Daniel Almeraz

Oscar Galindo

Final Project Part 2: Sophisticated Agents

This section of the final project brings two new aspects onto the table. The first is that we now need to create agents with uncertainty, and the second is being able to take advantage of playing against the same opponent multiple times. Following are agents that we designed for use with these specific types of scenarios.

Agents:

One-shot games with no uncertainty

MaxAverage

When it came down to playing one-shot games with uncertainty we decided to follow a slightly greedy approach. The way that our agent works is that it would look at all their possible moves and pick the one that would give them the highest average payoff. For example, if we were playing as the column player and the column's averages was 5, 6, 3 we would play the middle column since we would likely get the highest payoff from it. This agent was designed to be care about itself getting the highest payoff after all test are done.

MaxPayoff

As an even greedier approach than max average, max payoff cares only about trying to get the highest possible result. This hurts the agent since it disregards the all the possible bad moves but it has potential good results when dealing with small boards.

One-shot games with uncertainty -

VariancePayoff

In one-shot with uncertainty our approach was to use variance. The main reason we decided to go with variance when it came down to playing one-shot games with uncertainty was to create an agent that could "play it safe" and be predictable. The way it works is by finding the variance in each option and picking the one that had the lowest variance using:

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

where

μ = mean

N = Number of scores.

©easycalculation.com

This agent was designed to be able to tolerate high chances of uncertainty since it would not take rash decisions in what move to pick. In addition, this agent implemented a comparison of the averages of payoff for all columns/rows. This last quality allowed the agent to discern between options that represented just a small variance and those that represented a small variance as well as a good potential payoff. In the case of negative average payoffs, besides the variance analysis our agent chooses the most positive/least negative of all the options.

Repeated games with uncertainty

Combined

Inspired by the working of neural networks, our approach to solving a game that could allow for playing against the same opponent multiple time was actually to combine several different strategies and then "learn" which is the best one.

Version 1 - CombinedSoftPayoff

The three agents we decided to combine was the MaxAveragePayoff as described in the first section, The variance agent described in the second section, and a third agent called MaxPayoff that was as greedy as possible and picked the move that could potentially give it the highest payoff. The reason for these three agents is that we expect our agent to learn what playing style is best between one that tries to guarantee a safe predictable move(variance) against an aggressive agent, one that tries to reach the highest likely solution (max average) against an unpredictable agent, and one that tried to reach the highest possible value(maxpayoff) and unpredictable agent and in a small game. The way the agent learns is by using a softmax activation. In addition, the agent keeps track of how many games each possibility has won, which also allows the agent to be expandable since a softmax activation of an input vector is always 1. Additionally, thanks to the addition of probability to the agent, it means our agent can become unpredictable and thus prevent other agents from learning our (already several) strategies.

Version 2 - CombinedPayoff

Since we were able to trade out internal agents fairly easily within this agent, we decided to create an agent that used similar concept to that of version but with more specialized versions such as the ones learned in class. We combined maxmin payoff, minimax regret, and our own max average payoff to cover more general bases. Minimax regret would try to cover an

aggressive opponent, maxmin payoff would try to outsmart an opponent that was trying to make us pick the worse move, and max average payoff would attempt to deal with an unpredictable agent.

Sample Tournament Results:

All results are shown using 20 actions per game type.

Combined-SoftPayoff vs. CombinedPayoff

Zero Sum - no uncertainty one shot

Total Wins
Combined-Payoff 27.0
CombinedSoft-Payoff 33.0

Overall Average Expected Utility
Combined-Payoff -0.14383750568116155
CombinedSoft-Payoff 0.14383750568116155

Tournament Stabilities
Combined-Payoff 0.2876750113623231
CombinedSoft-Payoff -0.2876750113623231

General Sum - no uncertainty one shot

Total Wins
Combined-Payoff 29.0
CombinedSoft-Payoff 31.0

Overall Average Expected Utility
Combined-Payoff 4.126677818977945
CombinedSoft-Payoff 3.501296269932906

Tournament Stabilities
Combined-Payoff -2.9424281853995504
CombinedSoft-Payoff -1.6916650873094725

Risk v Reward - no uncertainty one shot

Total Wins
Combined-Payoff 29.0
CombinedSoft-Payoff 31.0

Overall Average Expected Utility
Combined-Payoff 33.91804137306518
CombinedSoft-Payoff 30.047295689851754

Tournament Stabilities
Combined-Payoff -25.10017239062376
CombinedSoft-Payoff -17.358681024196905

Risk v Reward - few outcomes changed with large intervals one shot

Total Wins
Combined-Payoff 33.0
CombinedSoft-Payoff 27.0

Overall Average Expected Utility
Combined-Payoff 34.94602195733099
CombinedSoft-Payoff 37.83833870208506

Tournament Stabilities
Combined-Payoff -23.365616522801066
CombinedSoft-Payoff -29.150250012309197

General Sum - few outcomes changed with large intervals one shot

Total Wins
Combined-Payoff 36.0
CombinedSoft-Payoff 24.0

Overall Average Expected Utility
Combined-Payoff 4.1716213962390505
CombinedSoft-Payoff 3.8162581805372096

Tournament Stabilities
Combined-Payoff -2.9709962875066234
CombinedSoft-Payoff -2.2602698561029406

Risk v Reward - many outcomes changed with large intervals one shot

Total Wins
Combined-Payoff 32.0
CombinedSoft-Payoff 28.0

Overall Average Expected Utility
Combined-Payoff 33.880575428777156
CombinedSoft-Payoff 26.25790326895342

Tournament Stabilities
Combined-Payoff -25.627812613112976
CombinedSoft-Payoff -10.3824682934655

Variance vs. UniformRandom

Zero Sum - no uncertainty one shot

Total Wins
UniformRandom 20.0
Variance 40.0

Overall Average Expected Utility
UniformRandom -0.32128580860041617
Variance 0.32128580860041617

Tournament Stabilities
UniformRandom 0.6425716172008323
Variance -0.6425716172008323

General Sum - no uncertainty one shot

Total Wins
UniformRandom 37.0
Variance 23.0

Overall Average Expected Utility
UniformRandom 3.7797394808932845
Variance 3.2097507473813467

Tournament Stabilities
UniformRandom -2.916700218429509
Variance -1.776722751405634

Risk v Reward - no uncertainty one shot

Total Wins
UniformRandom 20.0
Variance 40.0

Overall Average Expected Utility
UniformRandom 23.450701485784453
Variance 37.9727126425014

Tournament Stabilities
UniformRandom -6.236767242483982
Variance -35.28078955591787

Risk v Reward - few outcomes changed with large intervals one shot

Total Wins
UniformRandom 20.0
Variance 40.0

Overall Average Expected Utility
UniformRandom 23.450701485784453
Variance 37.9727126425014

Tournament Stabilities
UniformRandom -6.236767242483982
Variance -35.28078955591787

General Sum - few outcomes changed with large intervals one shot

Total Wins
UniformRandom 38.0
Variance 22.0

Overall Average Expected Utility
UniformRandom 3.796141025375803
Variance 3.2649161477195983

Tournament Stabilities
UniformRandom -2.900218955971705
Variance -1.837769200659296

Risk v Reward - many outcomes changed with large intervals one shot

Total Wins
UniformRandom 20.0
Variance 40.0

Overall Average Expected Utility
UniformRandom 23.676633639652746
Variance 36.96526634128614

Tournament Stabilities
UniformRandom -6.723416651664635
Variance -33.30068205493143

MaxAverage vs. UniformRandom

Zero Sum - no uncertainty one shot

Total Wins
UniformRandom 20.0
MaxAverage 40.0

Overall Average Expected Utility
UniformRandom -0.6319781538676932
MaxAverage 0.6319781538676932

Tournament Stabilities
UniformRandom 1.2639563077353864
MaxAverage -1.2639563077353864

General Sum - no uncertainty one shot

Total Wins
UniformRandom 21.0
MaxAverage 39.0

Overall Average Expected Utility
UniformRandom 3.7478131791521534
MaxAverage 4.369925428977844

Tournament Stabilities
UniformRandom -1.8689750510916525
MaxAverage -3.113199550743034

Risk v Reward - no uncertainty one shot

Total Wins
UniformRandom 20.0
MaxAverage 40.0

Overall Average Expected Utility
UniformRandom 23.450701485784453
MaxAverage 37.9727126425014

Tournament Stabilities
UniformRandom -6.236767242483982
MaxAverage -35.28078955591787

Risk v Reward - few outcomes changed with large intervals one shot

Total Wins
UniformRandom 20.0
MaxAverage 40.0

Overall Average Expected Utility
UniformRandom 23.450701485784453
MaxAverage 37.9727126425014

Tournament Stabilities
UniformRandom -6.236767242483982
MaxAverage -35.28078955591787

General Sum - few outcomes changed with large intervals one shot

Total Wins
UniformRandom 21.0
MaxAverage 39.0

Overall Average Expected Utility
UniformRandom 3.7478131791521534
MaxAverage 4.369925428977844

Tournament Stabilities
UniformRandom -1.8689750510916525
MaxAverage -3.113199550743034

Risk v Reward - many outcomes changed with large intervals one shot

Total Wins
UniformRandom 20.0
MaxAverage 40.0

Overall Average Expected Utility
UniformRandom 23.493048653039153
MaxAverage 37.701584826052084

Tournament Stabilities
UniformRandom -6.318722807129916
MaxAverage -34.735795153155784