Creating a New Angry Birds Competition Track

Rohan Verma, Xiaoyu Ge, Jochen Renz

Research School of Computer Science
The Australian National University, Canberra ACT 0200, Australia
jochen.renz@anu.edu.au

Abstract

This paper introduces the new competitive track of the Angry Birds Artificial Intelligence Competition that was most recently hosted at IJCAI in August 2015. The goal of the competition is to inspire the creation of AI that can predict the effects of physical actions in the real world. Agents in the competitive track will have to focus on the AI techniques that will be useful in facing this challenge. The game the agents play is the popular Angry Birds created by Rovio. First, we discuss how we designed the competitive track and modelled it as an extensive form game. We show the pure strategy Nash Equilibrium for a single level of the competitive track. We then show that a single strategy is not dominant by defining simple cooperative strategies that outperform the optimal agent in the competition.

Introduction

Angry Birds is a popular physics simulation game where the player flings birds on different trajectories at structures to destroy them and the pigs they protect. The Angry Birds Artificial Intelligence competition was held in 2012 and many students signed up to create agents who could play Angry Birds and compete to be the best. Since then it has been held each year with international teams competing and writing papers on their submissions (Renz 2015). This new competition involves two agents playing same level with alternating turns with a bid to determine who goes first. The goal of this competition is to create Artificial Intelligence agents that can predict the consequences of physical actions in the real world. Angry Birds represents a simplified version of this problem creating an excellent environment within which to test AI agents. Over the past three Angry Birds Artificial Intelligence Competitions, the agents have developed from solving basic problems such as computer vision and trajectory planning to the stage where they are able to play as well or better than many human players. However, simple techniques have proved

to be very effective at playing traditional Angry Birds. Heuristics that cannot be generalized to other problems are not useful for furthering Artificial Intelligence research in this area. The goal of the new competitive track is to provide a testing ground for agents that can analyze and know the consequences of their actions.

When creating a computer competition where agents attempt to outsmart their opponents, it is useful to draw comparisons with other famous AI competitions. The one we focus on is the Iterated Prisoner's Dilemma competition first run by Robert Axelrod in 1980 (Axelrod 1997). In this competition, agents play a simple strategic form game in which the optimal way to play is already known. The reason it has inspired so much research is that the agents win not by defeating each opponent but achieving the highest score overall. This allows for many different strategies to be created emphasizing cooperation and the risks involved (Wedekind and Milinski 1996).

In this work, we discuss how we created the new competitive track. This track was a part of the Angry Birds Artificial Intelligence Competition 2015 hosted at IJCAI in August. Its rule set was first designed to emphasize the goals of the competition. We then evaluated the design by modelling it as an extensive form game and drawing comparisons to iterated Prisoner's Dilemma. This paper shows the depth of the competition by demonstrating how simple strategies match up against each other. We discuss the creation of the simple agent to compete in the competition. This involves adapting the naive agent from the main competition to the new rule set. We created a simple heuristic for estimating the value of a level and provide this agent's source code to teams wishing to compete.

Background

Angry Birds

Angry Birds is a video game franchise originally launched as an iOS game by the developer Rovio Entertainment in

2009. It has since become analogous with casual gaming. This paper discusses the Chrome version of the original game. In this version of the game, a series of puzzles or levels are presented to the player. Each level has a determined minimum number of moves to reach a solution. The player is generally allowed quite a few more moves than necessary but rewarded for using less. The objective of each puzzle is to destroy all of the target pigs. The players' moves consist of flinging birds at the pigs or the structure protecting them. The levels are two dimensional representations of physics from a side on view. Pigs are usually protected by structures which can be destroyed or knocked over with precision by players to achieve high scores.

Game Theory

General Sum Two Player Extensive Form Game

The extensive form of a game is a mathematical model used in game theory. It is an extensive version of strategic form games which can be represented as a matrix of payoffs given the strategies of the two players. The extensive form attempts to capture more about games that have sequential positions and moves. This allows for the ideas of bluffing, signaling and sandbagging to be examined. It does this by introducing three new concepts, the game tree, chance moves and information sets (Ferguson 2014).

The game tree is a directed graph of vertices and edges where each leaf node has payoffs associated with it for each player. Each vertex of the game tree represents a game state, which is usually a player's turn, and may have one or more edges which are the moves the player can make from this state. Edges may also have probabilities associated with them that are called chance moves. These usually represent actions such as rolling a dice or drawing a card from a shuffled deck. Finally information sets allow us to group game states together that players cannot differentiate between. For example, in poker after the cards are dealt, the player cannot know what cards are in their opponent's hand and vice versa. Extensive form games can also be used to model how much information a player can remember from earlier in the game. We assume in this paper that players can recall all previously observed outcomes.

We are concerned with a specific type of extensive form game, the two player general sum extensive form game. This type of game has only two players and the payoff at a terminal vertex cannot be written as a single number. Instead it is represented as a pair of real numbers representing the amount won by Player I and the amount won by Player II.

Prisoners' Dilemma

The Prisoners' Dilemma is famous in game theory. The game models a situation in which two prisoners have a plan of escape. If they both cooperate, escape is possible and they are both rewarded. However, if one cooperates while the other defects, the defector receives the highest payoff and the other the lowest. If neither cooperates they both receive a low payoff. This is represented by the Figure 1.

		Player 1	
		Cooperate	Defect
Player 2	Cooperate	3 , 3	5 , 0
	Defect	0 , 5	1 , 1

Figure 1: Payoff Matrix for Prisoner's Dilemma.

This creates the dilemma such that both want to escape, but they cannot gain anything by cooperating if the other defects. In fact, if they only meet once it is always best to defect. This has inspired a lot of research toward cooperative solutions to Prisoners' Dilemma (Golbeck 2002; Wedekind and Milinski 1996). Cooperation is possible in Prisoners' Dilemma when the same two players must play repeatedly. In this case, the overall payoff can be much higher and agents may attempt to only defect if their opponent defects first. This has led to many of the strategies that were successful in the Iterated Prisoner's Dilemma tournament where the winner is determined by the highest total score (Axelrod 1997). One of the most successful simple strategies was tit-for-tat in which it will defect once if its opponent defects first and then try cooperating again. This strategy is able to cooperate with other cooperative agents but also adequately defend itself against other strategies.

Prisoners' Dilemma is traditionally played simultaneously where both players' strategies are revealed at the same time. It can also be played sequentially where players take turns revealing first. This gives the opportunity for the second player to react to the first defecting immediately. Sequential repeated Prisoners' Dilemma is the game we can be most closely related to in the design of the Angry Birds competitive track.

Nash Equilibrium

The Nash Equilibrium is a fundamental part of game theory and is often also referred to as best response (Osborne, Martin and Rubinstein 1994). In strategic form games, the Nash Equilibrium is defined as an action profile of each player where no player can obtain a higher payoff deviating from this profile while the others remain the same (Osborne, Martin and Rubinstein 1994). A pure strategy of a strategic form game, such as Prisoner's dilemma, is where each player chooses one of the strategies that map to rows or columns.

Example. The pure strategies in Prisoner's Dilemma shown in Figure 1 are to either cooperate or defect. As we can see, if both players are cooperating either can change their strategy to defect to gain a higher payoff. If

one is defecting and the other cooperating, the cooperator can change to defect and gain a higher payoff. However, if both are defecting neither can gain by deviating from this strategy and it is therefore a Nash Equilibrium. These games consider the best strategy in a single iteration of the game. The Iterated Prisoner's Dilemma competition changed the goal to having the highest overall score after many iterations of the game against different opponents. This is why cooperation became more important than the best response and the winner was defined as the one who best reacted to the actions of the group (Golbeck 2002).

Competitive Track Design

In this section we introduce the new track of the Angry Birds Artificial Intelligence Competition. In the competitive track, the agents only get one chance to complete a level with their opponent in each match. This makes it important to be able to complete a level from many unknown states created by the other agent's shots. In contrast to the main competition, each agent has a time limit to play a series of levels and achieve the highest overall score. The agent can play each level any number of times in order to improve its score. The competitive tracks goal is to further reward careful analysis of the levels and the possibilities from each state. We discuss the design of the competitive track and why it is important for the goals of this research. The requirement of this competition is that strategic actions made through analysis are the key to victory. We then show the design meets this requirement through the use of game theoretical principles.

Description

The competitive track describes the game where agents take alternating shots after first bidding for order of play. The agent who takes the winning shot wins all of the points they scored in the level but may need to pay their opponent the bid. The bidding stage allows an agent to attempt to play first or second by submitting positive or negative (or zero) numbers of points which remains hidden from the opponent. The agent with the higher bid goes first and the lower bid goes second. The agent's intention is defined by whether they submitted a positive bid or a negative bid. If they submitted a positive bid they are bidding to go first otherwise they are bidding to go second. If an agent bid successfully achieves its intention, then if they go on to win, they must pay their bid to their opponent and keep what points remain. This is shown in the examples below:

Example 1. Agent A bids +12000 points on level 1 and Agent B bids +10000 points. Agent A makes the first shot, Agent B the second shot, Agent A the third shot and so on, until a level is won or lost. If Agent A scores the winning shot and the score is 28000 points, then Agent A gets 28000 points, but has to give 12000 points to Agent B. If Agent B wins, it can keep all points.

Example 2. If Agent C bids -10000 points on level 1 and Agent D bids +18000 points, then Agent D makes the first shot and Agent C the second shot. If Agent C scores the winning shot worth 27000 points, then Agent C gets 27000 points, but has to pay 10000 points to Agent D, since Agent C got to make the second shot it was bidding for. If Agent D scores the winning shot, it gets 27000 points, but has to pay 18000 points to Agent C, since Agent D got to make the first shot it was bidding for.

At the bidding stage, it is important each agent analyses the potential returns of completing a level and assesses the number of shots, even or odd number required. This allows them to determine whether the winning shot will be from first or second shooter. The competition proceeds with each agent playing every other agent over three levels and the highest accrued points wins the contest. For this competition all contestants have access to the sample agent. New server software was created to host the competition. A new display interface was also created to be shown during the live competition. This display makes it clear which agent currently has control over the game. It has a special interface for bidding by animating counting up and then highlighting the winner. All of this helps to make the competition more exciting and hopefully will lead to better technology for analyzing consequences of physical actions.

Strategies

This design creates a similar environment to repeated Prisoner's Dilemma. Following the bidding stage, the loser of the bid then has the chance to continue cooperating and receive the bid of the other or to attempt to alter the amount of shots to complete the level, in a way defecting, and taking all the remaining points for themselves. This is similar to non-simultaneous Prisoner's Dilemma. An agent loses the bid, and then can choose to defect or cooperate given that the first agent has chosen to do one or the other. This creates a simple strategic form game of the competitive track on a level worth 60,000 points shown on the next page.

Angry Birds

		Player 1	
		Cooperate	Defect
Player 2	Cooperate	30000 , 30000	50000 , 0
	Defect	0 , 50000	0 , 0

Figure 2: Payoff matrix of Prisoner's Dilemma and a puzzle in the Angry Birds Competitive Track.

The Figure 2 shows a model that is similar to Prisoner's Dilemma and suggests that an agent always wins or ties if they always defect never and let the other take a winning shot. To avoid this, wasting a shot in Angry Birds loses 10,000 points overall and thus decreases the remaining score when the game is repeated. In a round robin, an agent who always beats their opponent, but not by many points, will likely not come out on top in total score.

We then evaluated the well-studied and simple strategies from computer competitions of Prisoner's Dilemma to justify some of the design choices for this competition. We have related them back to the goal of this new competitive track.

• Cooperative agents always try to achieve the highest total payoff and bid half that amount to attempt to obtain the winning shot.

• Defector agents always bid zero and try to obtain the winning shot by wasting their first turn.

• Optimal agents use the Nash equilibrium of the described game to defend against exploitative agents.

• Null agents always bid zero and only shoot when they can make the winning shot.

To create an interesting competitive environment, it was important to not create a single optimal strategy. We modelled the described competitive track as an Extensive Form General Sum Game as in the next section.

Evaluation

The hypothetical level modelled in Figure 3 contains three birds and requires at least two shots to complete. The first nodes represent the choices of bids which will eventually determine the payoffs. We can consider there to be an equivalent tree under each bid pair selection. The bids are not revealed other than being higher or lower and therefore each agent cannot tell the difference between any of the matching states above or below his bid. The tree beneath represents an abstraction of the choices the agents could make while playing Angry Birds. Each node represents an agent's turn and they can either waste a bird which means they do not reduce the shots required to complete the level but do decrease the number of birds remaining. Alternatively, they can play an optimal shot which reduces the number of shots required to complete the level as well as decreasing the number of birds remaining. Thus, the maximum depth of the tree is the number of birds in a level and the minimum depth is the number of shots required to complete the level.
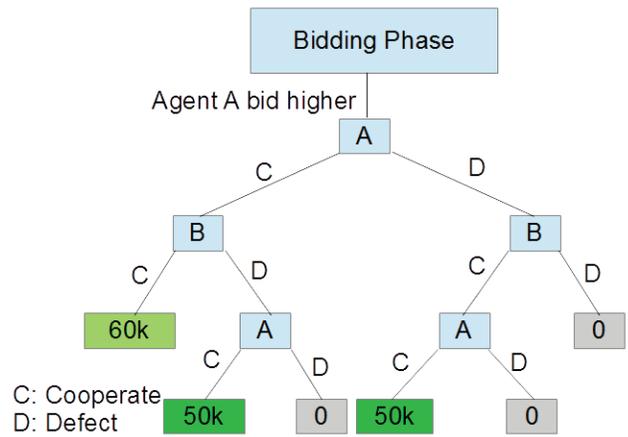

Figure 3: Extensive form of a hypothetical game between two agents in the Angry Birds Competitive Track with k denoting thousands of points.

A further simplification of the model is made such that every combination of shots leads to either a win with a set number of points only affected by the number of birds or a loss worth zero points. These payoffs are the terminal nodes containing 60,000, 50,000 and zero in Figure 3. Similarly to the actual game the reward is 50,000 for completing the level and an additional 10,000 points for each remaining bird.

Definition. *Pure Strategy in Extensive Form Games: a pure strategy in an extensive form game is a set of instructions describing which edge to follow at each node in the game tree.*

Observation 1. *The pure strategy Nash equilibrium of this game is to bid 25,000 points and always cooperate unless shooting first in which case defect.*

We use the optimal subgame principle to compare the opposing agent's pure strategies to the theorized equilibrium to show that they cannot gain by deviation. In our case, a pure strategy is a bid amount and whether to cooperate or defect at each node in the following two trees which represent going first or second after bidding higher or lower than the opponent.
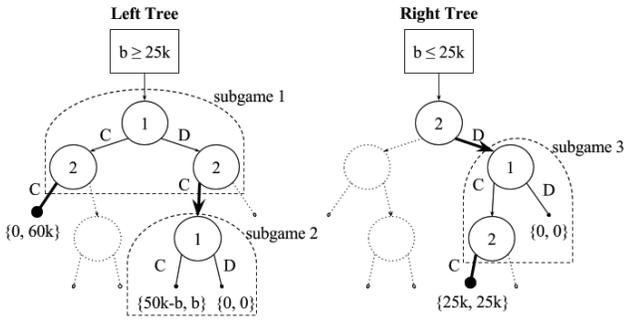
Figure 4: A representation of all possible pure strategies played against the Nash Equilibrium.

These trees show the three reachable subgames for the player, Player 1, who will attempt to deviate from the theorized equilibrium strategy. b is the bid made by Player 1, Player 2's bid is always 25,000. The maximum payoff achievable in each of these trees can be solved using backwards induction. This is shown by the formulas below.

max(left tree)
$= max_{subgame1}(0, max_{subgame2}(50k - b, 0)) = 50k - b \leq 25k$
max(right tree)
$= max_{subgame3}(25k, 0) = 25k \leq 25k$

This concludes that the maximum payoff achievable in this game where the opponent's strategy remains unchanged is 25,000 and thus this is a pure strategy Nash equilibrium.

The competition is, however, a repeated game against different opponents. If an agent were to play the equilibrium in a tournament where every other agent strictly cooperated, two cooperating agents would bid 30,000, half of their expected payoff, to go second and attempt to reach the highest payoff, 60,000, split evenly between them. They would then each earn 30,000 points in the games between each other and 25,000 against the optimal agent.

Observation 2. *The Nash equilibrium will not always be the best strategy to use in the competition.*

To show this observation we created Table 1 showing the payoffs two agents would receive after selecting each of the simple strategies defined earlier.

	Cooperative	Defector	Optimal	Null
Cooperative	30k , 30k	0 , 50k	25k , 25k	0 , 50k
Defector	50k , 0	0 , 50k	25k , 25k	0 , 0
Optimal	25k , 25k	25k , 25k	25k , 25k	0 , 0
Null	50k , 0	0 , 0	0 , 0	0 , 0

Table 1: Payoffs between each matchup of the simple strategies described earlier.

From this table we can see that if an agent used the optimal strategy in a competition where all other agents used the cooperative strategy it would receive a lower total payoff. However, if instead the agent used the defector strategy it would win over all the cooperative agents. This shows that this observation holds true in the round robin design and is important in promoting a variety of strategies. As mentioned earlier, the best player may not be the one that plays against each opponent best, but the one that understands the actions of the group (Golbeck 2002). In order to gain knowledge and make these strategic plays from each level in the tournament, an agent must have the ability to analyze the levels. The agents must be able to anticipate how destructive a shot might be so as not to let the opposing agent take the winning shot.

Theorem 1. *An agent that cannot predict the outcome of its actions cannot win against an agent that can when playing a level that requires multiple shots to solve.*

We consider being able to predict the outcome of an action to include knowing whether or not a shot will destroy all the remaining pigs, completing the level. This allows an agent to know whether or not it can reach a terminal node and allows it to then use the null strategy described earlier. Any strategy, when played against a null strategy, will always receive a payoff of zero and the null agent will receive a zero or higher payoff. The null agent strategy is simple and uninteresting but it demonstrates through this theorem that the competitions design requires agents to be able to analyze the consequences of a shot to win.

The Results of the 2015 Competition

There were two naive agents competing here by referred to as 'Zero' and 'Heuristic' and two team entered agents IH-SEV and s-Bird Returns. The two naïve agents differ only in the bids that they make. The 'Zero' agent always bids zero and the 'Heuristic' agent bids half the amount it expects to gain from the level if each pig takes one bird to destroy, which is the amount the pigs are worth plus what

the remaining birds are worth. Both of these naive agents then always shoot towards a random pig during their turn.

The naive agent's idea of shooting at a random pig was very successful in previous competitions where the goal was to get the highest score. The naive agent won the first AIBIRDS competition in 2012 and since then still scores decently in the main track. In this new competitive track, both naive agent variants ended with zero points overall. Meanwhile, the winner IHSEV scored 127610 points and second s-Birds Returns scored 61450 points overall. This is likely due to the naïve agents' strong performance in the main competitive track where they were allowed to repeat the same level multiple times. This allowed them to eventually randomly pick the right trajectories and the sequence of pigs to shoot to get a high score. An agent such as this performing badly compared to its more analytical competitors was one of the goals of this new competitive track. This gap is very clear in the results of this year's competition.

Discussion

The results of this year's competition saw the track successfully inspire a simulation module as a part of one of the agents entered. This was shown to be valuable as it allowed the team to implement a bidding strategy that helped their agent win the competition. This was in alignment with our claim that the competition will refocus the agents on analysis. Unfortunately, participation in this year's competition was limited. The time frame to develop an agent after the release of the software was very short and was most likely a factor in the low number of signups. This was also the first time the competition had been run and in the future it will hopefully see more success.

The extensive form game model we created was able to show the competition design rewards analysis but lost some of the complexities of Angry Birds. One example was in the second level of Poached Eggs. The level can be solved in a single shot by utilizing a domino effect. However, if an agent shoots the middle pig, removing a domino from the chain, the level can at best be completed in two additional shots. This changes the way the game tree might look to something more complex and generating this model may only be done by an agent who can analyze all the potential states of a level. Theorem 1 is also only supported by this model which assumes the agents have equal capabilities. If one agent could solve a level in half the number of shots of its opponent, it would be possible for it to win with only greedy shots.

Conclusions

In this paper we described the design of the new competitive track for the Angry Birds Artificial Intelligence competition. We gave insight into the motivations of this new track and highlighted the design decisions that drive agents to be more analytical. We modelled the new competitive track as an extensive form game and evaluated simple strategies to show that different configurations of participants in a tournament affected which was best. The design choices achieved the goal of creating a competition in which its competitors will contribute to the development of Artificial Intelligence predicting the consequences of real-world actions.

References

Golbeck, J. 2002. Evolving strategies for the Prisoner's Dilemma. *Adv. Intell. Syst., Fuzzy Syst., Evol. Comput., 299 -306.*

Ferguson, T. S. 2014. Game Theory, Second Edition. *Part III: 2-4.*

Axelrod, R. 1997. The Complexity of Cooperation. *Princeton: Princeton University Press.*

Wedekind, C., Milinksi, M. 1996. Human cooperation in the simultaneous and alternating prisoner's dilemma: Pavlov versus Generous Tit-for-Tat. *Proc. Natl. Acad. Sci. U.S.A. 93, 2686–2698.*

J. Renz, AIBIRDS 2015 The Angry Birds Artificial Intelligence Competition. *Proceedings of the 29th AAAI Conference (AAAI'15), Austin, TX, January 2015, 4326-4327.*

Osborne, Martin, J., Rubinstein, A. 1994. A Course in Game Theory. *Cambridge, MA: MIT Press.*