Exercise Sheet - Game Theory

Exercise 1 (MiniMax Algorithm)**.** The Minimax algorithm assumes that the opponent plays optimally. Suppose the opponent *does not play optimally*. Show the following:

1. For non-optimally playing opponents the Minimax value is at least as good as the computed Minimax value if played against an optimal opponent.
2. For non-optimally playing opponents, the Minimax value may not be optimal.

Solution:

1. Consider a MIN node whose children are terminal nodes. If MIN plays suboptimally, then the value of the node is greater than or equal to the value it would have if MIN played optimally. Hence, the value of the MAX node that is the MIN node's parent can only be increased or remains equal. This argument can be extended by a simple induction argument all the way to the root.
2. If the suboptimal play by MIN is predictable, then one can do better than a minimax strategy. For example, if MIN always falls for a certain kind of trap and loses, then setting the trap guarantees a win even if there is actually a devastating response for MIN. This is shown in Figure **??**. Max should pick the right tree if min plays optimal, but if min plays suboptimal in the left-subtree then it should pick the left tree to earn a utility of 1000.

▲

Exercise 2 (Heuristic Minimax)**.** This exercise is about the board game Tic-Tac-Toe, played on a 3x3 grid.

(a) Given a game tree, we also refer to a branch in the game tree as a *game* (thus, a game captures a possible play of Tic-Tac-Toe). Give an upper bound on how many possible games of Tic-Tac-Toe there are?

(b) Draw the game tree starting from an empty board down to depth 2 (i.e., 2-plies). Take symmetry into account; that is, do not draw any branches in the game that are symmetric.

(c) Define an appropriate *evaluation function* for Tic-Tac-Toe. Give arguments why it is a good one and compute the utility of the states at depth 2 from part (b).

(d) Based on the utility of the evaluation function of (c) use the Heuristic Minimax algorithm to compute the Heuristic Minimax value of each state of the game tree. Also, give the corresponding optimal actions.

Solution: For a, there are at most 9! games. (This is the number of move sequences that fill up the board, but many wins and losses end before the board is full.)

For b,c,d Figure 1 shows the game tree, with the evaluation function values below the terminal nodes and the backed-up values to the right of the non-terminal nodes.

The evaluation function is the number of lines that can be completed by X minus the number of lines that can be completed by O. It should be easier to win or at least draw when you have a lot more ways to complete your lines, since it gives you more options.

The values imply that the best starting move for X is to take the center. The values to the right of the non-terminal nodes show the Heuristic Minimax values.

▲

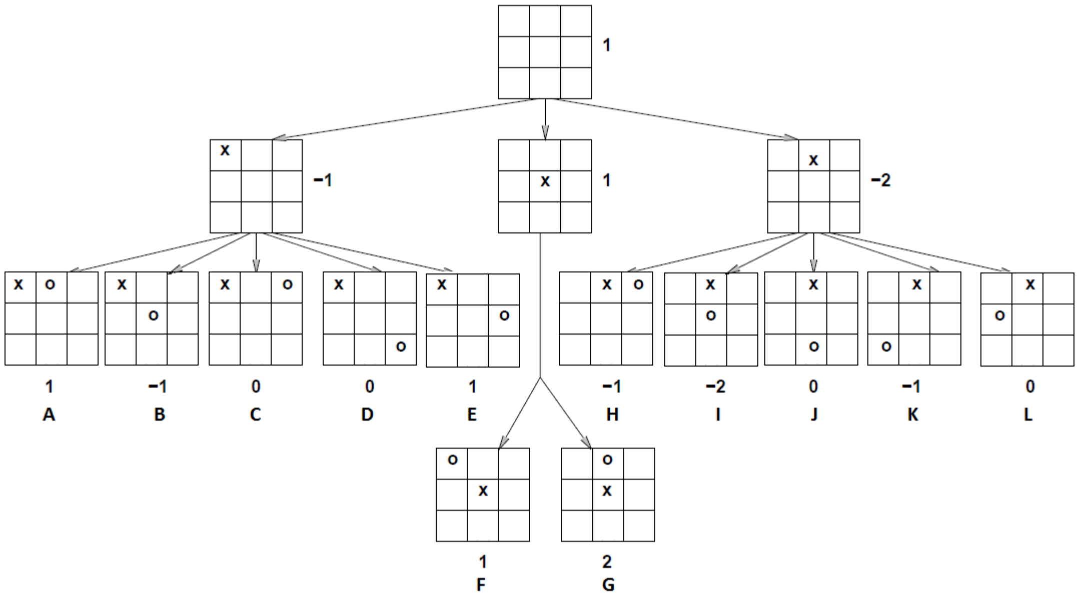


Figure 1: Solution for tic-tac-toe.

Exercise 3 (Zero sum game)**.** Consider the following game (Bombers and Fighters):

		Bomber Crew	
		Look Up	Look Down
Fighter Pilots	Hun-in-the-Sun	*0.95,0.05*	*1,0*
	Ezak-Imak	*1,0*	*0,1*

1. Compute the maximin strategies of both players using Von Neumann's Maximin technique.
2. Compute all (mixed) Nash equilibria deriving the answer from your solution to 1.
3. Compute the maximin value of the Fighter Pilots directly, using the alternative methods shown on slide 25 of the lecture.
4. Show that the value of the "Fighters and Bombers" game (i.e. the maximin value of the Fighter Pilots) is *strictly greater than* 0.95 (and thus that there is a mixed strategy that is better than the pure Hun-in-the-sun strategy).

Solution:

1. Suppose the Fighter player has the strategy $(Hun : p, Ezak : 1 - p)$. This is maximal if, regardless of the other player's strategy, Fighter has the same utility, i,e,

$$0.95p + 1(1 - p) = 1p + 0(1 - p)$$

$$0.95p + (1 - p) = p$$

This solves to: $p = \frac{1}{1.05} = \frac{20}{21}$ and so $(Hun : \frac{20}{21}, Ezak : \frac{1}{21})$ is the players maxmin strategy.
Similarly, we compute the maxmin strategy of the Bomber crew:

$$0.05p = 1 - p$$

Thus, $(Up : \frac{20}{21}, Down : \frac{1}{21})$ is the players maxmin strategy

2. We give the generic calculation. Suppose there is a NE where the Bomber's play $(Up : q, Down : 1 - q)$ with $0 < q < 1$. Then, if the Fighter Pilots play $(Hun : p, Eza; : 1 - p)$. Then, the Bomber's pure strategies Up and Down must give the same payoff for the Bombers, for otherwise it would be better to always play Up or Down, respectively.

$$0.05p \quad = \quad (1-p)1 \qquad \Rightarrow p = \frac{20}{21} \tag{1}$$

Similarly, for the Fighter Pilots, we get:

$$0.95p + (1-p) = p \qquad \Rightarrow p = \frac{20}{21} \tag{2}$$

Thus, the Nash equilibrium is: $((Hun : \frac{20}{21}, Ezak : \frac{1}{21}), (Up : \frac{20}{21}, Down : \frac{1}{21}))$.
This is exactly, the set of maxmin strategies (M_P, M_B). This is no surprise as both coincide for finite 2-player, constant sum games.

3. We want to find the strategy s_{fp} for the fighter pilots that maximizes the utility ut_{fp}, given that after choosing s_{fp}, the bomber crew will choose the strategy s_{bc} that minimizes the utility of the fighter pilots:

$$\begin{aligned}
s_{fp} &= \mathsf{argmax}_{s_{fp}} \mathsf{min}_{s_{bc}} \mathsf{ut}_{fp}((\text{Hun: } x, \text{Ezak: } 1-x), (\text{Up: } y, \text{Down: } (1-y)) \\
&= \mathsf{argmax}_{s_{fp}} \mathsf{min}_{s_{bc}} (\\
&\quad 0.95xy && \text{(Hun, Up)} \\
&\quad + x(1-y) && \text{(Hun, Down)} \\
&\quad + (1-x)y && \text{(Ezak, Up)} \\
&\quad + 0(1-x)(1-y) && \text{(Ezak, Down)} \\
) &= \mathsf{argmax}_{s_{fp}} \mathsf{min}_{s_{bc}} (-1.05xy + x + y)
\end{aligned}$$

The bomber crew tries to minimize the expression of ut_{fp}. The slope in y is given by the first derivative of ut_{fp} to y:

$$\frac{d\mathsf{ut}_{fp}}{dy} = -1.05x + 1$$

The figher pilots are best off when ut_{fp} does not slope at all, i.e. $-1.05x + 1 = 0$, which implies $x = \frac{20}{21}$.

4. The outcome of the Fighter pilots' maxmin strategy is

$$ut_P = 0.95\frac{20}{21} + \frac{1}{21} = \frac{20}{21} = 0,952380... > 0.95$$

▲

Exercise 4 (Pirates and Gold)**.** We are given the following puzzle. Five rational *pirates* a, b, c, d, e negotiate about how to share 100 *gold coins*. The pirates are ranked from a *(highest)* to e *(lowest)* The task is to distribute the coins such that:

1. the highest ranked pirate proposes a distribution.
2. each pirate can *accept* or *reject* the proposal, *majority* decides (highest ranked pirate *breaks ties*, otherwise the pirate takes part in the negotiation as any other pirate).
3. if the proposal is rejected the proposing pirate is *killed* and the next highest ranked pirate makes a proposal, and so forth.

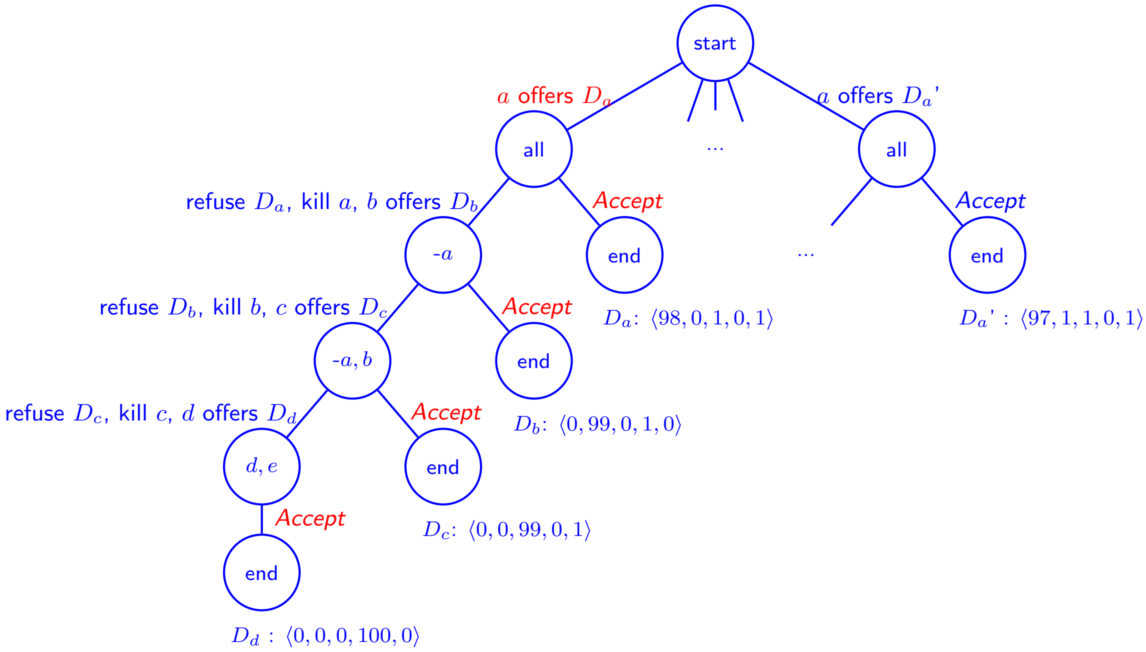
The *preferences* of the pirates (in this order and additive) are as follows: (i) stay alive, (ii) maximize the number of gold coins the pirate gets, and (iii) kill other pirates.

1. Model this as an extensive form game (informally) and explain how strategies are defined.
2. Compute the unique subgame perfect Nash equilibrium.

Solution:

In the nodes, the players who are voting.

In red, the best choice.



Backwards induction:

- at (d, e), e has no more bargaining power and the offer D_d cannot be refused. Consequently d will maximise his payoff and give nothing to e.
- At $-a, b$, to obtain a majority, all c has to do is offer e some nonzero sum. In fact, if c's offer is refused, we reach (d, e) and e gets nothing. So e will vote for c's offer and the game ends with D_c.
- At $-a$, d knows that if the game reaches $-a, b$, e and c will outvote him and he will get nothing. So he will accept any nonzero offer. So he will accept D_b. Since higher rank pirates break ties, D_b is accepted and the game ends.
- At all, everyone knows that if the game reaches $-a$, the game ends immediately with outcome D_b. So, to outvote the opponents, a simply has to offer some nonzero sum to those who would get 0 if the game reaches $-a$. So D_a is accepted and the game ends immediately.

▲