt C – average packet delay is 1ms
- If you both adopt D then the delay is 3ms
- If one adopts D and the other C then no delay for the first and 4ms delay for the second

Definition of games in normal form

		Column	player
		Correct	Defect
Row player	Correct	-1, -1	-4, 0
	Defect	0, -4	-3, -3

- The **normal form** , also known as the strategic form, is the most familiar representation of strategic interactions in game theory.
- Represent every player’s utility for every state of the world in the special case when states depend only on the player’s combined actions

Definition of games in normal form

- A (finite n-person) normal form game is a tuple (**N**,**A**,**U**) where:
  - **N** is a finite set of **n** players indexed by ***i***
  - $\mathbf{A = A_1 \times \ldots \times A_n}$ where $\mathbf{A_i}$ is a finite set of actions available to player ***i***.
  - Each vector $\mathbf{a=(a_1, .., a_n)} \in \mathbf{A}$ is called an **action profile** (or strategy profile)
  - $\mathbf{U=(u_1,..,u_n)}$ where u_i (s): $\mathbf{A} \rightarrow$R is a real-valued utility (payoff) function for player ***i***.

    (the state corresponds to **a**, $s=(a_1, ...a_n)$)
- A natural way to represent games is via an **n-dimensional matrix**

Examples of games in normal form

Prisoner's dilema

		Column player	
		Cooperate	Defect
Row player	Cooperate	3, 3	0, 5
	Defect	5, 0	1, 1

- Cooperate (keep silent)
- Defect (betray the other)

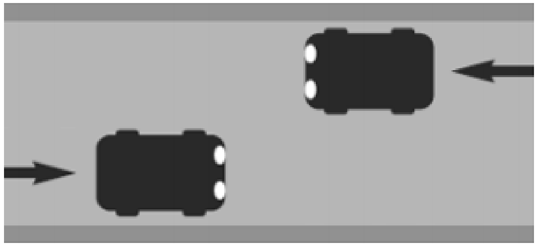
		Column player	
		Cooperate	Defect
Row player	Cooperate	a, a	b, c
	Defect	c, b	d, d

Prisoner's dilema = any normal form game for which **c > a > d > b**

Common payoff games

Coordination game

		Column player	
		Left	Right
Row player	Left	1, 1	0, 0
	Right	0, 0	1, 1



- 2 people driving toward each other in a country with no traffic rules
- Each driver independently decides whether to stay on the left or the right
- Need to coordinate your action with the action of the other driver

- A **common payoff game** is a game in which for all actions profiles **a**$\in A_1 \times \ldots \times A_n$ and for any pair of agents ***i***, ***j*** it is the case that $u_i(a)=u_j(a)$
- They are also called **pure coordination games** or **team games**
- Agents have no conflicting interests, their aim is to coordinate on an action that is maximally beneficial to all

Zero sum games

Matching pennies

	Heads	Tails
Heads	1, -1	-1, 1
Tails	-1, 1	1, -1

- Properly called constant sum games
- A game is a **constant sum game** if there exists a constant **c** such that for each strategy profile $a \in A_1 \times A_2$ it is the case that $\mathbf{u_1(a)+u_2(a)=c}$
- If **c=0** then **zero-sum games**
- Represent pure competition

Zero sum games (other example)

Rock-Paper-Scissors

- Two players: Each simultaneously picks an action: Rock, Paper or Sissors
- The payoff (utility):
  - **Rock** beats **Scissors**
  - **Scissors** beats **Paper**
  - **Paper** beats **Rock**

Rock, Paper, Scissors	Rock	Paper	Scissors
Rock	0, 0	-1, 1	1, -1
Paper	1, -1	0, 0	-1, 1
Scissors	-1, 1	1, -1	0, 0

Other examples

Battle of sexes

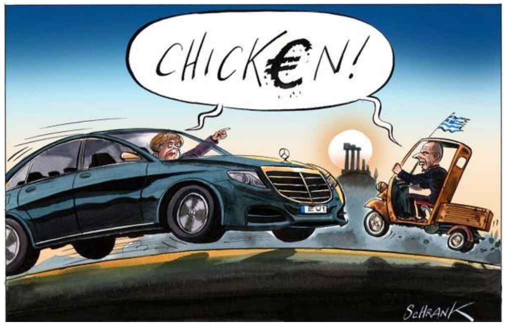
	Football	Movie
Football	2, 1	0, 0
Movie	0, 0	1, 2

- Game that have elements of both coordination and competition

Game of Chicken

		Column	player
		Swerve	Strait
Row player	Swerve	Tie, Tie	Lose, Win
	Strait	Win, Lose	Crash, Crash

		Column	player
		Swerve	Strait
Row player	Swerve	0, 0	-1, +1
	Strait	+1,-1	-10, -10



4. Strategies in normal form games

- Agents are supposed to behave rationally
- **Rational behavior** = an agent prefers a greater utility (payoff) over a smaller one
- **Payoff maximization**: what to maximize? From what point of view?

Strategies in normal form games

Outside observer: can some outcomes be considered better than some others?

- **Social welfare**
  - The sum of agents' utilities (payoffs) in a given solution.
  - Measures the global good of the agents
  - Problem: how to compare utilities
- **Pareto optimal solutions**

From the individual agent's point of view

- **Nash equilibrium**

- **Pure strategy:** select a single action to play

- **Mixed strategy:** randomize over the set of allowable actions according to some probability distribution

Let **(N,A,U)** be a normal form game.

with the **strategy (action) profile** –

the vector $\mathbf{s=(a_1, .., a_n) \in A}$

(the state is given by strategy profile and corresponds to the combination of actions)

Pareto optimality

A strategy profile **s** **Pareto dominates** the strategy profile **s'** if for all $\mathbf{i} \in N$, $\mathbf{u_i(s) \geq u_i(s')}$ and there is some $\mathbf{j} \in N$ for which $\mathbf{u_j(s) > u_j(s')}$

- Defines a partial order over strategy profiles
- Useful when we have a set of noncomparable optima

Strategy profile **s** is **Pareto optimal** or **strictly Pareto efficient** if there is not another strategy profile $\mathbf{s'} \in \mathbf{S}$ that Pareto dominates **s**.

- Every game must have at least one such optimum
- In zero-sum games all strategies are strictly Pareto efficient
- In common-payoff games all Pareto optimal strategy profiles have the same payoff

Pareto optimality

Alternate definition of Pareto optimality

- A solution **u**, i.e., a **utility vector $\mathbf{u(a_1)}$, …, $\mathbf{u(a_n)}$**, is **Pareto efficient**, i.e., **Pareto optimal**, if there is no other solution **u'** such that at least one agent is better off in **u'** than in **u** and no agent is worst off in **u'** than in **u**.
- Measures global good, does not require utility comparison
- Social welfare $\subset$ Pareto efficiency

Definitions

(N,A,U) a normal form game

- $A = A_1 \times \ldots \times A_n$ – A_i set of actions available to player ***i***.

Let **(N,A,U)** be a normal form game and, for any set **X,** let **P(X)** be the set of all probability distributions over **X**. Then the **set of mixed strategies** for player ***i*** is $S_i = \mathbf{P}(A_i)$

The **set of mixed strategy profiles** is the Cartesian product of the individual mixed strategy sets $S_1 \times \ldots \times S_n$.

$s_i(a_i)$ denotes the probability that an action a_i will be played under a mixed strategy s_i.

Definitions

The **support of a mixed strategy** $\mathbf{s_i}$ is

$$\{a_i \mid s_i(a_i) > 0\}$$

A **pure strategy** $\mathbf{s_i}$ is a strategy such that the support has size 1, i.e.

$$|\{a_i \mid s_i(a_i) > 0\}| = 1$$

A pure strategy plays a single action with probability 1

Expected utility of a mixed strategy

Given **(N,A,U)** a normal form game, the expected utility $\boldsymbol{u_i}$ of player ***i*** of the mixed strategy profile $s = (s_1,.., s_N)$ is defined as

$$u_i(s) = \sum_{a \in A} u_i(a) Pr(a|s)$$

Nash equilibrium

Definition $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$

Player's **i** **best response** to the strategy profile $\mathbf{s_{-i}}$ is a (mixed) strategy $\mathbf{s_i^* \in S_i}$ such that

$$u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i})$$

for all strategies $\mathbf{s_i \in S_i}$

Definition

A strategy profile $s = (s_1, \ldots, s_n)$ is a **Nash equilibrium** if for all agents ***i***, $\mathbf{s_i}$ is the best response to $\mathbf{s_{-i}}$

It is a **stable strategy**: no agent would want to change its strategy if he knew what strategies other agents were following

Nash equilibrium – alternate definition

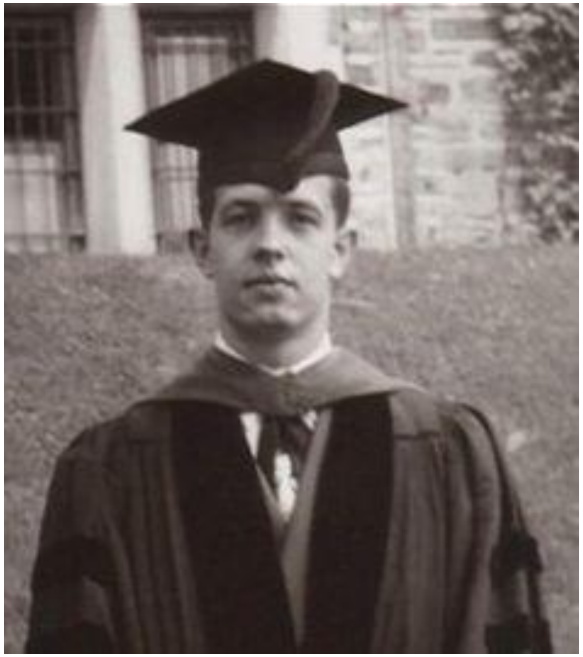
Definition

- Two strategies, S_1 of agent A and S_2 of agent B are in a **Nash equilibrium** if:
  - in case agent A follows S_1 then agent B can not do better than using S_2 **and**
  - in case agent B follows S_2 then agent A can not do better than using S_1.

Definition

- **The set of strategies** $\{S_1, S_2, \ldots, S_k\}$ used by the agents $A_1, A_2, \ldots, A_k$ is in a **Nash equilibrium** if, for any agent A_i, the strategy S_i is the best strategy to be followed by A_i if the other agents are using strategies $\{ S_1, S_2, \ldots, S_{i-1}, S_{i+1}, \ldots, S_k.\}$.

John Forbes Nash



Pure Nash Equilibrium

- no pure Nash equilibrum
- multiple Nash equilibria

Mixed strategy equilibrium

Theorem (Nash 1950)

- Every game with a finite number of payers and action profiles has at least one Nash strategy equilibrium

4. 1 Pure strategies

Prisoners' Dilemma

		Column player	
		Cooperate	Defect
Row player	Cooperate	3, 3	0, 5
	Defect	5, 0	1, 1

Cooperate = keep silent
Defect = betray the other

Pareto?
Pure Nash Equilibrium?

One pure Nash equilibrium

Pure strategies - Examples

Coordination game

		Column player	
		Left	Right
Row player	Left	1, 1	0, 0
	Right	0, 0	1, 1

Two pure Nash equilibria

BoS

	Football	Movie
Football	2, 1	0, 0
Movie	0, 0	1, 2

Pure strategies - Examples

Game of Chicken

		Column	player
		Swerve	Strait
Row player	Swerve	Tie, Tie	Lose, Win
	Strait	Win, Lose	Crash, Crash

		Column	player
		Swerve	Strait
Row player	Swerve	0, 0	-1, +1
	Strait	+1, -1	-10, -10

Two pure Nash equilibria

- Two firms are involved in developing a new technology that will allow consumers to taste food over the Internet. This has potential, for example, in restaurant promotion. Given the risks and the relatively small expected size of this market, compatibility of the technologies is very important.

- Firm **DigiTaste** is far advanced in developing its *RemoteTaste* technology .
- **WebOdor** has been expanding into the Internet taste arena with its product, *BitterWeb.*

- The two companies agree that if they both adopt the same technology, they each may gross $200M from the developing industry.
- If they adopt different technologies, consumers will make fun of both companies, and purchase neither product, leading to a gross of $0.
- Retooling one's factory to make the competing (nonproprietary) technology would cost **WebOdor** $100M and **DigiTaste** $250M.
- By the wave of an economist's wand, their production decisions must be made simultaneously.

- Set up the above scenario as a normal form (simultaneous) game.

What is the equilibrium outcome?

- Both adopt *RemoteTaste*
- Both adopt *BitterWeb*
- **DigiTaste** adopts *RemoteTaste* and **WebOdor** a dopts *BitterWeb*
- **WebOdor** adopts *RemoteTaste* and **DigiTaste** a dopts *BitterWeb*

4.2 Mixed strategies

Penny Matching

P2

P1

	Heads	Tails
Heads	1, -1	-1, 1
Tails	-1, 1	1, -1

- What are the equilibrium strategies now?
  - If P1 play Heads then P2 of course play Tails
  - but that makes P1 want to play Tails too
  - which in turn makes P2 want to play Heads
  - etc. etc. etc.

No pure Nash equilibrium

Penny Matching

	Heads	Tails
Heads	1, -1	-1, 1
Tails	-1, 1	1, -1

- But what if we can each (privately) *flip coins?*
  - – the strategy pair (1/2, 1/2) is an equilibrium

- There is a ***mixed strategies equilibrium***

- In a mixed Nash strategy equilibrium, each of the players must be <u>indifferent between any of the pure strategies</u> played with a probability.
- If this were not the case, then there is a profitable deviation: play the pure strategy with higher payoff with higher probability
- Remember that

Given **(N,A,U)** a normal form game, <u>the expected utility ***u_i*** of player ***i***</u> of the mixed strategy profile s = (s_1,.., s_N) is defined as

$$u_i(s) = \sum_{a \in A} u_i(a) Pr(a|s)$$

Mixed strategies

P2

Example

P1

	a (q)	b (1-q)
A (p)	2, 1	0, 0
B (1-p)	0, 0	1, 2

P1's expected payoff for A against (q, 1-q) = 2*q + 0*(1-q) = 2q

P1's expected payoff for B against (q, 1-q) = 0*q + 1*(1-q) = 1-q

If P1 is mixing on both strategies (A, B) then both must yield the same expected payoff

$$2q = 1-q \quad \rightarrow \quad q=1/3$$

P2's expected payoff for a against (p,1-p) = 1*p+0*(1-p) = p

P2's expected payoff for b against (p,1-p) = 0*p +2*(1-p) = 2–2p

$$p = 2-2p \quad \rightarrow \quad p=2/3$$

Mixed strategies

Matching Pennies

P2

P1

	Heads (q)	Tails (1-q)
Heads (p)	1, -1	-1, 1
Tails (1-p)	-1, 1	1, -1

$U_{P1}(H)=U_{P1}(T)$

$$1q + (-1)(1-q) = (-1)q + 1(1-q) \rightarrow q=1/2$$

$U_{P2}(H)=U_{P2}(T)$

$$(-1)p + 1(1-p) = 1p + (-1)(1-p) \rightarrow p=1/2$$

Mixed strategies

BoS

Husband

Wife

	Football (q)	Movie (1-q)
Football (p)	2, 1	0, 0
Movie (1-p)	0, 0	1, 2

Suppose both players randomise

- Wife is indifferent between F and M if she also mixes the 2 actions (otherwise she would be better off switching to a pure strategy according to which she only plays the best of her actions)

$U_{wife}(\text{Football})=U_{wife}(\text{Movie})$ if p is such that

$$2q + 0(1-q) = 0q + 1(1-q) \rightarrow p=1/3$$

$U_{hus}(\text{Football})=U_{hus}(\text{Movie})$ if q is such that

$$1p + 0(1-p) = 0p + 2(1-p) \rightarrow q=2/3$$

Mixed strategies

P1 / P2	A (x)	B (y)	C (1-x-y)
A (p)	1, 1	10, 0	-10, 1
B (q)	0, 10	1, 1	10, 1
C (1-p-q)	1, -10	1, 10	1, 1

- Consider Player 1. He plays column A with probability **p**, B with probability **q**, and C with probability **1−p−q**.
- Player 1 is indifferent between his pure strategies A,B,C. Same for Player 2
- He is indifferent between row A and row B if x,y are such that

$$1x+10*y+(-10)*(1-x-y) = 0*x+1*y+10*(1-x-y).$$

- He is indifferent between B and C if …

P1 / P2	A	B	C
A	1, 1	10, 0	-10, 1
B	0, 10	1, 1	10, 1
C	1, -10	1, 10	1, 1

- We construct the profile (p,q; x,y) such that the other player is indifferent between his pure strategies.

- So, no matter how the other player unilaterally deviates, his expected payoff will be identical to that in equilibrium (p,q; x,y)

Interpreting Mixed Strategy Equilibria

What does it mean to play a mixed strategy?

- Randomize to confuse your opponent
- Randomize when you are uncertain about the other's action
- Mixed strategies are a description of what might happen in repeated play