Version 01 as of May 12, 2014
Author: Steven Sagona
To be submitted to PRL (not really)

The Game Theory of Pokemon: AI implementing Nash Equilibrium

(Dated: May 12, 2014)

This paper is written to propose an algorithm that uses game theory to create an AI for multiple turn simulteneous games like Pokemon. The paper will introduce the concepts of Nash Equilibriums. [Reference to Technical Machine.]

PACS numbers:

INTRODUCTION

In Chess, we can use the "minimax algorithm" to determine what a player's best move is. But in "simultaneous games" like Pokemon, when players make their choices simulteneously, ie without knowing their opponent's decision - we cannot use the minimax algorithm.

And, in fact, there are situations in "simultaneous games" where there is no single best move.

Let's look at the following example from Pokemon: You have a weakened Charizard against a Lucario. If the Charizard stays in it is killed by Lucario's Quick Attack. But if Charizard switches out to avoid the Quick Attack, you risk Lucario getting off a Swords Dance. So do you predict Lucario's Swords Dance by attacking him with Flamethrower - or not risk losing your Charizard to a Quick Attack?

In this example, our intuition tells us that neither move should be used 100 percent of the time. This is because it would allow your opponent an opportunity to capitalize on your choice - should they predict your decision. If it's clear that if your opponent notices that you always play it safe and switch, they're going to take advantage of that fact and exploit your tendency.

So our intuition tells us that to avoid risk of being taken advantage of, some of the time we should use Flamethrower, and some of the time we should switch. But exactly how much of the time should we use Flamethrower and how much of the time should we switch? 50/50? 70/30? These probabilities are what we are interested in finding. Strategies that involve multiple possible actions with assigned probabilities are called a "mixed strategies." And we are interested in finding which mixed strategy is the "safest" mixed strategy - meaning the strategy that is impossible for our opponent to capitalize on by using prediction. This "safest mixed strategy" solution is famously known in economics as a "Nash Equilibrium."

First we will go over some simple examples to familiarize the reader with Nash Equilibriums, and then we will continue by seeing how we can find the Nash Equilibrium for games of Pokemon.

Intuitive Example

You and your friend Bob decide to split a cookie, but can't decide on how to split it. You both decide that you will cut the cookie in 2 pieces, but your friend Bob gets to choose which piece of the cookie he wants.

What is the safest way for you to cut the cookie, ensuring you get the largest piece possible? Your safest strategy would be to cut the cookie exactly in half - and you can think of this strategy as effectively being the "Nash Equilibrium". By splitting the cookie in half, you've made it so that it doesn't even matter which one he chooses - he will still get half a cookie. This is a particular example illustrating that the safest strategy makes the payoffs of all of your opponent's decisions identical - meaning that there is no particular choice that can be made by your opponent that can capitalize on your strategy.

But there is one key assumption here that's important to point out. Splitting the cookie exactly in half is only the best strategy if we assume that Bob /wants/ the cookie. If Bob didn't want the cookie, we could've split the cookie into two different sizes; he would gladly take the smaller piece and we would get a larger piece. This is a really important point to emphasize for the Nash Equilibrium. Even though we are considering the Nash Equilibrium when seeking to maximize the size of the cookie we will get, A Nash Equilibrium does NOT guarantee that we get the maximum size. It instead gives you the best possible, "safest" strategy - the strategy that prevents the opponent from being able to capitalize on your decision.

It might be a little bit confusing that this cookie example didn't say anything about probabilities. The Nash Equilibrium just represents the safest strategy. Sometimes this strategy is a "mixed strategy" - when you assign probabilities to each possible action and roll a die to determine which you will use. But a Nash Equilibrium can also just be a "pure strategy" meaning that you choose one action 100

2D Nash Solution

Let's go back to our Pokemon example and see if we can find the Nash Equilibrium, ie the safest mixed strategy. But first we have to assign some numbers to our situation. In the cookie example we just did, we sought to maximize the size of the cookie-piece that we'll get. But in Pokemon, it's not so clear what we should be maximizing. Health? Damage done? Number of Pokemon? This is one of the core questions we will try to tackle; and we will dive into this in detail later, but for now let's say that we're trying to maximize "difference in pokemon." Table 1 illustrates the numbers we've assigned to our pokemon game. Using your Flamethrower against the enemy's Lucario's Swords Dance results in Lucario fainting, which results in you having one more pokemon than the opponent. Likewise if he uses Quick Attack when you use Flamethrower, it results in a loss in one of your pokemon. And if Lucario uses Swords Dance when you switch, he will then be able to sweep two of your pokemon, hence the negative two.

TABLE I: Simple Rock Paper Scissors Payout Matrix

	Swords Dance	Quick Attack
Flamethrower	1	-1
Switch	-3	0

What is the Nash Equilibrium for this system. Well earlier we stated that the nash equilibrium is the "safest strategy" - the choice that makes it so that the opponent is unable to capitalize on our decision. This means that he payoff for the opponent using Swords Dance should be identical to the payoff of our opponent using Quick Attack. [explain payoff]

$$\begin{aligned}
&\text{Payoff of Opponent using Swords Dance:}\\
&\quad = \text{Probability using Flamethrower}\\
&\qquad\quad \bullet (\text{result of Swords Dance vs. Flamethrower})\\
&\qquad + \text{Probability of switching}\\
&\qquad\quad \bullet (\text{result of Sword Dance vs switch})\\
&\quad = P(F) \bullet (1) + P(S) \bullet (-3)\\
&\quad = P(F) - 3P(S)
\end{aligned}$$

Using our matrix of Table 1, we can write out the expected payoff of the opponent using Swords Dance. We know what the "result of Swords Dance vs. Flamethrower" is (winning +1 pokemon), and the result of Swords Dance vs. switch (losing -3 pokemon) - but we don't still don't know what our probabilities of using Flamethrower and Swords Dance should be. We'll notate P(F) - probability of using Flamethower and P(S) probability of using Swords Dance.

But we know that the Nash Equilibrium means that the expected payoff of using Swords Dance should be identical to the expected payoff for Quick Attack. And we can find the expected payoff for Quick Attack:

$$\begin{aligned}
&\text{Payoff of Opponent using Quick Attack:}\\
&\quad = \text{Probability using Flamethrower}\\
&\qquad\quad \bullet (\text{result of Quick Attack vs. Flamethrower})\\
&\qquad + \text{Probability of switching}\\
&\qquad\quad \bullet (\text{result of Quick Attack vs switch})\\
&\quad = P(F) \bullet (-1) + P(S) \bullet (0)\\
&\quad = -P(F)
\end{aligned}$$

Now by setting these two payoffs equal to eachother we are asking: "What probabilities should I choose such that these two payoffs are equal?" (This set of probabilities is the so-called Nash Equilibrium!)

$$\begin{aligned}
&\text{Payoff of Opponent using Quick Attack}\\
&\quad = \text{Payoff of Opponent using Flamethower}\\
&\Longrightarrow P(F) + -3P(S) = -P(F)
\end{aligned}$$

We see that we have two variables we're interested P(F) and P(S), but only one equation. But we remember that P(S) and P(F) are the only two possible strategies, meaning that the sums of their probability should add up to 1. And now we have 2 equations and two unknowns, and can solve.

$$\begin{aligned}
&P(F) + -3P(S) = -P(F)\\
&P(F) + P(S) = 1\\
&\Longrightarrow P(F) = 3/5\\
&\Longrightarrow P(S) = 2/5
\end{aligned}$$

This says that your "safest strategy" is to use Flamethrower 3/5'ths of the time and switch $\frac{2}{5}$ of the time.

$$\begin{aligned}
&\text{Payoff when opponent uses Swords Dance:}\\
&\quad = P(F) \bullet (1) + P(S) \bullet (-3)\\
&\quad = \frac{3}{5} \bullet (1) + \frac{2}{5} \bullet (-3)\\
&\quad = -\frac{3}{5}
\end{aligned}$$

Payoff when opponent uses Quick Attack:

$$= P(F) \bullet (-1) + P(S) \bullet (0)$$
$$= \frac{3}{5} \bullet (-1)$$
$$= -\frac{3}{5}$$

And as expected we see that our payoffs are identical no matter what our opponent's choice is. In fact, since the expectation values are the same for Swords Dance and Quick Attack, any "mixed strategy" is *also* going to to have the same expectation (For example, if the opponent does 25% Swords Dance, 75% Quick Attack - we still get -$\frac{3}{5}$ as our payoff).

Now let's interpret the meaning of this expected payoff. The payoff is always an *average value*. This means that for our example, if you were to encounter this situation 50 times in total, we expect you to have lost around $\frac{3}{5} \bullet 50 = 30$ pokemon.

And this is our best strategy? It clearly doesn't seem like it. For example, we can see that if the opponent does Swords Dance 50% of the time and Quick Attack 50% of the time, then we can capitalize on our opponents strategy by picking Flamethrower 100% of the time, and now our payoff is 0 on average instead of $-\frac{3}{5}$. The bottom line is that by not picking a Nash Equilibrium strategy, we very well may get a higher payoff - but by doing so we are risking our opponent capitalizing on our strategy.

So we found the Nash Equilibrium mixed strategy for the Charizard. But what about the Nash Equilibrium for the Lucario?

Payoff using Flamethrower = Payoff using switch

$$P(SD) - P(Q) = -3P(SD)$$
$$P(Q) + P(SD) = 1$$
$$\implies P(SD) = 1/5$$
$$\implies P(Q) = 4/5$$

Lucario should use Swords Dance 1/5'th of the time and Quick Attack 4/5'ths of the time. Notice that if we calculate the expected payoff of either Flamethower or Swords Dance we get the same result as before: -3/5.

So there is an important but somewhat strange concept that's important to identify here. If I decide to implement my Nash equilibrium it means that no matter what my opponent does - my expected outcome will be the same. This is a curious, and somewhat frustrating conclusion. By choosing the Nash equilibrium, you are giving up the idea of capitalizing on your opponents choices while also simultaneously making it so that the opponent's choices are irrelevant.

Even though it seems as though you are giving up opportunities to maximize your payoff, Nash Equilibriums are still very useful. Let's say that you instead chose to use Flamethower half the time, and switch half the time. But by choosing this mixed strategy that's not at Nash equilibrium, your opponent has an opportunity to capitalize on your decision. If they use Swords Dance 100% of the time, this means that you have an expected payoff of -1 (which is worse than the Nash payoff of -3/5'ths).

So in summary, if you choose a Nash equlibrium strategy, no matter what strategy your opponent picks, your expected payoff does not change. If you go outside of a Nash Equilibrium strategy, you risk your opponent being able to capitalize on your payoff and do better than the Nash equilibrium's payoff. But by deciding to capitalize on you being outside of a Nash equilibrium strategy, they must also go outside of their Nash equilibrium mixed strategy. And by doing so, they also are risking being capitalized on themselves.

[not including bad strategies]

Nash Equilibrium for Rock-Paper-Scissors

We will do one more example. We will cover the Nash Equilibrium for the game of rock-paper-scissors. Can you guess what the Nash Equilibrium rock-paper-scissors is? We can solve for it explicitly, but we would expect the safest strategy to be exactly 1/3 rock, 1/3 paper, 1/3 scissors.

TABLE II: Simple Rock-Paper-Scissors Payout Matrix

	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

Table II illustrates the payout matrix for rock paper scissors. Notice that I chose 1 and -1 to be the numbers that signify a win or a loss - and it can be shown that these numbers can be anything so long as they are "zero-sum," meaning that a win for me is a loss for you (and 15 "points" for me is -15 "points" for you in your perspective; 103.3 banana's I gain is the same as you losing 103.3 bananas).

We seek the Nash Equilibrium for Player 1. Therefore we want to express the payoffs of Player 2 and find the probabilities Player 1 should implement to make Player 2's expected payoffs equal.

$$P(R1) = \text{Probability of Player 1 using Rock}$$
$$P(P1) = \text{Probability of Player 1 using Paper}$$
$$P(S1) = \text{Probability of Player 1 using Scissors}$$

Using the notation listed above we can express the various payoffs for Player 2 as follows:

$$\begin{aligned}
&\text{Player2's Payoff using Rock} = \\
&= 0 * P(R_1) + -1 * P(P_1) + 1 * P(S_1) \\
&\text{Player2's Payoff using Paper} = \\
&= 1 * P(R_1) + 0 * P(P_1) + -1 * P(S_1) \\
&\text{Player2's Payoff using Scissors} = \\
&= -1 * P(R_1) + 1 * P(P_1) + 0 * P(S_1)
\end{aligned}$$

And now we can set these payoffs equal to each other. Note that we set Rock = Scissors, and Paper = Scissors, but don't set Rock = Scissors. This is because the third equation is useless because it doesn't tell us any new information we don't already know (since a=b, b=c $\implies$ a=c). But we can obtain a third equation using the fact that our probabilities add up to 1:

$$\begin{aligned}
&\text{P2's payoff using Rock} = \text{P2's payoff using Scissors} \\
&\text{P2's payoff using Paper} = \text{P2's payoff using Scissors} \\
&P(R_1) + P(P_1) + P(S_1) = 1
\end{aligned}$$

We now have 3 equations and 3 variables, and we can now solve for $(P(R_1), P(P_1), \& P(S_1))$ via whatever algebra method is easiest. For higher-dimension examples, an understanding of linear algebra is particularly helpful. So we'll show you how to solve this example using linear algebra.

Rewriting our variables for simplicity,($P(R_1)$= R, etc.) our 3 equations can be written as:

$$\begin{aligned}
R - S &= -P + S \\
R - S &= -R + P \\
R + S + P &= 1
\end{aligned}$$

We can then move all the variables to the left hand side and express this as a matrix:

$$\begin{bmatrix} 1 & 1 & -2 \\ 2 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \bullet \begin{bmatrix} R \\ P \\ S \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We happen to be very lucky in this case. Since the determinant of this matrix is not zero, it means that an inverse of this matrix exists (this is a well-known linear algebra theorem, which you can google for the proof). And since the inverse exists, we can multiply both sides by the inverse. Lo and behold, the result is 1/3, 1/3, 1/3 - exactly what we expect for a game of rock-paper-scissors.

Weighted Rock-Paper-Scissors

We can make things more interesting and change the numbers around to create a game of weighted Rock-Paper-Scissors:

TABLE III: Symmetric Weighted Rock-Paper-Scissors

	Rock	Paper	Scissors
Rock	0	-1	5
Paper	1	0	-2
Scissors	-5	2	0

We can interpret the above table as a game of Rock-Paper-Scissors, where player's receive 5$ for winning with Rock, 1$ for winning with Paper, and 2$ for winning with Scissors. We could use the same analytical methods to calculate the Nash equilibrium, and we'd find that we should use Rock 2/8ths of the time, Paper 5/8ths of the time, and Scissors 1/8th of the time. It's interesting to note that the weakest move, Paper, is used more than twice as much as Rock even though it has one fifth the payoff of Rock.

Additionally we can do this for an asymmetrical weighted rock-paper-scissors. In the following game, players gain different amounts. For example, in the table we see that Player 1 gains 10$ when winning with Scissors, while Player 2 gains 2$ when winning with Scissors.

TABLE IV: Asymmetric Weighted Rock-Paper-Scissors

	Rock	Paper	Scissors
Rock	0	-3	3
Paper	.5	0	-2
Scissors	-.3	10	0

Again, we can use the same methods as before, but this time each player will have their own different mixed Nash Equilibria.

SIMPLEX ALGORITHM

HUERISTICS STATES FOR POKEMON

So we keep writing out these tables of numbers, but what do these numbers mean in the case of Pokemon? If you recall, we chose "difference in Pokemon" as the numbers that go in our table to calculate our Nash equilibrium. This means we chose "difference in Pokemon" as our heuristic function. A heuristic function is just a function that provides a good rule of thumb for deciding who's winning and who's losing. So knowing that person A has one more Pokemon against person B is a pretty good rule of thumb to decide who's winning. But it's clear that this is clearly not perfect. For example, if you have the option of dealing half health to a Charizard or 100% to a Magikarp - our heuristic function would indicate that it's a better idea to KO the Magikarp than to deal half health to the Charizard. So

We propose a heuristic function for Pokemon called "Offensive Presence."

TABLE V: Asymmetric Weighted Rock-Paper-Scissors

	Vileplume	Exeggutor	Articuno
Infernape	1	1	1
Chansey	1/5	1/5	1/10

TABLE VI: Asymmetric Weighted Rock-Paper-Scissors

	Infernape	Chansey
Vileplume	1/2	1/7
Exeggutor	1/2	1/3
Articuno	1/2	1/6

In the following example, We take

[1] http://wanda.fiu.edu/teaching/courses/Modernlabmanual/images/comptonscattering.png