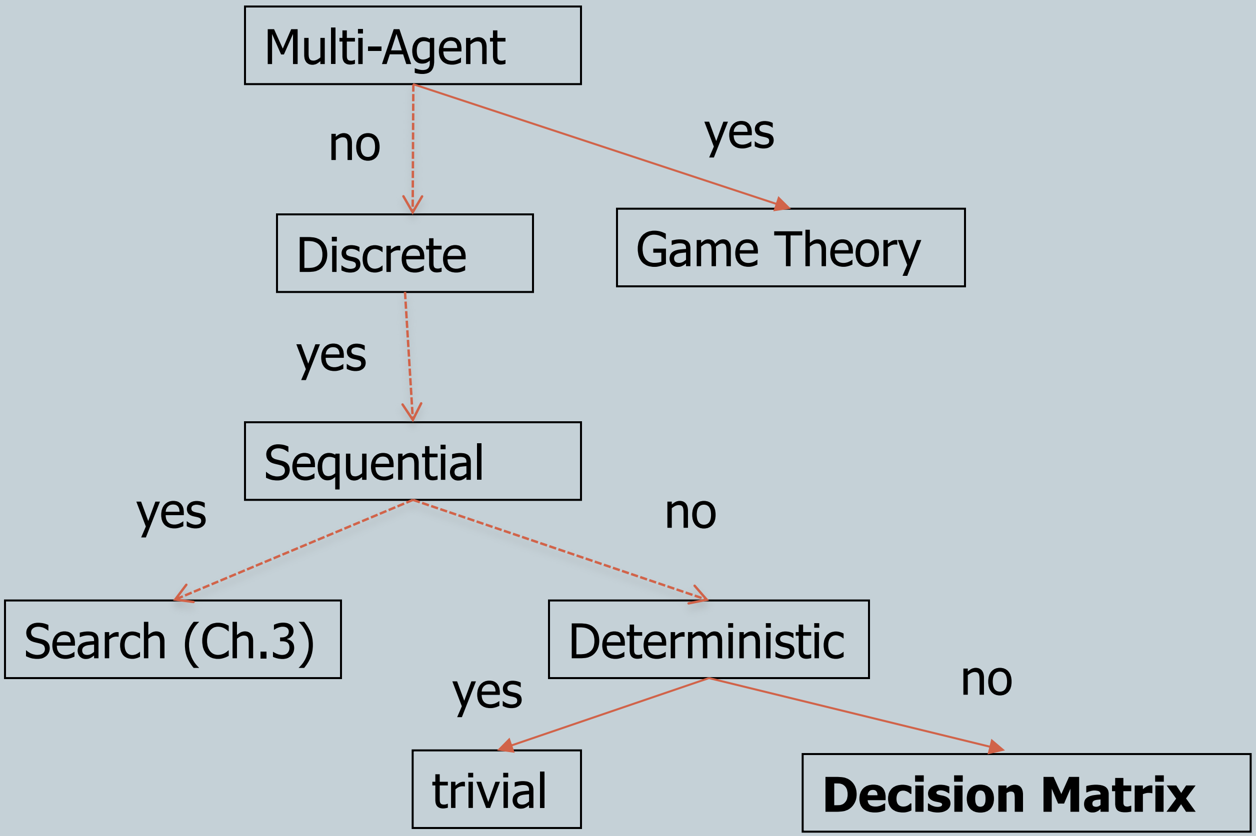
Rational Choice Theory

INTRO TO SINGLE AGENT DECISION MAKING

OLIVER SCHULTE

CMPT 310

Environment Type



Choice in a Deterministic Known Environment

- Without uncertainty, choice is trivial in principle: choose what you know to be the best option
- Trivial if the problem is represented in a look-up table

Option	Value
Chocolate	10
Wine	20
Book	15

- This is the standard problem representation in decision theory (economics).
- For examples see file intro-choice.doc
- Humans get confused even about these type of problems, see file bdt.doc and sunk cost survey.

Choice in a Stochastic Environment

- Need to incorporate uncertainty about the outcomes of choices.
  - There are different possible *states of the world*.

Acts/States of the World	Coin comes up heads	Coin comes up tails
Pay to play	$1	$0

What is the average or **expected value** of paying to play?

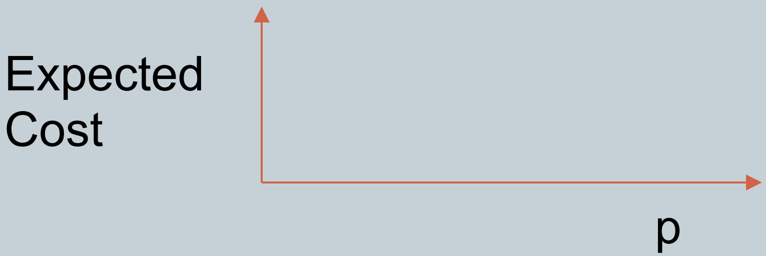
Paying for parking

Which option minimizes expected cost?

Acts/States	Get Caught (p)	Not get caught (1-p)
Pay	-\$3.00	-\$3.00
Don't pay	-\$100	\$0

Exercise:

1. Calculate the expected cost for Pay with p = 0%, 10%, 50%, 90%
2. Plot the expected cost for Pay in a graph
3. Calculate the expected cost for Don't Pay with p = 10%, 50%, 90%
4. Plot the expected cost for Don't Pay in the same graph



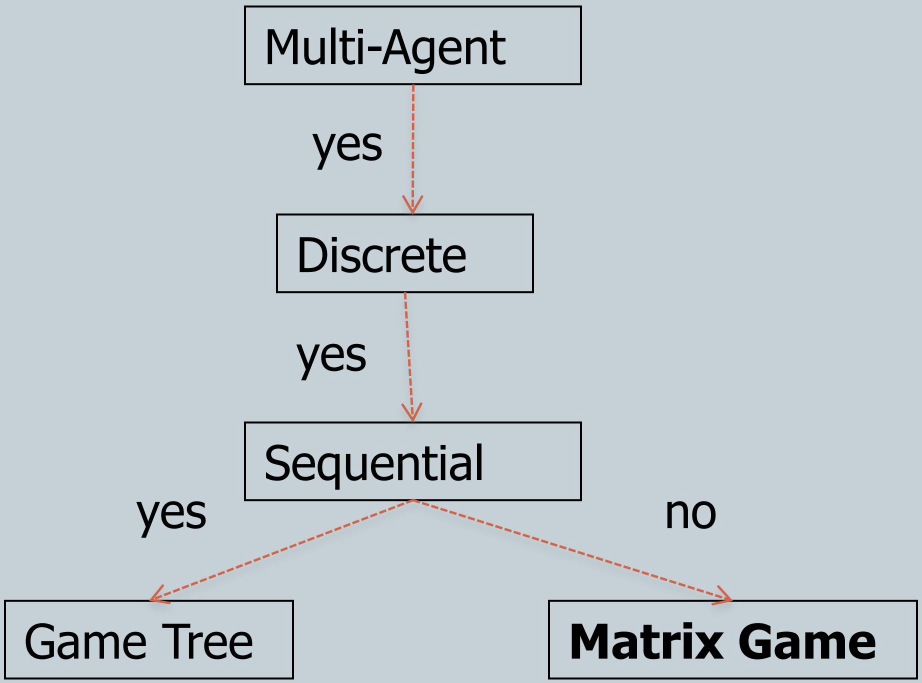
Game Theory

MULTI-AGENT DECISION MAKING

OLIVER SCHULTE

CMPT 310

Environment Type



Multi-Agent Interactions

- Game theory is *not* about games!
- Von Neumann-Morgernstern games (1944): a very general formalism for representing **multi-agent interactions**.
- Extensively used in economics, decision theory, operations research, political science.
  - Increasingly computer science, e.g., auction design.
- Zero-sum "parlour" games as special case (chess, poker etc.).
- 1997 Nobel prize (Nash, Selten, Harsanyi).

Multi-Agent Environment

- Recall from Ch.2: An environment is a multi-agent environment for an agent if it contains other agents whose decisions affect the agent's performance measure.
- Soccer Goalie

Matrix Games

Definition: 2-player matrix games

- Aka as strategic-form game.
- Each player has a set of options or **strategies**.
- Each player chooses one <u>independently</u> of the other player.
- Given a choice of strategies (s1,s2) by (player 1, player 2) each player receives a payoff aka **utility**.
- The utilities are the performance measure for the game.

Coordination Game

	Column Player	
Row Player	L	R
T	1,1	0,0
B	0,0	1, 1

A Story

	Driver 1	
Driver 2	Left	Right
Left	1,1	0,0
Right	0,0	1, 1

Driving on either side of the road is fine as long as it is the same side.

Another Story

	Carl	
Cecile	AQ	ASB
AQ	1,1	0,0
ASB	0,0	1, 1

- Carl and Cecile agreed to meet for coffee at a renaissance stand.
- They didn't specify which stand.
- They are okay with either as long as they meet up.

Battle of the Sexes

	Column Player	
Row Player	L	R
T	1,2	0,0
B	0,0	2, 1

A Story

	Martin	
Marcella	Theatre	Hockey
Theatre	1 ,2	0,0
Hockey	0,0	2, 1

- Martin and Marcella want to go on a date.
- Marcella prefers hockey, Martin the theatre.
- The worst outcome for both is if the date does not happen.

Another Story

	Patricians	
Plebeans	Boss	Worker
Worker	1,2	0,0
Boss	0,0	2, 1

- Early in its development, Republican Rome was divided into Patricians, the ruling class, and Plebeans, who were the workers, businesspeople, etc.
- The Plebeans were dissatisfied with this order, and one day they left the city and set up camp nearby. The Patricians argued that they should return.
- "In a healthy body, the stomach looks lazy, but it is necessary to digest the food that the other parts of the body procure. In the same way a healthy society must have a part that does not deal with everyday labour." The Plebeans were persuaded and returned.

The Prisoner's Dilemma

Prisoner's Dilemma

	Column Player	
Row Player	L	R
T	0,0	-2,2
B	2,-2	-1,-1

The Story

	Column Player	
Row Player	Not Confess	Confess
Not Confess	5 years each	Row: 9 years Col: 1 year
Confess	Row: 1 year Col: 9 years	7 years for each

- Two players are interrogated separately and offered a deal to confess to the crime.
- Exercise: Can be linearly transformed to numbers from slide before: what matters is relative utilities.

Another Story

	Pakistan	
India	As Is	Build Up
As Is	0,0	weaker: -2, stronger: 2
Build Up	stronger: 2, weaker: -2	-1 (billion), -1 (billion)

- Pakistan and India have often been military enemies.
- Each country can either build up its military, say spend $1 billion dollars on it, or stay at its present state of military capability.
- Assume that they have no important potential enemies except for each other, so that military might is only an advantage for a country to the extent that the country has more military strength than its neighbour.

Yet Another Story (I)

- In emergency situations often a general panic ensues that is bad for everyone.
- A common example is a fire breaking out in a closed room; amidst the general trampling and rushing, many fewer people get out than if they had maintained order.
- Let's look at the problem decision-theoretically.

The Panic Exercise

- Olivia is inside a disco with 200 other people when a fire breaks out. She does not know how everybody else is going to behave. To simplify things, let's say that either there will be order or there will be a general rush to the exit.
- If there is a general rush to the exit, Olivia will certainly want to rush too or else she will be the last person out with low chances of survival.
- But if there is order, Olivia would still prefer to rush herself because then she will get past anyone else who may be waiting and maximize her chances of getting out.
- If she is not going to have an advantage over others, she would rather that everybody wait than that everybody rushes, because many more people, including herself, are likely to get out if there is order.

Fill in the payoff matrix below to represent these preferences.

	Possible World States	
Options	General Rush	Order and line-up
Rush	2, 2	4
Wait	0	3, 3

Free-Riding and the Prisoner's Dilemma

- Free-riding opportunities create Prisoner's Dilemma
- An example is paying taxes.
  - According to some estimates, Canada loses between $10 billion and $15 billion annually due to corporate tax dodging using tax havens. (see Panama and Paradise papers)
  - Domestically, some estimates say that 60% of tax breaks intended for small businesses go to individuals earning over $150 K per year.

	Possible World States	
You	Other people pay their taxes	Other people avoid taxes
Pay Taxes	pay taxes + receive services	pay taxes – receive no services
Avoid Taxes	avoid taxes + receive services	avoid taxes - receive no services

Preparing for National College Entrance Exam

- Some Chinese students study 15 hrs/day for the gaokao, for an entire year.
- Why so much? fill in the matrix

	Possible World States	
You	Other people study 15 hours	Other people study 10 hours
15 hours	chance of a top 1% score = 1% - sleep	
10 hours		chance of a top 1% score = 1% + sleep

Adair Turner on the "zero-sum economy"

- Increasingly economic activity is about getting a bigger slice of the pie, not growing it

➢ Another Prisoner's Dilemma

Dominance

Strong Dominance

- An option *a* **strongly dominates** another option *a*' if

  for all possible states of the world, option *a* yields a better outcome than *a* '.
- With a strongly dominant option "you can only win".
- Also called strict dominance in game theory.

Example

Harvey has been thinking for weeks about asking Sabrina on a date. Finally he decides that he would rather ask her and be rejected than not ask her and live in the agony of hesitation. Obviously he prefers her saying yes to not asking at all. Given these preferences, asking Sabrina strongly dominates not asking her.

	Possible States	
Options	Sabrina likes Harvey	Sabrina doesn't like Harvey
ask out	she says yes	she says no
don't ask out	hesitation	hesitation

Example

	Column Player Choices	
Row Player	L	R
T	0	-2
B	2	-1

Prisoner's Dilemma from the Row Player's Point of View.

Weak Dominance

An option *a* **weakly dominates** another option *a*' if

1. for all possible worlds, the outcome of option *a* is at least as good the outcome for option *a*'.
2. for some possible worlds, the outcome of option *a* is better than the outcome for option *a*'.

With a weakly dominant option "you can't lose and you might win".

Example

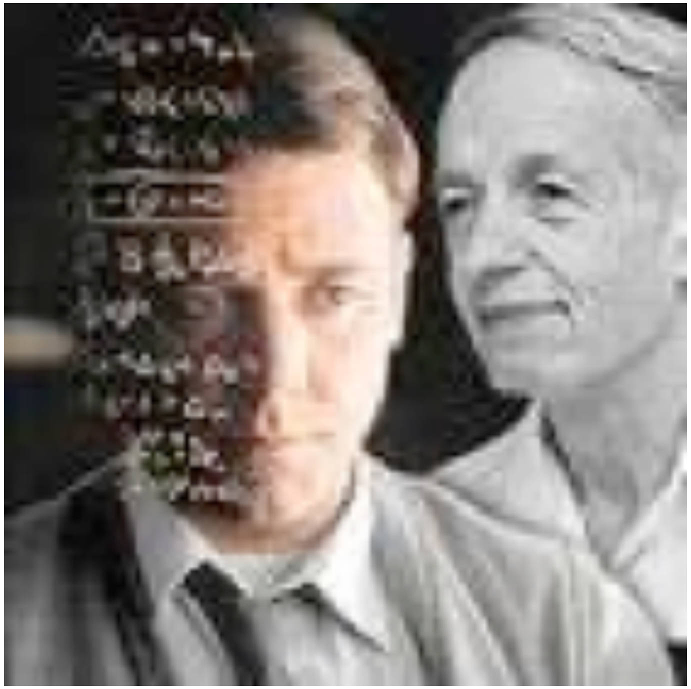
- Somebody sends you some books to check out, "with no obligation to buy".

	Possible States	
Options	like book	don't like book
check out book	got good book 👍	no book, no money spent
return book	no book, no money spent	no book, no money spent

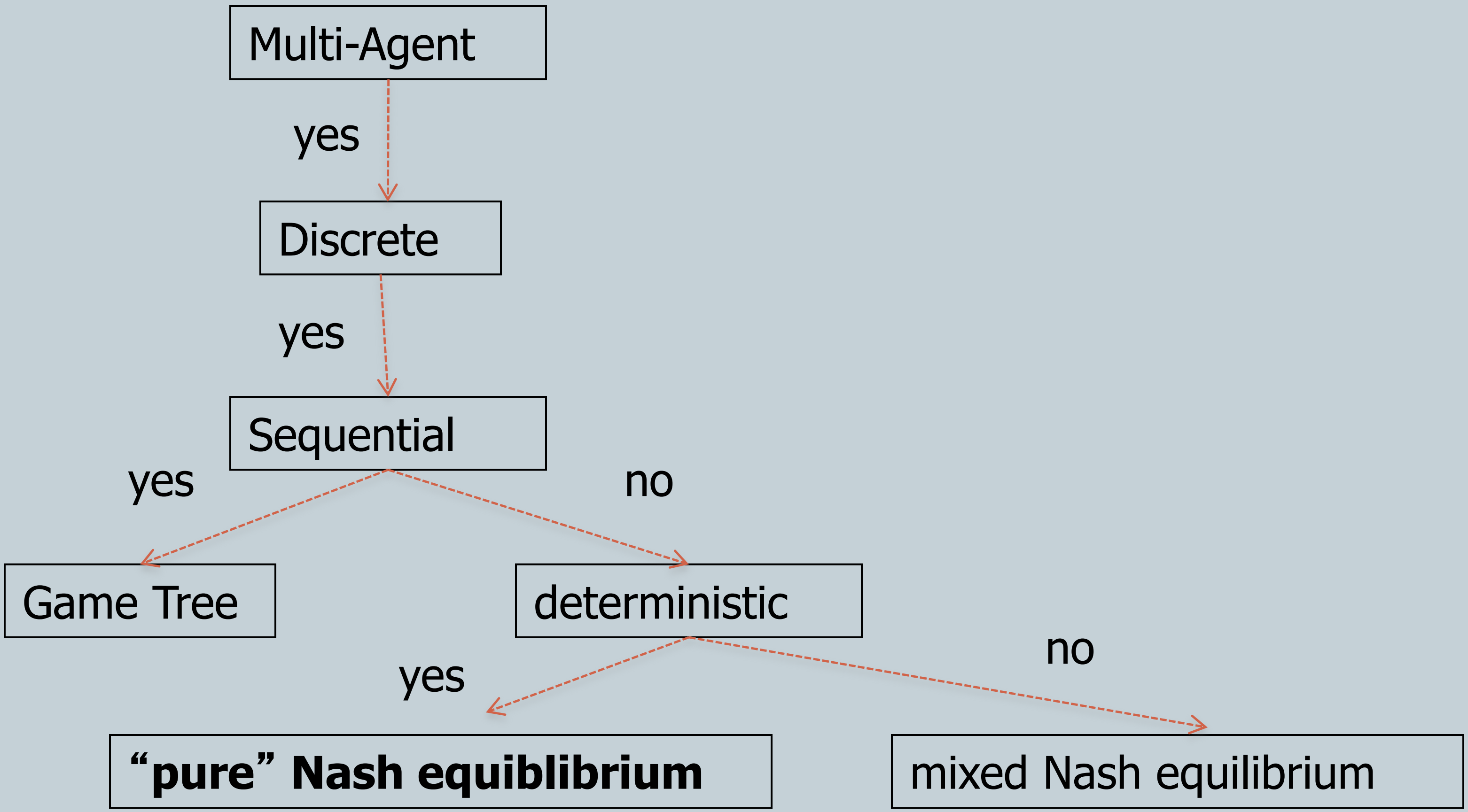
Exercise

- Jennifer is not satisfied with her grade on an assignment.
- She considers two options, accepting the grade, and taking up the instructor's office hour with a discussion of her grade.
- Model Jennifer's situation with a decision matrix with two options and two possible worlds.
- Assume that Jennifer prefers a better grade, and does not mind spending an hour discussing with her instructor.
- Show that going to discuss the grade weakly dominates not discussing it.

Nash Equilibrium



Environment Type Discussed In this Lecture



Definition

- A strategy s_1 for player 1 is a **best response** against a strategy s_2 for player 2 iff there is no strategy s'_1 that does better against s_2 than s_1 does.
- Similarly, a strategy s_2 for player 2 is a **best response** against a strategy s_1 for player 1 iff there is no strategy s'_2 that does better against s_1 than s_2 does.
  - Note that a player can have more than one best response against another player's strategy.
- A **pair of strategies** (s_1, s_2) is a **Nash equilibrium** iff s_1 is a best response against s_2 **and** s_2 is a best response against s_1.

Example: Coordination Game

	Column Player	
Row Player	L	R
T	1,1	0,0
B	0,0	1, 1

- The pairs (T,L) and (B,R) are each a Nash equilibrium.

Exercise

	Column Player	
Row Player	L	R
T	1,2	0,0
B	0,0	2, 1

- How many pure (deterministic) Nash Equilibria are there in the Battle of the Sexes?

Exercise

	Column Player	
Row Player	L	R
T	**0,0**	-2,2
B	2,-2	**-1,-1**

- How many Nash Equilibria are there in the Prisoner' s Dilemma? (See Canvas quiz)

Justifications of Nash equilibrium

Economists predict routinely that players will play Nash equilibria. There are two main justifications for this.

1. **The steady state interpretation.** If players were to play the game repeatedly, they would keep improving their responses until they are each playing a best reponse.
2. **The self-enforcing agreement interpretation.** Suppose the players agreed in advance how they were going to play. Then they could trust each other to keep their agreement if and only if that agreement is a Nash equilibrium.
Social conventions can establish Nash equilibria.

Games With Unpredictability

Matching Pennies

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

The Story

- The row player wins if both player make the same choice.
- The column player wins if the players make different choices.

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

Exercise

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

- How many Nash Equilibria are there in Matching Pennies?

Exercise

- Army A has a single plane with which it can strike one of two possible targets.
- Army B has one anti-aircraft gun that it can assign to one of the targets.
- If Army A attacks a target, A destroys the target if Army B' s anti-aircraft gun does not defend the target. If B' s gun is defending the target that A attacks, the target is safe.
- Model this situation as a game where A has 2 options "attack target 1" and "attack target 2", and B has 2 options "defend target 1" and "defend target 2".
- Army A does not care which target it destroys, but prefers destroying a target to not destroying one.
- Army B prefers both targets to be safe, but is indifferent between target 1 and target 2 being destroyed.

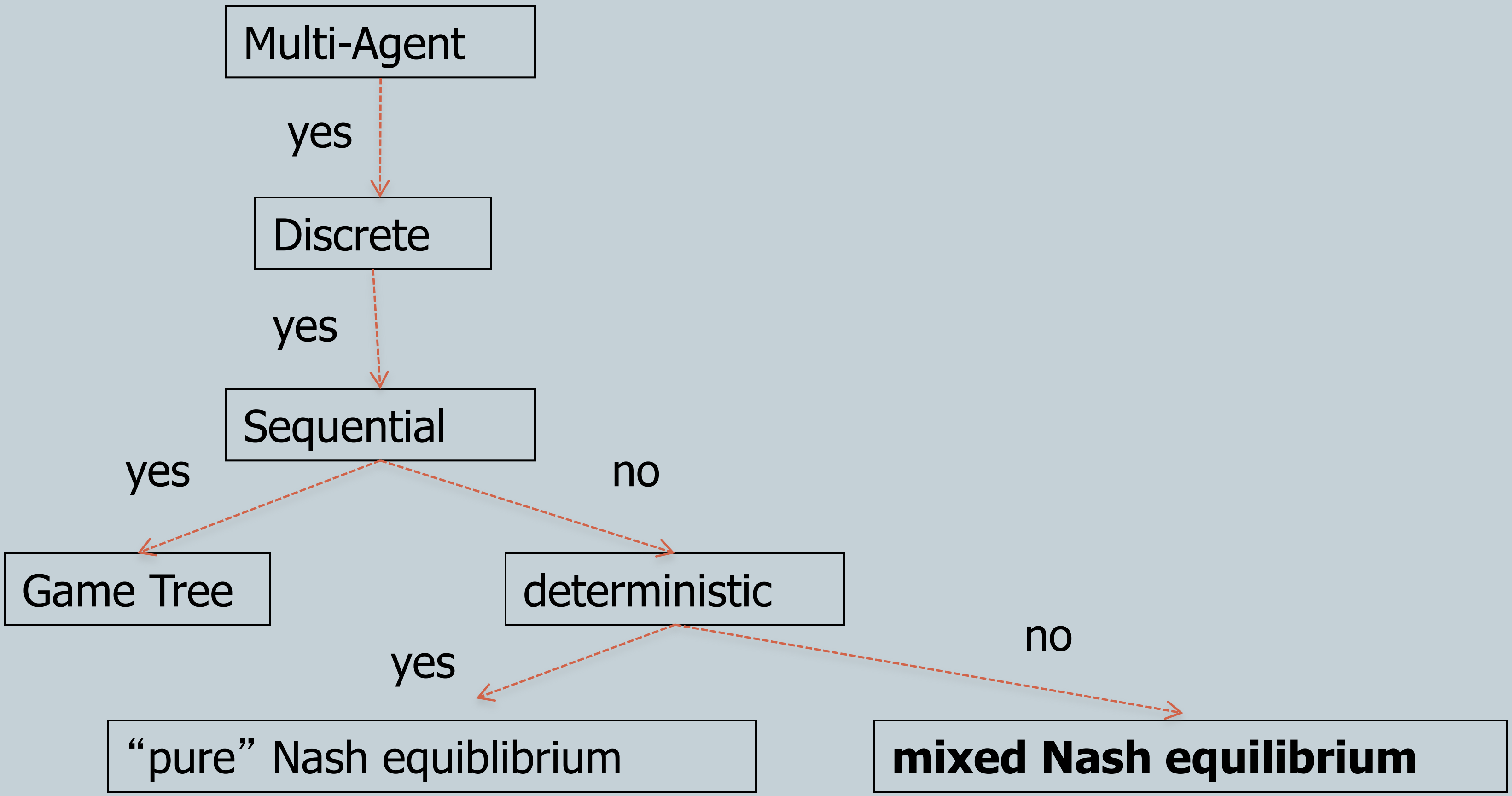
	Army B	
Army A	defend target 1	defend target 2
attack target 1		
attack target 2		

Exercise

- Represent Rock, Paper, Scissors as a 3x3 matrix game.
- How many pure (deterministic) Nash equilibria are there in RPS?

Mixed Strategies

Environment Type Discussed In this Lecture



Motivation

- Some games like Rock, Paper, Scissor don't have a Nash equilibrium as defined so far.
- Intuitively, the reason is that there is no steady state where players have *perfect knowledge* of each other's actions: knowing exactly what the other player will do allows me to achieve an optimal payoff at their expense.
- Von Neumann and Morgenstern observed that this changes if we allow players to be **unpredictable** by choosing **randomized strategies.**
- Random choices are used in many areas of AI, recall vaccum robot.

Definition

- Consider a matrix game with finitely many strategies.
- A **mixed strategy** is a probability distribution over the strategies.
- Given two mixed strategies (σ_1, σ_2), the payoff to each player is the **expected payoff** from randomizing in accordance with the strategies.
- Mixed Strategies in Soccer Penalties

Example: Matching Pennies

	Heads: 1/2	Tails: 1/2
Heads: 1/2	1,-1	-1,1
Tails: 1/2	-1,1	1,-1

- Row's mixed strategy is to play Heads ½ the time, Tails ½ the time.
- Column's mixed strategy is the same.

Payoff From Mixed Strategies

- Given a pair of mixed strategies (σ_1, σ_2), the payoff to each player is the **expected payoff** from randomizing in accordance with the strategies.
- Consider for example the row player in a 2x2 game.
- The expected payoff for T is the expected value given that each action by column is chosen with a given probability.
- The expected payoff from σ_1 is given by payoff(T, σ_2) x σ_1(T) +payoff(B, σ_2) x σ_1(B).

Example: Matching Pennies

	Heads: q = 1/2	Tails: 1/2
Heads: p = 1/2	1,-1	-1,1
Tails: 1/2	-1,1	1,-1

row payoff(Heads,q) = 1/2 x 1 + 1/2 x -1 = 0
row payoff(Tails,q) = 1/2 x -1 + 1/2 x 1 = 0
column payoff(Heads,p) = 1/2 x 1 + 1/2 x -1 = 0
column payoff(Tails,p) = 1/2 x -1 + 1/2 x 1 = 0

For Row player: payoff(p,q) = 1/2 x payoff(Heads,q) + 1/2 x payoff(Tails,q) =0
For Column player: payoff(p,q) = 1/2 x payoff(Heads,p) + 1/2 x payoff(Tails,p) = 0.

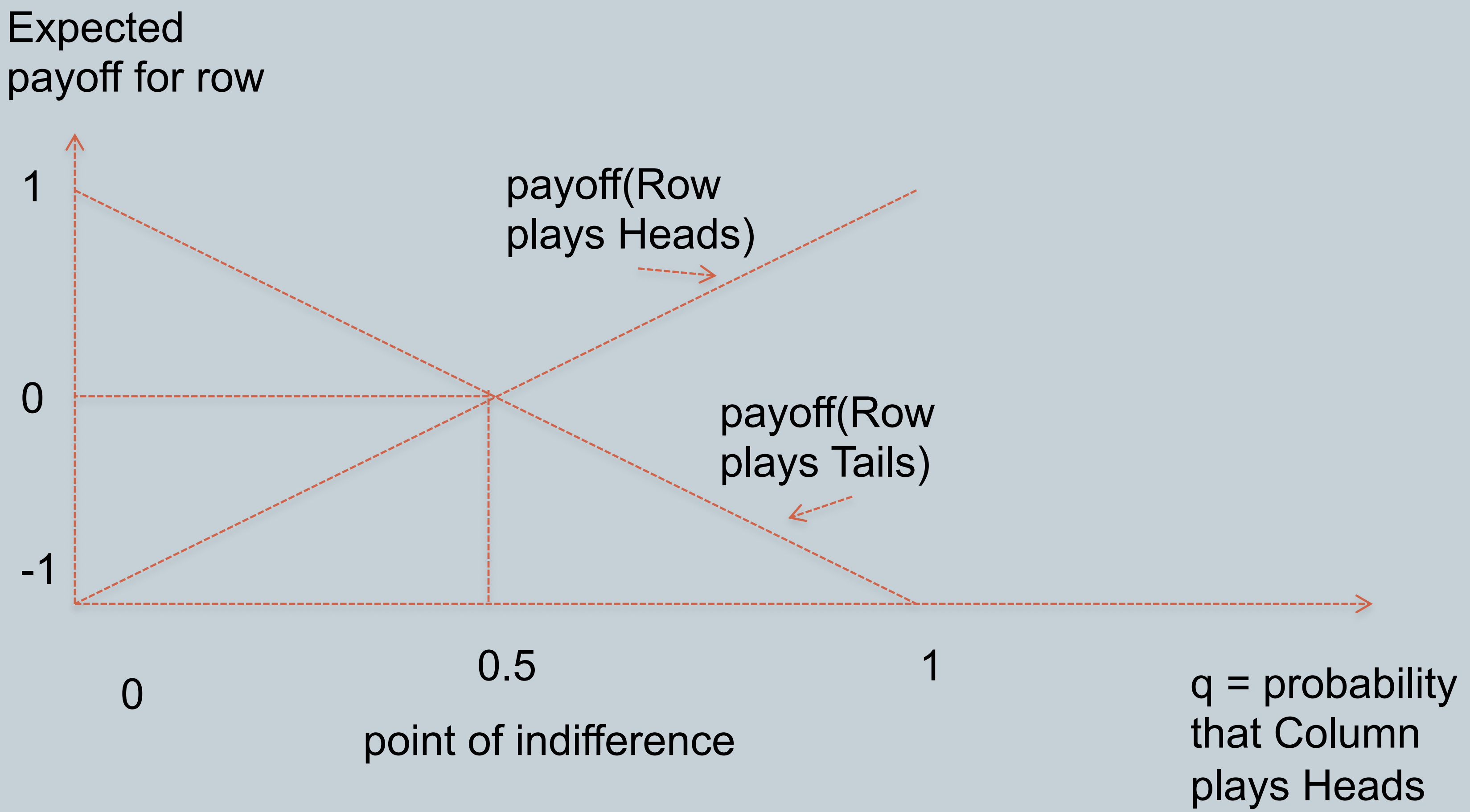
Example: Matching Pennies

	Heads: q = 1/2	Tails: 1/2
Heads: p = 1/2	1,-1	-1,1
Tails: 1/2	-1,1	1,-1

row payoff(Heads,q) = 1/2 x 1 + 1/2 x -1 = 0
row payoff(Tails,q) = 1/2 x -1 + 1/2 x 1 = 0
column payoff(Heads,p) = 1/2 x 1 + 1/2 x -1 = 0
column payoff(Tails,p) = 1/2 x -1 + 1/2 x 1 = 0

For Row player: payoff(p,q) = 1/2 x payoff(Heads,q) + 1/2 x payoff(Tails,q) =0
For Column player: payoff(p,q) = 1/2 x payoff(Heads,p) + 1/2 x payoff(Tails,p) = 0.

Visualization of Row Payoff Function



Abstract Representation

	L: q	R: 1-q
T: p	a1,b1	a2,b2
B: 1-p	a3,b3	a4,b4

- The mixed strategy of player 1 is represented by p, for player 2 by q.

payoff(T,q) = q x a1 + (1-q) x a2
payoff(B,q) = q x a3 + (1-q) x a4
payoff(L,p) = p x b1 + (1-p) x b3
payoff(R,p) = p x b2 + (1-p) x b4

For Row player: payoff(p,q) = p x payoff(T,q) + (1-p) x payoff(T,q)
For Column player: payoff(p,q) = q x payoff(L,p) + (1-q) x payoff(R,p).

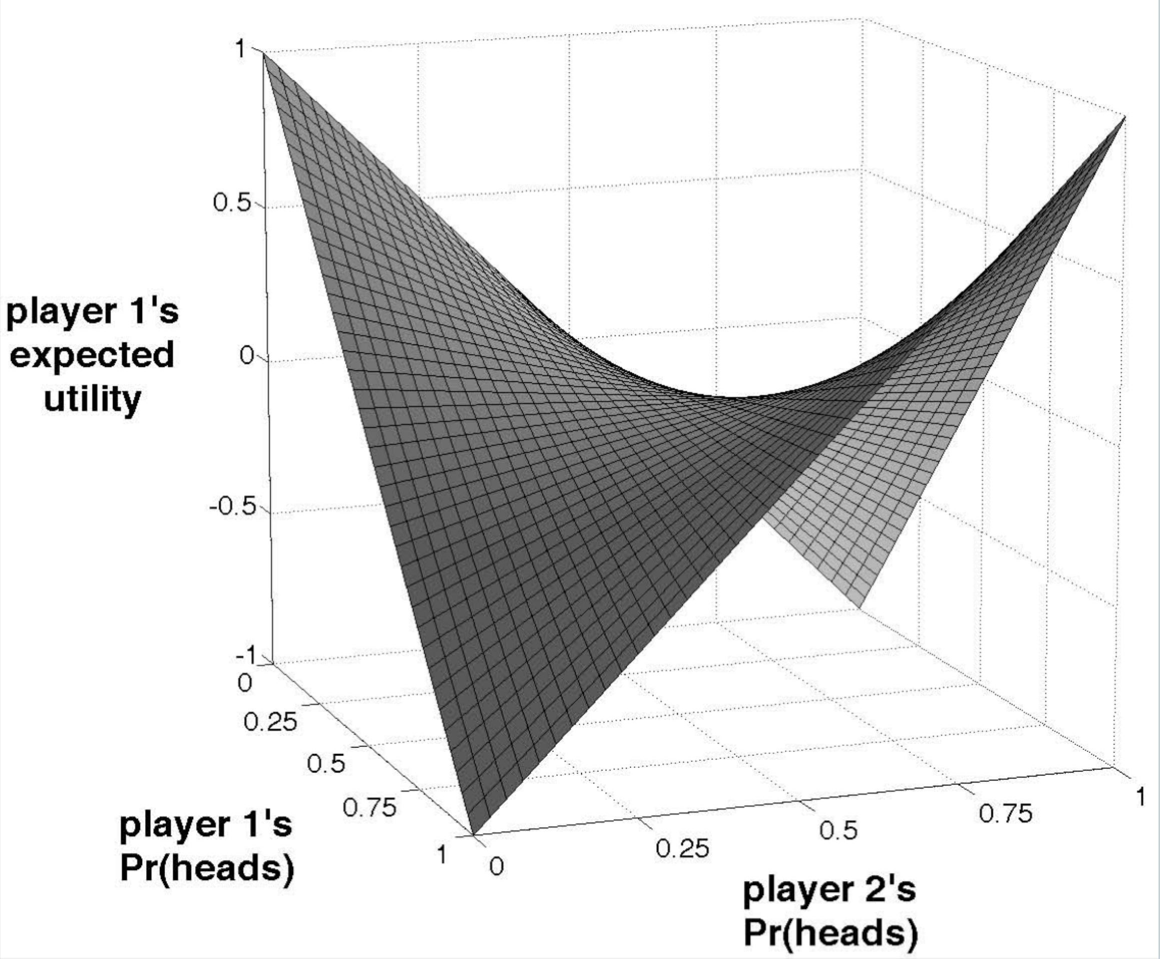
Finding Mixed Strategies

NASH EQUILIBRIUM

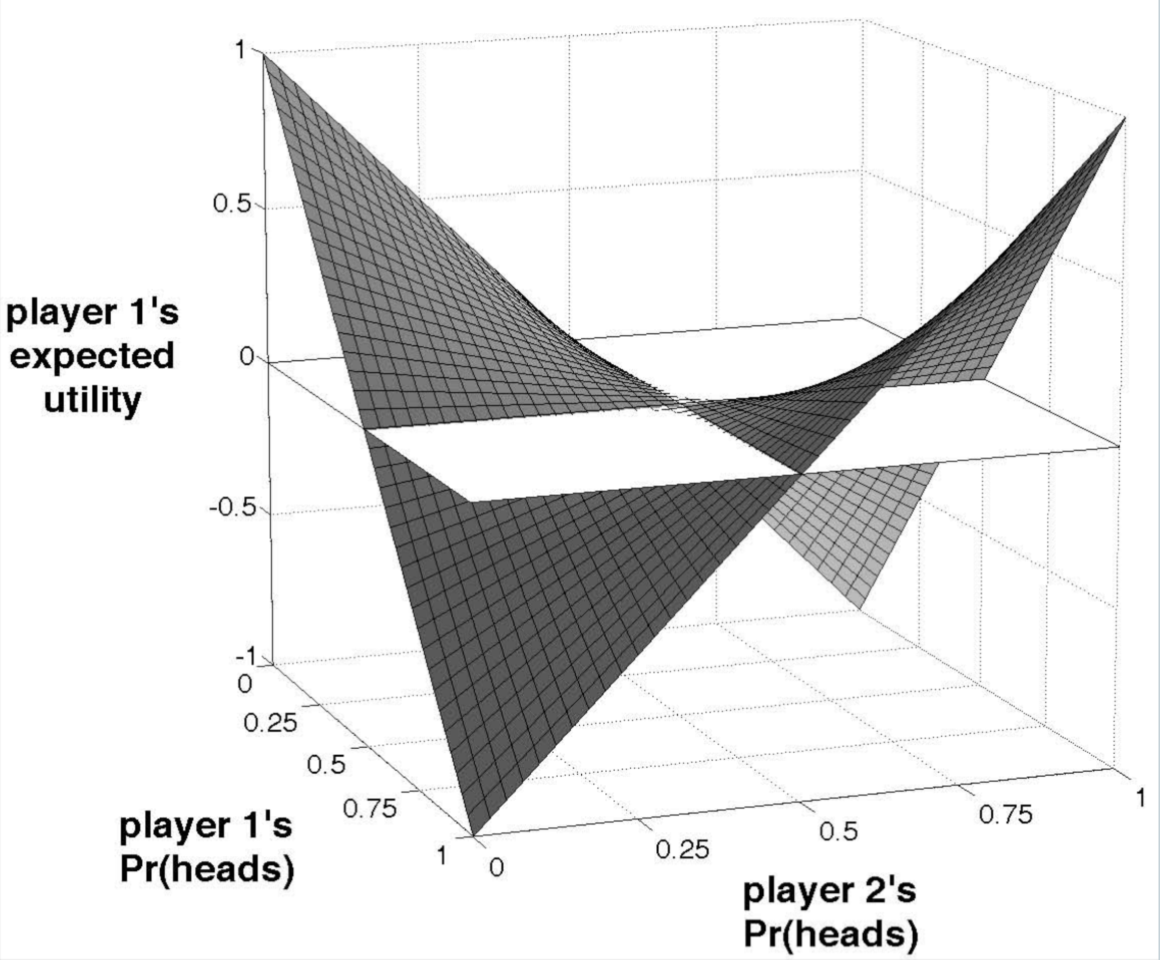
Nash equilibrium in mixed strategies

- The definition of best response is as before: a mixed strategy σ_1 is a **best response** to σ_2 if and only if there is no other mixed strategy σ'_1 with $payoff(\sigma'_1, \sigma_2) > payoff(\sigma_1, \sigma_2)$.
- Similarly, two mixed strategies (σ_1, σ_2) are a **Nash equilibrium** if each is a best response to the other.
- Example: In Matching Pennies, a Nash equilibrium obtains if each player chooses Heads or Tails with equal probability. See payoff function graph.

Visualization of Matching Pennies Payoff Function



Visualization of Matching Pennies equilibrium



Exercise: Coordination Game

	Column Player	
Row Player	L 1/2	R 1/2
T	1,1	0,0
B	0,0	1, 1

- Draw row' s payoff function graph.
- Use this to verify that a Nash equilibrium obtains if each player chooses their option with equal probability.

Exercise

- Guess a mixed strategy equilibrium for Rock, Paper, Scissors.

Existence of Nash Equilibrium

- Theorem (Nash 1950). In any finite game (any number of players) there exists a Nash equilibrium.
- Short proof by Nash.

Computing Mixed Strategy Equilibria

Basic Insight

- You have 10 free hours/wk to work.
- You need to allocate the 10 hours between two jobs.
  1. Job 1 pays $30/hr
  2. Job 2 pays $25/hr
- What fraction of your time goes to Job 1?
- What if you have three job offers?
  1. Job 1 pays $30/hr
  2. Job 2 pays $25/hr
  3. Job 3 pays $30/hr

The Support Theorem

- **Definition** The **support** of a probability measure *p* is the set of all points *x* s.t. *p*(*x*)>0.
- **Theorem**. A mixed strategy pair (σ_1, σ_2) is an N.E. if and only if for all pure strategies a_i in the support of σ_i, the strategy a_i is a best reply so σ_{-i}.
  σ_{-i} is the mixed strategy for the player who is not i.
- **Corollary.** If (σ_1, σ_2) is an N.E. and a_1,b_1 are in the support of σ_1, then $payoff(a_1, \sigma_2) = payoff(b_1, \sigma_2)$. That is, player 1 is indifferent between a_1 and b_1. The same for a_2,b_2 in the support of s_2.
- In Nash equilibrium, all strategies that can be chosen with positive probability must be optimal, therefore *equally good*.
  - *Like Job 1 and 3 in the previous example.*

A general NP procedure for finding a Nash Equilibrium

1. Choose support for player 1, support for player 2.
2. Check if there is a Nash equilibrium with those supports.

Example: Coordination Game (I)

	Column Player	
Row Player	L	R
T p	1,1	0,0
B 1-p	0,0	1, 1

1. Find all deterministic equilibria where neither side randomizes. These are (B,R) and (T,L).

Coordination Game, Column's Point of View

Is there an equilibrium where both players randomize?

	Column Player	
Row Player	L	R
T p	1	0
B 1-p	0	1

1. Suppose that Row chooses T with probability p where $0<p<1$. Let us find the optimal replies for Column. We have
   1. payoff(L,p) = p x 1+ (1-p) x 0 = p
   2. payoff(R,p) = 1-p .
2. The support theorem says that if Column randomizes, both L and R must be equally good. So payoff(L,p) = payoff(R,p), thus p = 1-p which entails that p = 50%.
3. **So in any mixed equilibrium, row player chooses T or B with equal probability.**

Coordination Game, Rows' s Point of View

Is there an equilibrium where both players randomize?

	Column Player	
Row Player	L q	R 1-q
T	1	0
B	0	1

1. Suppose that Column chooses L with probability q where 0<q<1. Let us find the optimal replies for Row. We have
   1. payoff(T,q) = q x 1+ (1-q) x 0 = q.
   2. payoff(B,q) = 1-q .
2. The support theorem says that if Row randomizes, both T and B must equally good. So payoff(T,q) = payoff(B,q), thus q = 1-q which entails that q = 50%.
3. **So the only mixed equilibrium is one where both players choose both options with equal probability.**

Example: Matching Pennies

Find all the equilibria where both players randomize.

	L: q	R: 1-q
T: p	1,-1	-1,1
B: 1-p	-1,1	1,-1

1. Suppose that Row chooses T with probability p where 0<p<1. Let us find the optimal replies for Column. We have
   1. payoff(L,p) = p x -1+ (1-p) x 1 = 1 -2p
   2. payoff(R,p) = p x 1 + (1-p) x -1 = 2p - 1 .
2. The support theorem says that if Column randomizes, both L and R must equally good. So payoff(L,p) = payoff(R,p), thus 1-2p = 2p-1 which entails that p = 50%.
3. The game is symmetric between Row and Column. Thus the same argument shows that if Column randomizes, then q = 50%.
4. **So the only mixed equilibrium is one where both players choose completely randomly.**
5. There is no deterministic equilibrium, so there is only this random equilibrium.

Abstract Representation

	L: q	R: 1-q
T: p	a1,b1	a2,b2
B: 1-p	a3,b3	a4,b4

payoff(T,q) = q x a1 + (1-q) x a2
payoff(B,q) = q x a3 + (1-q) x a4
payoff(L,p) = p x b1 + (1-p) x b3
payoff(R,p) = p x b2 + (1-p) x b4

- If both players randomize, then $0 < p < 1$ and $0 < q < 1$.
- At Nash equilibrium, players choose only optimal strategies so must have

1. payoff(T,q) = payoff(B,q) so q x a1 + (1-q) x a2 = q x a3 + (1-q) x a4
2. payoff(L,p) = payoff(R,p) so p x b1 + (1-p) x b3 = p x b2 + (1-p) x b4.

Examples

Real-World Examples

- Choosing serve in tennis.
- Penalty kick strategy in soccer.
- Dung Fly Mating. Mating Video
  - Dung Flies mate at cowpats.
  - Females visit less as cowpat gets stale.
  - Male choice: leave early, leave late.
  - The percentage of males leaving early is exactly that predicted by mixed Nash equilibrium.
- Ratio of orange, blue, yellow side-blotched lizards predicted by mixed Nash equilibrium.

Population and Traffic Models

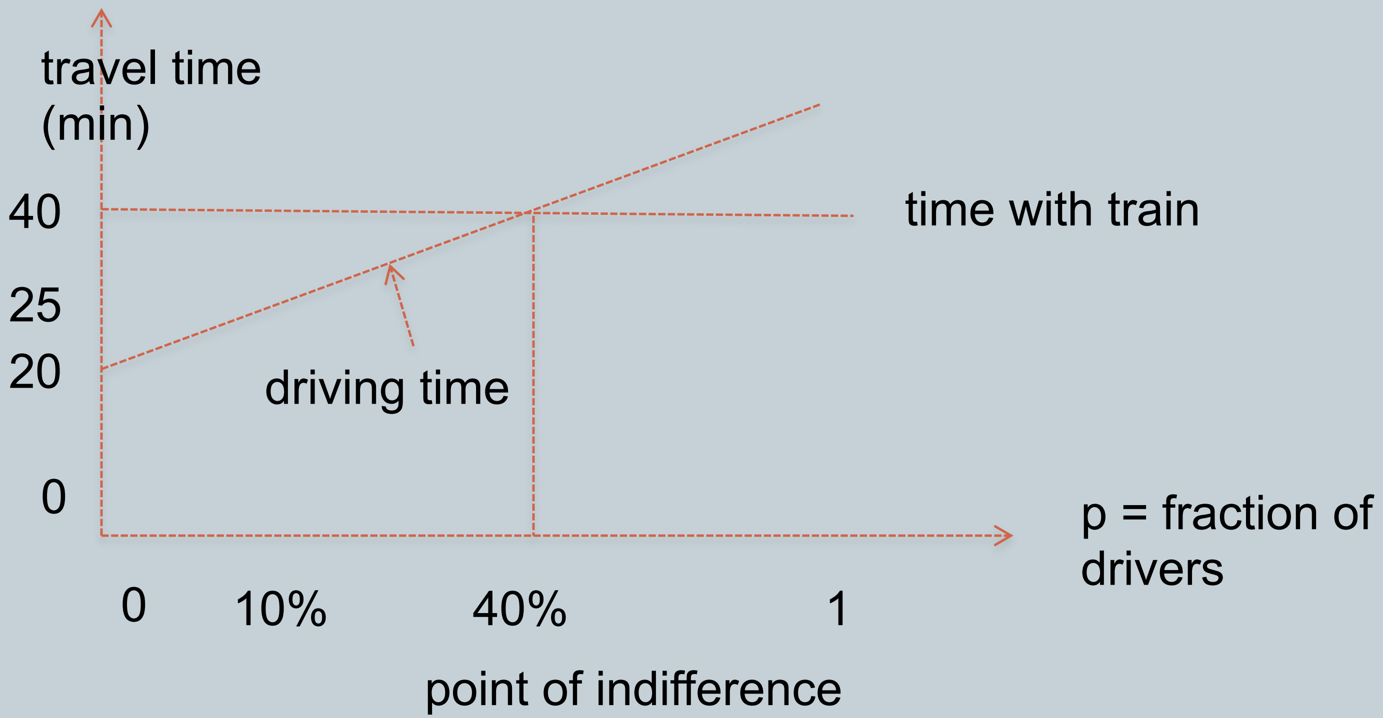
- Consider 1,000 people commuting from Coquitlam to downtown.

1. The train takes 40 min no matter what.
2. Driving takes 20 min if there are 0 other cars. Every 100 cars adds 5 min.

Number of Cars	Driving Time
0	20 min
100	25 min
400	40 min

1. With 1000 cars total, 1% of cars is 10 cars. So how many minutes does each 1% of cars add?
2. At equilibrium, the driving time on the highway is the same as the commuting time on the train. What is the percentage of driving commuters at which we reach equilibrium?

Visualization of Traffic Equilibrium



Conclusion

Conclusion (I)

- Matrix games model many frequent scenarios in Multi-Agent interactions.
- Combine cooperation and competition.
  - Coordination.
  - Battle of the Sexes.
  - Prisoner’ s Dilemma.
  - Matching Pennies.
  - Chicken (not discussed)

Conclusion (II)

- Nash equilibrium predicts outcome of the game.
  - Self-enforcing choices by both players.
  - Mixed equilibrium always exists.
- Dominant strategies are optimal in multi-agent interactions.
- Rational strategies are not always best for the group
  - In the Prisoner’s Dilemma, both players choosing a strongly dominant strategy does not lead to a Pareto-optimal outcome
    - Rationality makes everybody worse off than irrationality