Slide 1

Maurice Pagnucco
Knowledge Systems Group
Department of Artificial Intelligence
School of Computer Science and Engineering
The University of New South Wales
NSW 2052, AUSTRALIA
`morri@cse.unsw.edu.au`

Slide 2

Game Trees

Matrix games assume players choose strategy simultaneously without knowledge of what other player is choosing

In real situations decisions made sequentially and information about previous choices becomes available to players as situation develops

We shall introduce another way of modelling situations

Slide 3

Structure of Game Trees

A game tree is structured as follows:

Each *node* labelled by player making choice

Each *branch* labelled with particular choice (of action) made by player

Each *leaf node* labelled with payoff to players (convention: payoff to one of the players)

Chance events (e.g., roll of die, dealing of cards, . . .) must also be represented

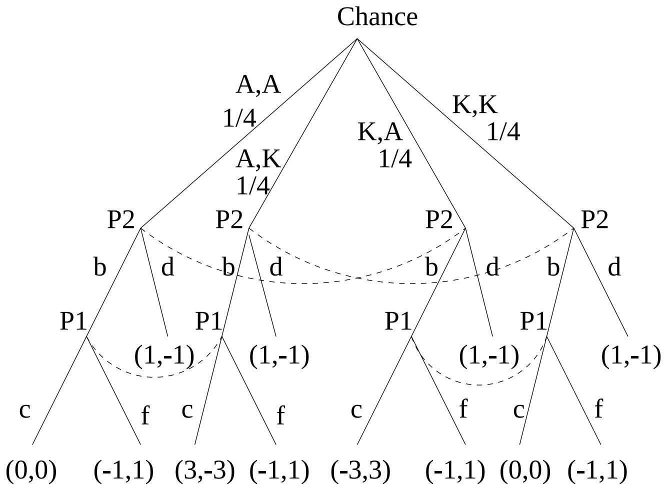
In this case, node labelled *Chance* and branches labelled with probability that Chance will come up with that choice

Slide 4

Example (Straffin 1993, p. 38)

Two players start by putting $1 in the pot. Each player is dealt one card from a deck of aces and kings. Player 2 must either bet $2 to continue or drop their hand letting Player 1 win the pot. If Player 2 bets, Player 1 must either call by matching Player 1's bet or fold. If Player 1 folds, Player 2 wins the pot. If Player 1 calls, the two players compare cards with the higher card winning or the pot is split evenly in the case of a draw

Slide 5



b = bet; d= drop
c = call; f = fold

Slide 6

Information Sets

In some situations a player may not know where they are in the tree (e.g., after the deal)

Nodes which represent the player's current situation given the information at their disposal but are distinct in the tree due to chance factors form an *information set*

Nodes in the same information set are linked via dotted lines (sometimes in the literature circled in dotted regions)

Slide 7

Strategy in Game Tree

A strategy (action) in a game tree corresponds to a player's complete description of choice to be made at any information set in the tree

Knowing strategies of players we can determine course of play (except for Chance)

Knowing Chance's probabilities we can calculate expected payoffs

Mapping Game Tree to Game Matrix

1. label rows and columns of matrix with players' possible strategies
2. place expected payoffs in entries of matrix

However, number of strategies may be enormous!

Slide 8

Continuing Straffin Example

Examine strategy where Player 2 bets only holding an Ace and Player 1 calls only holding an Ace

Prob	**Hands** (1, 2)	**Outcome**	**Payoff** (to 1)
$\frac{1}{4}$	A, A	2b, 1c	0
$\frac{1}{4}$	A, K	2d	1
$\frac{1}{4}$	K, A	2b, 1f	-1
$\frac{1}{4}$	K, K	2d	1

Expected payof $= \frac{1}{4}.0 + \frac{1}{4}.1 + \frac{1}{4}.(-1) + \frac{1}{4}.1 = \frac{1}{4}$

Game matrix:

Slide 9

Player 2 bets

		Always	**A Only**	**K Only**	**Never**
	Always	0	$-\frac{1}{4}$	$\frac{5}{4}$	1
Player 1 calls	**A Only**	$\frac{1}{4}$	$\frac{1}{4}$	1	1
	K Only	$-\frac{5}{4}$	$-\frac{1}{2}$	$\frac{1}{4}$	1
	Never	-1	0	0	1

Slide 10

Games of Perfect Information

– No nodes labelled Chance
– Information sets all consist of a single code

Chance has no role in the game

Players know all preceding moves

Such games can be analysed by *truncation* (or *tree pruning*)

Slide 11

Truncation

Start at leaves of game tree.

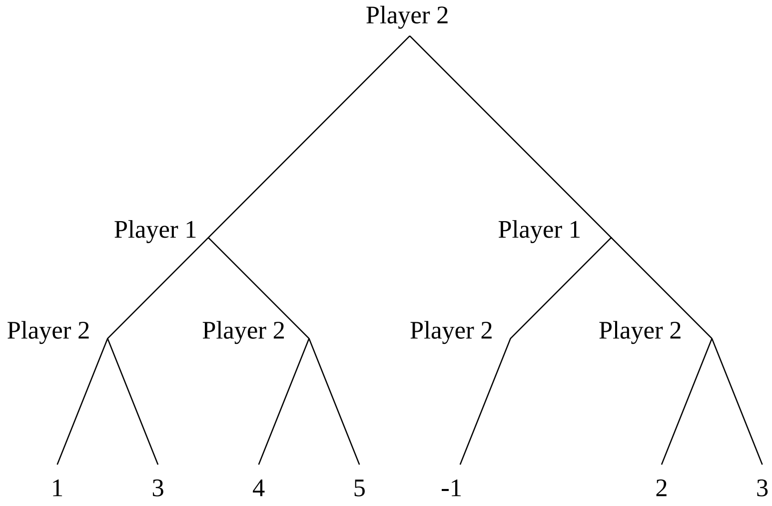
For all leaves connected by branches to the same node one level higher up in the tree select the value at that leaf representing the best choice for the player labelling the node

Delete all leaves of this node and propagate the best value up to the node

Continue this process all the way to the root of the tree

Slide 12

Example



Slide 13

Games of Perfect Information

Tree truncation is akin to (iterated) removal of dominated strategies in matrix games to find a saddle point

This process will work for any two-person zero-sum games of perfect information (just follow the truncation procedure)

Therefore, all two-person zero-sum games of perfect information have a saddle point

(Zermelo 1912) In a finite game of complete information one of the two players has a strategy that can force a win no matter what the other player does

Slide 14

Utility Theory

Where do the numbers come from?

How important are they?

How do we assign numbers to outcomes?

Utility theory: science of assigning numbers to particular outcomes so as to reflect agent's underlying preferences

Conside a game with a saddle point

		Red	
		1	**2**
Blue	**A**	4	3
	B	7	1

All we need, however, to locate this equilibrium pair is that the outcome be the smallest in its row and the largest in its column

That is, row player prefers this outcome to any other outcome in the same column but prefers all other outcomes in the same row

Therfore, all we need is the ordering and not the actual numbers!

Ordinal Utilities

Row player must be able to compare outcomes (indifference is ok) and that comparison must be transitive

For game to be zero-sum, column player must also be able to order outcomes and, moreover, column player's ordering must be reverse of row player's ordering

If only order matters (not magnitude) we have an *ordinal scale* and the numbers are said to represent *ordinal utilities*

These are sufficient for locating saddle points and dominated strategies

Slide 17

Cardinal Utilities

For mixed stratgies, however, we need to calculate ratios.

E.g., using Williams method:

		Kershaw			
		f	**x**	*Diffs*	*Probs*
	a	-2	4	-6	$\frac{3}{9}$
Goldsen	**i**	1	-2	3	$\frac{6}{9}$
	Diffs	-3	6		
	Probs	$\frac{6}{9}$	$\frac{3}{9}$		

If ratios are important we have an *interval* (or cardinal) *scale* and the numbers are said to be *cardinal utilities*

Slide 18

Ordinals to Cardinals

(von Neumann and Morgenstern)

Suppose an ordering

$$A < B < C$$

assign arbitrary numbers to A and C

$$A = 0,\ C = 100$$

Ask question: “is it preferable to have B for certain or a lottery giving 50% chance of A and 50% chance of C?

If answer is to prefer B, then B is greater than mid-point (50%)

If answer is to prefer lottery, then B is less than mid-point (50%)

Continue using “binary chop”

Slide 19

(This is for the row player; for the column player, negate these values for a zero-sum game.)

Cardinal utilities usually more difficult to ascertain than ordinal ones

Need cardinal utilities for meaningful mixed-strategy solution! Why?

Slide 20

Observation: end points of cardinal utilities are arbitrary

How does this affect the choice of numbers?

Different utility assignments represent the same cardinal preferences as long as there is a positive linear transformation from one to the other

Positive linear transformation: $f(x) = m.x + b$ with $m > 0$ (gradient)

Can transform cardinal utilities by such a function without “altering” the information they contain

Zero-Sum Equivalent Games

What does this all mean?

Games that are seemingly non-zero sum may be equivalent to zero-sum games

E.g., constant sum games

		Red 1	*Red* 2
Blue	**A**	(10,5)	(7,8)
	B	(12,3)	(1,14)

		Red 1	*Red* 2
Blue	**A**	(10,-10)	(7,-7)
	B	(12,-12)	(1,-1)

Subtract sum from one player's payoffs

Another Graphical Method

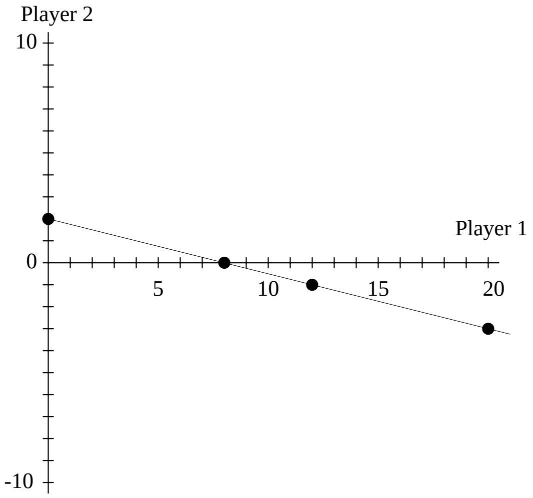
How do we tell whether a game is equivalent to a zero-sum game?

		Player 2 1	*Player 2* 2
Player 1	**A**	(0, 2)	(12, -1)
	B	(20, -3)	(8, 0)

Plot payoffs on a cartesian plane and if a straight line (with negative slope) can be drawn through the points, the game is zero-sum

Therefore, utilities are only defined for individuals and are arbitrary up to a positive linear scale

Slide 23



Transform Player 1 by $f(x) = \frac{1}{4}(x - 8)$

Slide 24

Games Against Nature

If one of the players in the game is incapable of "reasoning" and has no interest in outcome, minimax may not apply

Does solution still apply?

Depends on the goals of the reasoning player

Slide 25

Imagine playing a game against Nature.

How would you proceed?

Idea: if you know Nature's probabilities use the expected value principle

		Nature	
		1	**2**
Player 1	**A**	2	0
	B	1	1

Suppose Nature chooses **1**: $\frac{2}{3}$ **2**: $\frac{1}{3}$

Expected values **A**: $\frac{2}{3}.2 + \frac{1}{3}.0 = \frac{2}{3}$ **B**: $\frac{2}{3}.1 + \frac{1}{3}.1 = 1$

Slide 26

What if you're not sure of Nature's probabilities?

If you're confident in a range of probabilities can try varying them and calculating expected value again to see whether things change

What if you have no idea at all about how Nature behaves?

Various suggestions have been made; Milnor (1954) compared a number of these:

Laplace (1812) *principle of insufficient reason* — choose row with highest sum (or highest average payoff)

Assumes all strategies equally likely

Wald (1950) choose row with largest minimum outcome

Prepare for the worst. If game has a saddle point, this will choose strategy containing one

Hurwicz select "coefficient of optimisim" $\alpha \in [0,\ 1]$. Calculate for every row

$$\alpha(rowmax)\ +\ (1-\alpha)(rowmin)$$

Take row with highest value

Mixes "best" and "worst"

Savage minimise maximum regret

Regret matrix: for every outcome subtract value of maximum outcome in column; in each row determine the maximum regret and select row with minimal maximum regret

Axiomatic Method

These can all give different answers!

What would you do?

More mathematical approach: *axiomatic method*

Identify axioms (principles) that a good method for playing against Nature should adhere to

Evaluate which methods satisfy these axioms

Slide 29

Axioms

Milnor (1954) suggest the following.

1. *ordering:* all actions must be completely ordered
2. *symmetry:* ordering is independent of labelling of rows and columns
3. *strong domination:* A is preferred to B if A strongly dominates B
4. *linearity:* ordering unchanged by linear transformation (all multiplied by constant or constant added)
5. *continuity:* If $A > B$ in a sequence of decision problems under uncertainty, then $B \not> A$ in limiting decision problem under uncertainty

Slide 30

6. *column duplication:* adding an identical column does not change ordering
7. *column linearity:* ordering not changed by adding constant to column entries
8. *row adjunction:* ordering between old rows not changed by adding new rows
9. *convexity:* If A, B indifferent in ordering, then neither A nor B preferred to $\frac{1}{2}A$, $\frac{1}{2}B$
10. *special row adjunction:* adding a weakly dominated action does not change ordering of old actions

Slide 31

Comparison of Methods

Axiom	Laplace	Wald	Hurwicz	Savage
1. Ordering	*	*	*	*
2. Symmetry	*	*	*	*
3. Strong domination	*	*	*	*
4. Linearity	X	*	*	*
5. Continuity	X	X	*	X
6. Column duplication	*	*	*	–
7. Column linearity	*	–	–	*
8. Row adjunction	–	*	*	*
9. Convexity	X	*	–	*
10. Special row adjunction	X	X	X	*

Slide 32

(previous slide) * – characterise method; X – method satisfies this

No method satisfies all axioms

Milnor showed these axioms are incompatible!

That is, there can be no method that satisifies all of them

Slide 33

Equilibrium Pair of Strategies

Pair of strategies where player unilaterally deviating from their equilibrium strategy will worsen their expected payoff.

Minimax theorem implies equilibrium pair and minimax pair coincide *for zero-sum game.*

Nash's theorem extends this to non-cooperative games (be they zero-sum or nonzero-sum).

Slide 34

Nonzero-Sum Games

Each player has distinct payoffs which may not be reducible to a zero-sum game

They may no longer be in direct conflict

They may benefit from cooperation

Nash's Theorem: Any n-person, noncooperative game (zero-sum or nonzero-sum) for which each player has a finite number of pure strategies has at least one equilibrium set of strategies

Slide 35

Nash Equilibrium

Most important and widespread equilibrium concept

A strategy combination is a Nash equilibrium if no player has incentive to deviate from their strategy given the other player(s) does not deviate

Formally:

$$u_i(S_1^*, S_2^*, \ldots, S_i^*, \ldots, S_n^*) \geq u_i(S_1^*, S_2^*, \ldots, S_i, \ldots, S_n^*) \geq$$

for all players i where $u_i(S_1^*, \ldots, S_n^*)$ is the utility to player i if players adopt strategies $S_1^*, \ldots, S_n^*$.

Slide 36

There may be several Nash equilibria

They don't have to be fair

There may be no obvious way to choose one

Even if unique equilibrium exists it may be detrimental to both parties (cf. prisoner's dilemma)

Interpretation: In a mutually suspicious world, if such a state was "imposed", it will endure

Nash Equilibrium

Motivation: dominance implies a chain of reasoning: if Player 1 plays X, Player 2 plays Y, hence Player 1 should play X', then Player 2 should try $\ldots$, then Player 1 might $\ldots$

The hope is that the procedure "zeros-in" on a unique strategy for each player

Reason: looking for a stable state — equilibrium — where no player would want to deviate

Finding Pure Strategy Nash Equilibria

In two player games:
  – for each strategy of opponent, underline own best reply
  – a cell with both entries underlined represents a (pure-strategy) Nash Equilibrium

E.g., prisoner's dilemma

		Player 2	
		Loyal	**Fink**
Player 1	**Loyal**	(-1, -1)	(-3, $\underline{0}$)
	Fink	($\underline{0}$, -3)	($\underline{-2}$, $\underline{-2}$)

Fink Fink is a (pure-strategy) Nash Equilibrium

As in zero-sum games pure-strategy equilibria don't always exist

Therefore, look for mixed-strategy Nash Equilibria — next week

Slide 39

Pareto Optimality

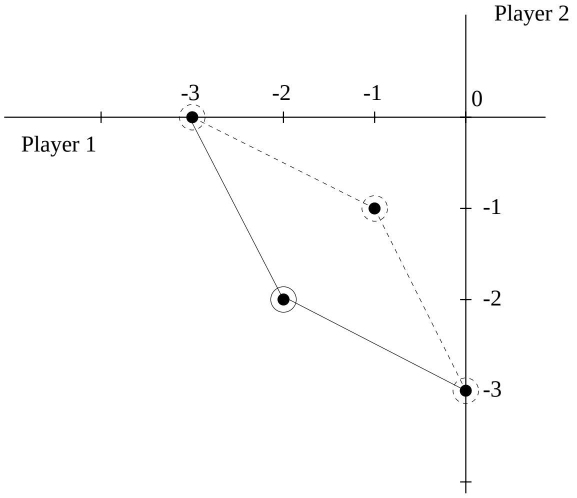
Definition: An outcome is *Pareto optimal* if there is no other outcome which would give both players a higher payoff or would give one player the same payoff and the other player a higher payoff.

		Player 2	
		Loyal	**Fink**
Player 1	**Loyal**	(-1, -1)*	(-3, $\underline{0}$)*
	Fink	($\underline{0}$, -3)*	($\underline{-2}$, $\underline{-2}$)

Pareto optimal points are on the north east boundary of the payoff polygon

Slide 40

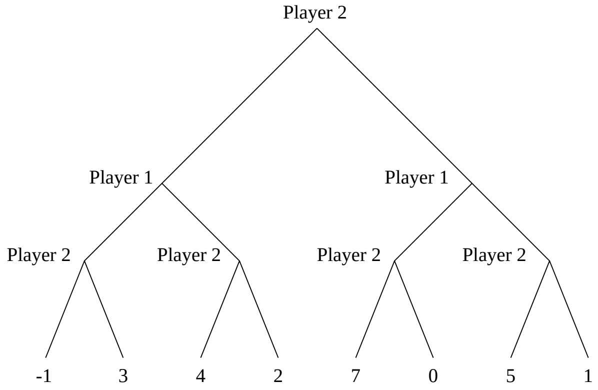
Payoff Polygon



Slide 41

Exercise 1

Find solution by truncation, write matrix for game and solve matrix checking it against answer to first part.



Slide 42

Exercise 2

(Straffin 1993, Ex9.4) Are the following equivalent to zero-sum games?

		Colin	
		1	**2**
Rose	**A**	(0,10)	(1,-10)
	B	(3,-50)	(-1,30)

		Colin	
		1	**2**
Rose	**A**	(-2,3)	(2,0)
	B	(3,-2)	(0,2)

Slide 43

Exercise 3

(Straffin 1993, Ex 10.1) Determine what each of Laplace, Wald, Hurwicz and Savage would recommend to this manager.

Economy

		Wy Up	**Sltly Up**	Sltly Dwn	**Wy Dwn**
	Hold Stdy	3	2	2	0
Mgr	**Exp Sltly**	4	2	0	0
	Exp Grtly	6	2	0	-2
	Div	1	1	2	2

Slide 44

Exercise 4

Find the pure-strategy Nash equilibria and Pareto optimal outcomes

Player 2

		A	**B**	**C**
Player 1	**1**	(0, 4)	(4, 0)	(5, 3)
	2	(4, 0)	(0, 4)	(5, 3)
	3	(3, 5)	(3, 5)	(6, 6)

Jane

		Opera	**Fight**
John	**Opera**	(2, 1)	(0, 0)
	Fight	(0, 0)	(1, 2)