1 STRATEGIES OF PLAY

The Minimax algorithm is the most well-known strategy of play of two-player, zero-sum games. The minimax theorem was proven by John von Neumann in 1928. Minimax is a strategy of always minimizing the maximum possible loss which can result from a choice that a player makes. Before we examine minimax, though, let's look at some of the other possible algorithms.

1.1 Maximax

Maximax principle counsels the player to choose the strategy that yields the best of the best possible outcomes. For example, let's consider a zero-sum game where two players simultaneously put either a blue or a red card on the table. If player 1 puts a red card down on the table, whichever card player 2 puts down, no one wins anything. If player 1 puts a blue card on the table and player 2 puts a red card, then player 2 wins $1,000 from player 1. Finally, if player 1 puts a blue card on the table and player 2 puts a blue card down, then player 1 wins $1,000 from player 2.

The payoff matrix for player 1 is shown in this table:

Player 2 Blue
Red
Player 1
Blue
1,000
-1,000
Red
0
0

Going by maximax principle, player 1 will always play the blue card, since it has the maximum possible payoff of 1,000. But as can be clearly seen from the table, assuming player 2 is rational, he will never play the blue card, since the red card gives him the dominant strategy. In such a case, if player 1 plays by the maximax rule, he will always lose.

The maximax principle is inherently irrational, as it does not take into account the other player's possible choices. Maximax is often adopted by naive decision-makers such as young children.

1.2 Maximin

Maximin is solely a one-person game strategy, i.e. a principle which may be used when a person's "competition" is nature, or chance. Whereas the maximax principle is ultra-optimistic, expecting the best possible payoff, the maximin is ultra-pessimistic, expecting the worst possible payoff. It involves choosing the best of the worst possible outcomes.

A simple example of a slot machine game may be used. A player has only two choices to make – to gamble or not to gamble. If he gambles, he risks losing his bet (say, 1), *but also has a chance to win the maxi* If he does not gamble, he can neither win nor lose.

The payoff matrix looks like this:

Chance Win
Lose
Player

Gamble

10000

-1

Not gamble

0

0

According to the maximin principle, the player should never gamble, because he faces a risk of losing \$1. It is clear that the maximin principle is quite inefficient, since it discourages taking any risks, no matter how high the reward may be.

1.3   Minimax for One-Person Games

The Minimax Regret Principle is based on the Minimax Theorem advanced by John von Neumann, but is geared only towards one-person games. It relies on the concept of regret matrices. To demonstrate, consider an example of a company trying to decide whether or not it should support a research project. Let's assume that the research project costs c units. If it succeeds, it will bring in a return of r units. If it fails, it will obviously not bring in anything.

The payoff matrix for the company looks like this:

Research Succeeds

Fails

Company

Supports research

r - c

-c

Neglect research

0

0

By the maximax principle, a company should always support research if the expected return on it is greater than its cost. By maximin, the company should never support research, since it is risking the cost of the research. Minimax is slightly more complicated.

We need to come up with a matrix that shows the "opportunity cost," or regret, of the player, depending on each possible decision. For example, if the company supports the research and it fails, the company's regret will be c, the price of research. If the company supports the research and it succeeds, the company will have no regrets. If the company neglects research and it would have succeeded, its regret value is r-c, the return on the research. So, the minimax regret matrix will look like this:

Research Succeeds

Fails

Company

Supports research

0

c

Neglect research

r-c

0

The object is to minimize the maximum possible regret. It is not obvious from the above matrix what the maximum value is. That is, is c greater than r-c? If (r-c) ¿ c, the company

should support research. If (r-c) ¡ c, the company should not. In other words, the company should support research if c ¡ r/2, or, if the expected return on research is more than twice its cost.

2 Minimax for Two-Person Games

In a two-person, zero-sum game, a person can win only if the other player loses. No cooperation is possible. Andrew Colman's Game Theory and Experimental Games shows the following historical example:

In 1943, the Allied forces received reports that a Japanese convoy would be heading by sea to reinforce their troops. The convoy could take on of two routes – the Northern or the Southern route. The Allies had to decide where to disperse their reconnaissance aircraft – in the north or the south – in order to spot the convoy as early as possible. The following payoff matrix shows the possible decisions made by the Japanese and the Allies, with the outcomes expressed in the number of days of bombing the Allies could achieve with each possibility:

Japanese Route North
South
Allies Reconnaissance
North
2
2
South
1
3

By this matrix, if the Japanese chose to take the southern route while the Allies decided to focus their recon planes in the north, the convoy would be bombed for 2 days. The best outcome for the Allies would be if they placed their airplanes in the south and the Japanese took the southern route. The best outcome for the Japanese would be to take the northern route, provided the Allies were looking for them in the south.

To minimize the worst possible outcome, the Allies would have to choose the north as the focus of their reconnaisance efforts. This ensures them 2 days of bombing, whereas they risk only 1 day of bombing if they focus on the south. Therefore, by minimax, the best strategy for them would be to focus on the north.

The Japanese can use the same strategy. The worst possible outcome for them is the 3 days of bombing which might occur if they took the southern route. Therefore, the Japanese would take the northern route.

It is, in fact, what had occurred: both the Allies and the Japanese chose the north, and the Japanese convoy was bombed for 2 days.

The previous matrix had a saddle point, meaning that both the Japanese and the Allies settled on the (North, North) square as the best outcome for both of them. Neither could do any better if the opponent was rational. In this case, the maximin and the minimax strategies produce the same result. Notice that if the Allies were following maximax, they would choose the South, and surely forfeit one day of bombing.

2.1   Mixed Strategies and Randomization

In some cases, there is no saddle point, and the players have to choose their strategies with a degree of randomness, as in the following simple game, called "***Matching Pennies***". Two players simultaneously place a penny on a table, either heads up or tails up. If the pennies are facing the same way, player 1 gets to keep both pennies. Otherwise, player 2 gets to keep both. The payoff matrix for player 1 looks like this:

Player 2 Heads
Tails
Player 1
Heads
1
-1
Tails
-1
1

- There is no clearly defined strategy for each player. The best way to play is to choose the position of the coin randomly. If either player follows this strategy, then in the long run, the payoffs for each will be 0.

- Notice that if, say, player 1 uses a 50/50 strategy, while player 2 plays heads 75% of the time, in the long run, both players will still have payoffs of 0.

- But if player 2 follows the 75/25 strategy, then player 1 can easily take advantage of it by playing heads more frequently, and therefore winning more frequently. So, it is important for each player to not only maintain a random strategy, but to also analyze the strategy of the other player.

2.2   Applications to Computing

- In computer simulations for cases such as this, it is important not to program the computer with a specific strategy in advance, but let it decide it at run-time. If the computer does not maintain unpredictability, then the opposing player may use this knowledge to his advantage. Many computer games suffer because although the computer is programmed with a strong strategy, it becomes predictable and easy to take advantage of.

- On the other hand, the computer might very well benefit if it recognizes a predictable strategy on the part of an opponent. Even in such a simple game as "Matching Pennies," where a 50/50 is called for, while the computer may follow a 50% algorithm for deciding whether to play heads or tails, a human cannot come up with completely random numbers. In fact, it has been observed that humans tend to play heads slightly more often. If a computer recognizes that the probability of its opponent of picking heads is slightly higher, it may adjust its own strategy to have an advantage.