Towards Optimal Play of Three-Player Piglet and Pig

François Bonnet
School of Computing
Tokyo Institute of Technology
Japan

Todd W. Neller
Department of Computer Science
Gettysburg College
USA

Simon Viennot
Araya, Inc., Tokyo
Japan

Abstract

While two-player optimal pure strategy play of the jeopardy games Piglet and Pig is known, 3-player optimal play was unknown. In this article, we first compute optimal play for the 3-player jeopardy coin game Piglet, and demonstrate a surprising possibility of coalition between players that have no direct interaction. Further, we examine various unfair coin scenarios and find the existence of multiple Nash equilibria. Finally, we turn our attention to the 3-player jeopardy dice game Pig, demonstrate the failure of value iteration to converge for high enough goal scores, and conjecture the existence of mixed strategy Nash equilibria for 3-player Pig.

Piglet

Piglet is a simple coin game played between two or more players. "*The object of Piglet is to be the first player to reach 10 points. Each turn, a player repeatedly flips a coin until either a tail is flipped or else the player holds and scores the number of consecutive heads flipped.*" (Neller and Presser 2004)

Piglet is Neller and Presser's coin simplification of the dice game Pig, the simplest known folk ancestor of the jeopardy dice game family. The commercial games Pass the Pigs (a.k.a. Pigmania) and Farkle are perhaps the best known games of this family.

One can represent a game state of Piglet as a 2-tuple (a, b) where a and b are number of points remaining to reach the winning goal score for the current player and their opponent, respectively. Let k denote the turn total, i.e. the number of consecutive heads flipped so far on a turn. A player's play policy is then a mapping $\pi(a, b)$ to a hold value. Beneath this hold value, the player continues to flip; at this hold value, the player holds and scores that many points. Let us assume both players play with optimal policy π^*. We can then describe optimal play for 2-player Piglet through equations based on the probability of winning $P(a, b, k)$:

$$P(a,b,k) = \begin{cases} 1 & \text{if } a = k \\ \max\left(\begin{array}{c} 1 - P(b, a-k, 0), \\ \frac{(1-P(b,a,0)) + P(a,b,k+1)}{2} \end{array} \right) & otherwise \end{cases}$$

As described in (Neller and Presser 2004), optimal policy may be computed with a value-iteration based method that iteratively takes arbitrary initial probabilities for each $P(a, b, k)$, evaluates right-hand sides of the equations above with these estimates, and computes better estimates of left-hand side values. This process converges, yielding the following optimal hold values in playing Piglet for a goal score of 10:

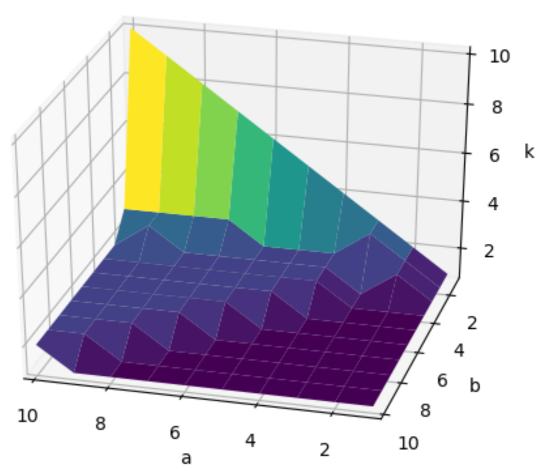

Figure 1: Optimal 2-player Piglet hold values where a and b represent the distance from the goal score for the current player and opponent, respectively, and k represents the turn total at which the current player should hold.

In most cases, when one is ahead in score, one should hold at 1. When one is behind, one generally holds at 2. However, one should take greater chances when the opponent is close to winning. In particular, when $b = 1$ then the optimal hold value $k = a$, i.e. the current player should play to win in this turn.

Three-Player Piglet

Although Neller and Presser proposed a value-iteration-based method to compute the optimal strategy of the 2-player version of Piglet and Pig (Neller and Presser 2004), and of many Pig variants (Neller and Presser 2005), no analysis with more players has been done.

In this research, we first investigate the 3-player version of Piglet. While we expected it to be a trivial generalization of the 2-player version, it appears to be quite different. Indeed 3-player games are generally much harder to analyze. As described for example in (Bewersdorff 2005, Chapter 43) for a Poker-like game, coalitions between players are possible in 3-player games, so it is difficult to even define “optimal play”.

Playing Independently

Instead of extending the value iteration method used in (Neller and Presser 2004), we design a new computation method for Nash equilibria (NE) (Nash 1950) by exhaustively listing the payoffs of each possible pure strategy for each player. This allows us to compute the exact rational NE of 3-player Piglet for the first time.

Score of players

The game, as described in (Neller and Presser 2005), starts in a state $(0, 0, 0)$ where the score of each player equals zero and then all scores progressively increase until one player reaches the goal G. In this paper, we use a reverse scoring system in which each player initially starts with G and the scores decrease until one player reaches 0.

The advantage of using such system is that the state of the system is completely defined with a triplet (a, b, c). Whatever the initial variable goal G, the target 0 is always the same and therefore the state (a, b, c) is always the same.

Reward system

Multiple reward system could be considered when playing this game. In this paper, we assume first one (1W2L) described below, but other options could also be investigated.

Fixed reward system

- Partial ranking with one Winner, two Losers (1W2L). The winning player gets a $+2$ reward and each loser gets a -1 reward.
- Full ranking:
  - Game continues between the two remaining players. The first winning player gets a reward of x ($x > 0$), the second winning player gets a reward of y ($y \leq x$), and the loser gets a reward of $-(x + y)$.
  - Game terminates after first winner; then players get reward $\{x, y, -(x + y)\}$, or $\left\{z, -\frac{z}{2}, -\frac{z}{2}\right\}$ in case of tie between the two losers.

Adaptive reward system One could also design a more complex system in which the reward of the winner depends on the final score of other players (as in the game of UNO). For example, let consider a game with goal G and assume that player P1 reaches G while players P2 and P3 have score s_2 and s_3. It is possible to take $-s_2$ for P2 reward, $-s_3$ for P3 reward, and finally $s_2 + s_3$ for P1.

Such reward system is much more complex to analyze. Techniques proposed in this paper cannot be applied easily.

Analysis P1*vs*P2*vs*P3 under 1W2L fixed reward system

Definitions

Expected gain Let us fix a strategy Σ_1, Σ_2, and Σ_3 for each player P1, P2, and P3 respectively. Given these strategies and an arbitrary current score (a, b, c), let us define $E(\Sigma_1, \Sigma_2, \Sigma_3, i, j, a, b, c)$ as the expected gain of player Pi given that player Pj is about to play (no coin-flip yet). To simply notation, we will omit Σ_1, Σ_2, and Σ_3 and simply write $E(i, j, a, b, c)$ when there is no ambiguity.

Optimal expected gain Assuming *independent players* (*i.e.* no coalition), there exists strategy Σ^* which corresponds to a Nash Equilibrium of the game if all three players follow Σ^*. Let us denotes by E^* the expected gains when all players follow Σ:

$$E^*(i, j, a, b, c) = E(\Sigma^*, \Sigma^*, \Sigma^*, i, j, a, b, c) \qquad (1)$$

Remark Due to symmetries, the following equalities hold:

$$\begin{cases} E^*(1, 1, a, b, c) = E^*(2, 2, c, a, b) = E^*(3, 3, b, c, a) \\ E^*(1, 2, a, b, c) = E^*(2, 3, c, a, b) = E^*(3, 1, b, c, a) \\ E^*(1, 3, a, b, c) = E^*(2, 1, c, a, b) = E^*(3, 2, b, c, a) \end{cases} \qquad (2)$$

Any expected gain can be expressed when the first player is about to play (*i.e.* using $E^*(i, 1, \cdot, \cdot, \cdot)$ notations). To simplify notations, we use $E^*_i(\cdot, \cdot, \cdot)$ as an alias for $E^*(i, 1, \cdot, \cdot, \cdot)$ and will write all equations using only $E^*_i(a, b, c)$.

Analysis

Compute expected gains Given strategies Σ_1, Σ_2, and Σ_3 for each player P1, P2, and P3 respectively, we can compute expected gains progressively (as explained in (Neller and Presser 2005) for iteration value method). Figure 2 depicts a situation in which player P1 is about to play and current score is (a, b, c). In score states, underlined number indicates the player who has to play. The initial state is indicated by the bold arrow and possible final states are drawn in bold. While score remains (a, b, c), each player is trying to holds at α, β, and γ heads respectively. These values α, β, and γ are given by strategies Σ_1, Σ_2, and Σ_3 respectively.

Since the starting state is $(\underline{a}, b, c)$, we want to compute $E(i, 1, a, b, c)$ for $1 \leq i \leq 3$. To do so, we only need to know $E(i, 2, a-\alpha, b, c)$, $E(i, 3, b-\beta, c)$, and $E(i, 1, a, b, c-\gamma)$ for all $1 \leq i \leq 3$. Solving the simple Markov chain leads to the following equations, that can be used to compute the expected gains for each player given a strategy triplet $(\Sigma_1, \Sigma_2, \Sigma_3)$.

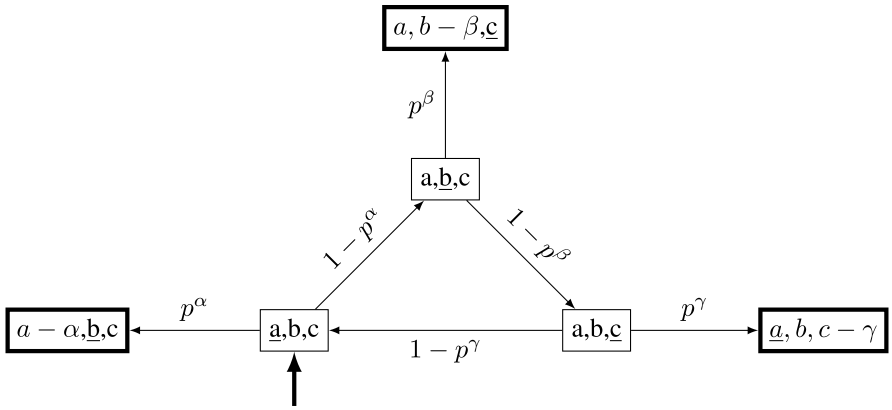

Figure 2: From score (a, b, c) until some player successfully holds, when P1, P2, and P3 holds at α, β, and γ heads respectively.

$$\begin{aligned}\forall i \in \{1,2,3\} & E(i,1,a,b,c) \\ &= \frac{p^\alpha E(i,2,a-\alpha,b,c)}{p^\alpha + (1-p^\alpha)p^\beta + (1-p^\alpha)(1-p^\beta)p^\gamma} \\ &+ \frac{(1-p^\alpha)p^\beta E(i,3,a,b-\beta,c)}{p^\alpha + (1-p^\alpha)p^\beta + (1-p^\alpha)(1-p^\beta)p^\gamma} \\ &+ \frac{(1-p^\alpha)(1-p^\beta)p^\gamma E(i,1,a,b,c-\gamma)}{p^\alpha + (1-p^\alpha)p^\beta + (1-p^\alpha)(1-p^\beta)p^\gamma}\end{aligned} \quad (3)$$

Compute NE expected gains Theoretically, it is easy to find the optimal strategy Σ^*; it is sufficient to compute expected gains $E(\Sigma, \Sigma, \Sigma, \cdot, \cdot, \cdot, \cdot)$ for all possible Σ and then find which strategy/strategies correspond to a Nash Equilibrium. It is of course impossible to do in practice, even for small instances of the game (except when the goal is 1 or maybe 2).

We can use Figure 2 to help us and compute progressively the NE strategy. Let us consider a score triplet (a, b, c). Let us assume that we already know partially the NE strategy Σ, more precisely we already know the moves for all triplet (a', b', c') such that $a' + b' + c' < a + b + c$. It implies that we already know the expected gains $E^*(i, 2, a-\alpha, b, c)$, $E^*(i, 3, b-\beta, c)$, and $E^*(i, 1, a, b, c-\gamma)$ for $1 \leq i \leq 3$ and $1 \leq \alpha \leq a$, $1 \leq \beta \leq b$, and $1 \leq \gamma \leq c$.

Given the score triplet (a, b, c), player P1 can choose among a strategies; he can try to reach either 1 head, or 2 heads, …, or a heads. Similarly players P2 and P3 can choose among b and c strategies respectively. There are therefore abc triplet of strategies to analyze. Using Equations 3, we can compute the expected gains for each player from state $(\underline{a}, b, c)$ assuming that players play strategies α, β, and γ respectively.

Comparing these abc options, we can find (one of) the Nash Equilibrium which corresponds to a triplet of strategies $(\alpha^*, \beta^*, \gamma^*)$.

Results

Figure 3 shows the expected gain of each player when a win is rewarded $+2$ and a loss -1, for goal scores from 1 to 30. Without any surprise, expected gains decrease according to player order. Note that the gains are not monotonic, e.g. the expected gain of Player 2 decreases slightly when the goal score goes from 3 to 4.

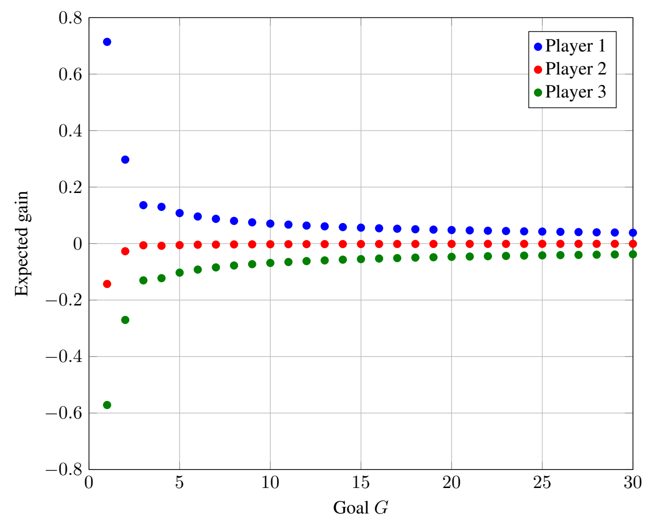

Figure 3: Expected gain versus goal score for three players using NE strategy.

Playing with Coalitions

Our computation method allows us to consider different objectives for the players. Assuming $\max_n$ play, i.e. that each player seeks to maximize their expected utility, we obtain the results of the "Playing Independently" section, yet if we instead assume that two players seek to minimize the expected utility of the third, our computations show that such two-player coalitions are possible and effective in the game of Piglet.

Figure 4 shows the expected gain of Player 1 with and without a coalition of Players 2 and 3 for different goal scores. Though the difference is small, a coalition between Players 2 and 3 is an effective possibility above a goal score of 4 points. When Players 2 and 3 collude against Player 1, the expected gain of Player 1 decreases. It is worth mentioning that the expected gain of Player 3 also decreases in this coalition, representing self-sacrifice for the coalition's goal.

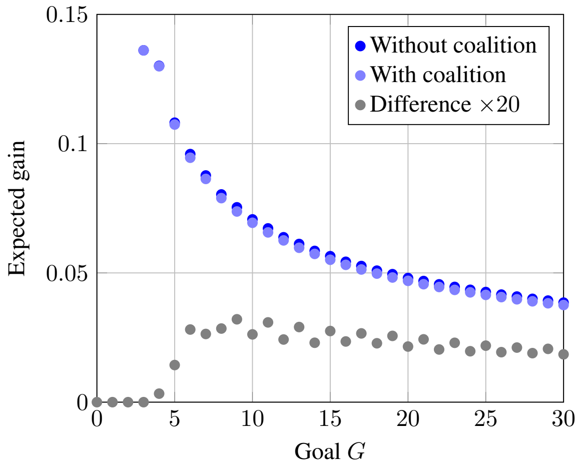

Figure 4: Expected gain of P1 with/without coalition of P2 and P3.

This possibility of coalitions is surprising because players have no direct interaction and thus no direct way to help each other in a game of Piglet. Still, coalitions are possible because contentment with a coalition partner winning changes play decisions. For example, consider a situation where two players in a coalition are leading a third they are cooperating against. Under normal non-coalition circumstances, the two players would aggressively push each other to greater risk-taking to be first to the goal. However, if their common objective is to beat the third player, both can more conservatively with respect to their common adversary.

While not so absolute, this can be viewed as a probabilistic generalization of the "Kingmaker scenario". A player trailing the leader may cooperate/compete with the leader to make that leader's victory more/less likely. While this effect is not large, it is nonetheless interesting that two players who collude against a third across games of Piglet can achieve better performance than the third player on average.

Playing with a biased coin

We also computed what happens when playing Piglet with a biased coin. Figure 5 shows the expected gain of Player 1 as a function of the probability p of flipping a Head. Each curve corresponds to a given goal score G.

Figure 5 shows that when the probability of flipping a Head increases, the first player has globally a higher probability of reaching the goal score first. This is not surprising, since a greater chance of flipping a Head is likely to help the player increase its score faster. However, it is counter-intuitive that this expected gain is not monotonic with respect to the Head probability. For example, for a goal score $G = 10$ (red curve), Player 1's expected gain is higher when $p = 0.45$ than when $p = 0.5$.

Also in Figure 5, some curves for different goal scores cross at some points. As noted in the "Playing Independently" section, this shows that the expected gain of Player 1 does not always monotonically decrease as the goal increases.

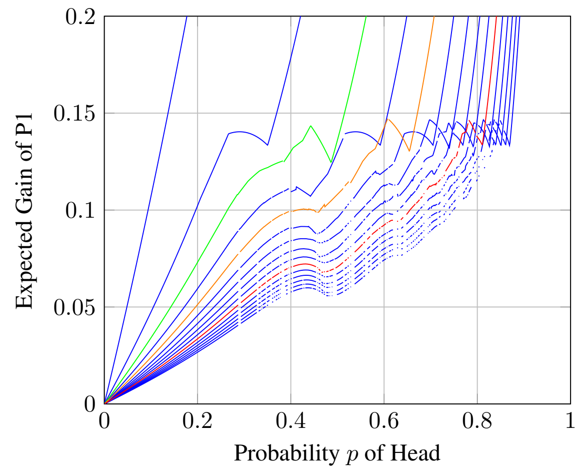

Figure 5: Each curve corresponds to a fixed goal G. Top curve is for $G = 1$, green curve is for $G = 3$, orange curve is for $G = 5$, red curve is for $G = 10$.

Some parts of the curves are missing in Figure 5. The reason is that for these instances of the game, there are multiple Nash equilibria with different payoffs, and thus no single optimal solution. Which solution concept to prefer is under consideration.

Two-Player Pig

The game of Pig is a simple jeopardy dice game in which the objective is to be the first player to score at least 100 points. Each turn, a player repeatedly rolls a die until either a 1 is rolled or the player holds and scores the sum of the rolls (i.e., the *turn total*). At any time during a player's turn, the player is faced with two choices: *roll* or *hold*. If the player rolls a 1, the player scores nothing and it becomes the opponent's turn. If the player rolls a number other than 1, the number is added to the player's turn total and the player's turn continues. If the player instead chooses to hold, the turn total is added to the player's score and it becomes the opponent's turn.

In forming optimality equations for two-player Pig, let $P_{i,j,k}$ be the player's probability of winning if the player's score is i, the opponent's score is j, and the player's turn total is k. In the case where $i+k \geq 100$, $P_{i,j,k} = 1$ because the player can simply hold and win. In the general case where $0 \leq i, j < 100$ and $k < 100 - i$, the probability of an optimal player winning is

$$P_{i,j,k} = \max\left(P_{i,j,k,roll}, P_{i,j,k,hold}\right)$$

where $P_{i,j,k,roll}$ and $P_{i,j,k,hold}$ are the probabilities of winning if one rolls and holds, respectively. These probabilities are given by:

$$P_{i,j,k,roll} = \frac{1}{6}((1 - P_{j,i,0}) + \sum_{2 \le r \le 6} P_{i,j,k+r})$$

$$P_{i,j,k,hold} = 1 - P_{j,i+k,0}$$

The probability of winning after rolling a 1 or holding is the probability that the other player will not win beginning with the next turn. All other outcomes are positive and dependent on the probabilities of winning with higher turn totals.

The value-iteration-based algorithm of (Neller and Presser 2004) can then solve for all state win probabilities, and derives the following optimal roll/hold boundary:

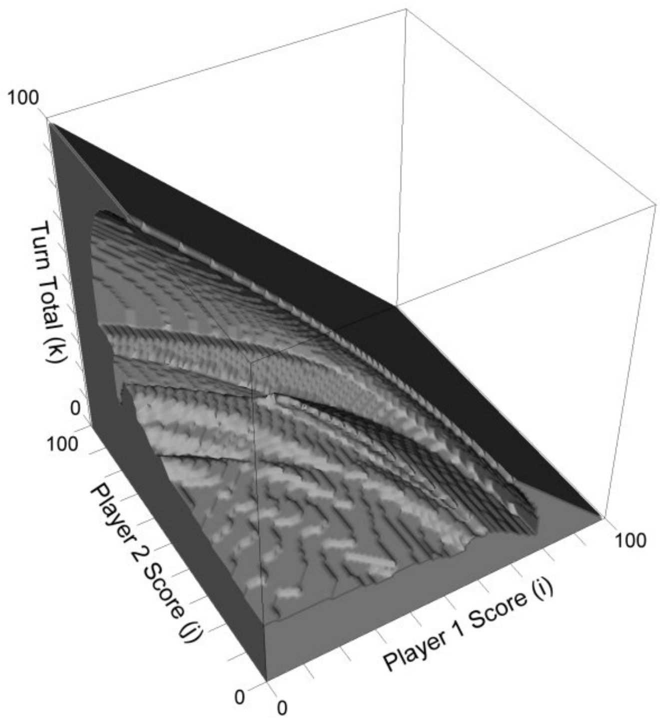

Figure 6: Optimal 2-player Pig roll/hold boundary, where a player playing a state should roll or hold if they are inside or outside of the surface, respectively.

Three-Player Pig

The same value-iteration-based approach applied to 3-player Pig converged to an optimal policy for goal scores up to 61, but did not converge for goal scores observed at 62 at above. Our equations were formed as above, but with j_1 and j_2 representing the scores of the next player and the player thereafter respectively in the 3-player game. Variables i and k retain the same interpretation as above.

The main difference is that we must compute a win probability distribution over all players for each non-terminal game state. Let $P_{i,j_1,j_2,k}[p]$ be the probability that the player p turns into the future will win. Thus, with this 0-based indexing, $p = 0$, 1, and 2 refers to the current, next, and next next players, respectively.

$$P_{i,j_1,j_2,k} = \max(P_{i,j_1,j_2,k,roll}, P_{i,j_1,j_2,k,hold})$$

$$P_{i,j_1,j_2,k,roll} = \frac{1}{6}(\text{right rotate}(P_{j_1,j_2,i,0}) + \sum_{2 \le r \le 6} P_{i,j_1,j_2,k+r})$$

$$P_{i,j_1,j_2,k,hold} = \text{right rotate}(P_{j_1,j_2,i+k,0})$$

In-place updates: Performing in-place updates for a goal score of 62 leads to a cycle of policy changes taking turns alternating between roll and hold for (i, j_1, j_2) values $(0, 1, 30)$ and $(30, 0, 1)$.

Separate policies: Next, we experimented with value-iterating separate policies for each player. Even though player decisions are symmetric, we were interested to see if differing policies could settle to an equilibrium. However, this approach also had non-convergence cyclic behavior for a goal score of 62 for states $(0, 28, 36)$ and $(28, 36, 0)$.

Randomized updates: Next, we added randomization to the previous experiment to see if the strict ordering of updates was responsible for the observed policy update cycling. Selection of non-terminal states for a value-iteration update was entirely randomized. Interestingly, the same $(0, 28, 36)$ and $(28, 36, 0)$ cycle appears, albeit with less frequent flip-flopping.

Gamma decay: Bellman's equations are not guaranteed to converge when one has either unbounded rewards or no discounting, i.e. discount factor $\gamma = 1$. We experimented with $\gamma = .9$. This only changed our cycle to occur between $(0, 6, 24)$ and $(24, 0, 6)$.

Regret minimization: After all prior attempts, we conjectured that it may be the case that no pure Nash equilibrium exists for a goal score of 62. To test this, we implemented a value-iteration-based approach that updates mixed strategies through regret minimization. This approach also cycled output policies with our original pair of states $(0, 1, 30)$ and $(30, 0, 1)$, that is, the output mixed strategy alternately fluctuates between preferring roll and hold for each of these two states. It may be that this algorithm would converge to an equilibrium of an equiprobable choice of roll or hold for these states, but we have not observed such convergence at this time.

We still conjecture that a mixed Nash equilibrium exists for 3-player Pig with a goal score of 62. What has become clear is that the dynamics of these iterative approaches leads us into repeatable policy orbits despite all of these variations. Work to determine a best approach to solving 3-player Pig is still in progress.

Conclusion

While Piglet is one of the simplest jeopardy games, optimal 3-player strategy is quite complex. We computed Nash equilibria for many instances of 3-player Piglet and observed that coalitions between players are possible. Also, in some cases, there are multiple NE with different payoffs, so that it is not obvious what should be considered "optimal play". There are still many open questions about 3-player Piglet.

Three-player Pig remains unsolved. At this point, we are inclined to believe that iterative approaches of many types tend to orbit among approximately-optimal policies. It may be that a different class of solution technique is required for such problems.

An interesting question to consider is whether there are potential relationships between the policy orbits observed for 3-player Pig, and the concept of coalitions seen in 3-player Piglet. Whether or not the flip-flopping of 3-player Pig policies corresponds to a type of making and breaking of coalitions is a possible future question to address.

Just as 2-player Pig offered up surprisingly complex optimal play policy for such a simple ruleset, 3-player Pig offers surprises as we seek its solution.

References

Bewersdorff, J. 2005. *Luck, Logic and White Lies*.

Nash, J. F. 1950. *Non-Cooperative Games*. Ph.D. Dissertation, Princeton University.

Neller, T. W., and Presser, C. G. 2004. Optimal play of the dice game pig. *The UMAP Journal* 25:25–47.

Neller, T. W., and Presser, C. G. 2005. Pigtail: A pig addendum. *The UMAP Journal* 26:443–458.