Extracting Learned Discard and Knocking Strategies from a Gin Rummy Bot

Benjamin Goldstein, Jean-Pierre Astudillo Guerra, Emily Haigh, Bryan Cruz Ulloa, Jeremy Blum
The Pennsylvania State University
{bjg5493, jpa5180, evh5311, bbc5256, jjb24}@psu.edu

Abstract

Various Gin Rummy strategy guides provide heuristics for human players to improve their gameplay. Often these heuristics are either conflicting or contain ambiguity that limits their applicability, especially for discard and end-of-game decisions. This paper describes an approach to analyzing the machine learning capabilities of a Gin Rummy agent to help resolve these conflicts and ambiguities. There are three main decision points in the game: when to draw from the discard pile, which card to discard from the player's hand, and when to knock. The agent uses a learning approach to estimate the expected utility for discards. An analysis of these utility values provides insight into resolving ambiguities in tips for discard decisions in human play. The agent's end-of-game, or knocking, strategy was derived using Monte Carlo Counterfactual regret minimization (MCCFR). This approach was applied to estimate Nash equilibrium knocking strategies under different rules of the game. The analysis suggests that conflicts in the end-of-game playing tips are due in part to different rules used in common Gin Rummy variants.

Introduction

In the past several years, algorithms to find Nash Equilibria in zero-sum, extensive-form games have been the subject of extensive research; however, the game of Gin Rummy has received little attention. Moreover, strategy guides for Gin Rummy, targeting human players, often contain conflicting or ambiguous suggestions for gameplay. The process of developing an agent for the EAAI Gin Rummy Competition (Neller 2019) provides an opportunity to study these strategy tips, and their alignment with equilibrium strategies.

There are three main decision points for a player in the game of Gin Rummy. In the first decision point, a player decides whether to draw the face-up card. The strategy used by our agent is a simple heuristic-based one, similar to a conservative player, where a face-up card is picked up under a limited set of favorable circumstances. The discard decision is driven by an approach that classifies each card based on its potential to improve the agent's score and its potential to improve the opponent's score. The agent used the Monte Carlo Counterfactual Regret (MCCFR) minimization algorithm to approximate equilibrium end-of-game strategies.

There is conflicting and ambiguous advice for discard and knocking strategies. When determining whether an opponent may be able to use a potential discard, suggested factors to consider include the card's rank, discarded cards of the same rank, and discarded cards in the same suit with nearby ranks (Brown 2019, Kotsckowski 2020B). However, the relative importance of these factors differs between sources. Similarly, some sources claim that the player should usually attempt to end the game as soon as possible (Kotsckowski 2020A). Others stress knocking early in the game and avoiding "weak" knocks later (Brown 2019).

The discard and knocking strategies, learned by the agent, provide insight into resolving these conflicting strategy tips. The agent's learned discard strategy provides the relative importance of a discard's rank and the state of cards that could meld with the discard. MCCFR minimization was used to approximate equilibrium knocking strategies under common variations in Gin Rummy rulesets. These knocking strategies varied significantly between rulesets. This result suggests that the rule configurations are important to consider when formulating an end-of-game strategy.

Background

The Game of Gin Rummy

Gin Rummy is part of the Rummy game family, which consists of hundreds of distinct games. Gin Rummy is believed to have been played publicly for the first time in 1910, when it was introduced by Elwood Baker (Jacoby 1959).

Gin Rummy is a two-player game, played with a standard deck of 52 cards; the goal is to be the first player to reach a score of 100 points (Richard 1945). After shuffling the deck, one player deals 10 cards to each player, and then turns over one card to start a discard pile. The remaining cards are called the stock pile.

Players then take turns drawing a card and discarding a card. At each turn, a player can choose to draw the face-up card from the discard pile, or draw the face-down card from the top of the stock pile. After drawing a card, the player must then choose a card to discard, such that their hand contains 10 cards at all times. During the first round, there is a small variation on the game play. If the first player chooses not to draw the face-up card, then the other player is given the option to draw it. Should either player choose to draw it, they will then discard a card, and play will continue starting from the opposing player.

To win, a player must form one of two types of groupings, called melds. A meld called a run is a sequence of 3 or more consecutively ranked cards in the same suit. The card rankings are ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, jack, queen, and king. For example, a run could consist of the 6 of hearts, 7 of hearts, and 8 of hearts. It is important to note that the player can add more cards of the same suit at either end of the run. An ace can only be used as a low card in a sequence; thus, a sequence cannot be made from a queen, a king, and an ace. The other type of meld, called a set, consists of three or four cards of the same rank. For example, a set could be made from the 7 of hearts, 7 of diamonds, and 7 of spades.

Within a player's hand, the cards that are not part of a meld are referred to as deadwood cards. Each deadwood card has an associated deadwood penalty, determined by its rank. Cards of rank 2 through 10 give a deadwood penalty equal to their rank, while aces give a penalty of 1 and face cards give a penalty of 10. The sum of the deadwood penalties for all the deadwood cards in a player's hand is referred to as the player's deadwood.

There are two ways that a round can end. If the stock pile is reduced to 2 cards, then the game ends in a tie and neither player earns any points. The other way occurs when a player decides to "knock." In order to knock, a player must have 10 or fewer deadwood points. A player indicates that they are knocking by placing their chosen discard face-down on the discard pile, and announcing that they are knocking. Then, they must display their deadwood cards and melds for the opponent to see.

At this point, the opponent also shows their melds. The opponent has the ability to add any of their deadwood cards into a compatible meld in the other player's hand, through a process called laying-off. After choosing cards to lay off, the opponent's deadwood is recounted. If the knocking player's deadwood is still lower than their opponent's, the knocking player wins the round, adding the absolute value of the difference between both players' deadwoods to their total score; however, if the opponent's deadwood is less than or equal to that of the knocking player, then the opponent receives the points instead, plus an additional undercut bonus. If the knocking player's deadwood is equal to zero, then they have gin, in which case they win the round unconditionally. If a player gets gin, then the other player is not allowed to lay off any cards. In addition to receiving points equal to their opponent's deadwood, they will get a gin bonus.

In the version of Gin Rummy popular in the United States, both the gin and undercut bonuses are typically worth 25 points. In the version popular in the United Kingdom, the gin bonus is 20 points and the undercut bonus is 10 points.

Regret Minimization in Games with Imperfect Information

Regret minimization provides a mechanism to find Nash equilibrium strategies for zero-sum, normal-form games (Hart and Mas-Colell 2000). Intuitively, this approach is analogous to a human player thinking back on past actions in a game and wondering "Well, what would have happened if I had acted differently?" This would allow the player to find decisions that they regretted making, and decide what would be a better course of action if a similar decision ever came up in the future.

The Counterfactual Regret (CFR) minimization algorithm extends this approach for imperfect information in extensive form games (Zinkevich and Johanson 2008, Neller and Lanctot 2013). Since players in these games have limited information available to them, multiple game states are indistinguishable from the perspective of a player. A set of indistinguishable states is called an information set, or infoset. During CFR training, the probability of taking an action at a particular infoset is equal to the counterfactual regret of not taking this action when at this infoset earlier in the training, where counterfactual regret is equal to the weighted difference in the expected utility of taking an action at this game node minus the expected utility of reaching this game node. This difference in utility is weighted by the probability that the other player would play to the node, and therefore, the regret is counterfactual in the sense that we assume that the current player plays to this game node with probability of one. The average of the strategies used over the training will converge to a Nash equilibrium.

In Monte Carlo CFR (MCCFR), rather than exploring the entire game tree during training, each training run samples just a portion of the game tree (Lanctot et al. 2009). For example, when the player who is training is at a decision point, they can choose just one action from the available set of actions. The probability for selecting each potential action is set to some ε, divided by the number of actions, plus the probability that they would usually take that action weighted by $1 - \varepsilon$. That is, with probability ε, they will sample uniformly from these actions, and with probability $1 - \varepsilon$ they will choose based on their regret matching strategy. The authors note that empirical studies found that a value of 0.6 worked well for ε. The MCCFR approach ensures that the sampled counterfactual regret will be equal to the actual counterfactual regret in expectation (Lanctot et al. 2009).

One of the challenges of applying CFR to large extensive form games arises from the number of states that must be captured. Rather than distinguishing between infosets based on all of the game states available to a player, typically a smaller set of attributes is used to characterize the infoset. However, the choice of attributes that are used in the abstraction can have a significant influence on the performance of the system (Waugh *et al.* 2009). In this work, we limit the attributes in the abstraction to ones that would be commonly used by human players.

Gin Rummy Agent Design

The Gin Rummy agent used a heuristic approach for draw and discard decisions, and a counterfactual regret minimization strategy for the knocking decision. The drawing strategy consists of a simple approach that draws the face-up card only if it would create a significant improvement in the deadwood points. The discard strategy uses estimates of the utility of potential discards to both the agent and the opponent. The utility for the agent is estimated based on an approach that looks ahead three turns in the game tree, while the utility for the opponent is estimated based on the cards known to be in the opponent's hand and the cards known to be unavailable to the opponent. The agent was written in Java, using no external libraries, and is available at https://github.com/bendstein/Gin-Rummy.

The knocking strategy is derived using Monte Carlo Counterfactual Regret (MCCFR) minimization. MCCFR was applied only to knocking decisions for computational reasons. In order to converge to an equilibrium strategy, MCCFR must sample from many nodes in a large number of possible game trees. For discard and draw decision nodes, at least two branches of the tree arise at each of these decision nodes. Consequently, the game tree will grow exponentially with the number of these decisions considered. However, a decision to knock will end the game. Thus, including just knocking decisions ensures that the game tree grows only linearly in the number of knocking decision points.

Draw Strategy

To decide whether to draw the face-up card or to draw from the stock pile, the agent calculates how the deadwood in the hand would change after drawing the face-up card and discarding the card that minimizes the resulting deadwood. If the face-up card would be a part of a meld and the resulting deadwood would improve, the face-up card is drawn. Otherwise, the agent draws a card from the stock pile.

This approach is consistent with advice given to new players, such as "don't draw from the discard pile unless it completes a run" (Brown 2019, Game Colony 2011), as drawing the face-up card has two disadvantages (Game Colony 2011). First, this action can indicate to the opponent what cards the agent is collecting. This information can help the opponent avoid future discards that might be added to a meld, as well as allow them to collect cards that can be laid off on the agent's melds, in the case that the agent knocks without gin. Second, drawing from the stock pile has a chance of enabling the agent to draw a card that is useful to the opponent. Since the opponent has discarded the face-up card, it is likely not of use to them.

Note that the full discard strategy is not run when evaluating the deadwood improvement from the face-up card. The full discard strategy involves an expensive search of the game tree, so instead, an upper bound on the deadwood improvement is established by assuming that the player will discard the most expensive deadwood card.

Discard Strategy

The discard strategy considers, as potential discards, all cards that are not currently melded. If there are multiple ways to create melds that minimize the deadwood in the hand, a card is considered a potential discard if they do not appear in a meld in at least one of the configurations. Each potential discard is assigned a score, and the potential discard with the lowest score is chosen for discard. The score, $s(c)$, for a potential discard c, is given by the formula:

$$s(c) = d_i(c) + u_{-i}(c) \quad (1)$$

where:

- $d_i(c)$ is the estimate of the expected value of the deadwood that will remain in the agent's hand if we discard card c, calculated as described below.
- $u_{-i}(c)$ is the estimate of the utility of card c for the opponent, also described below.

1. Estimating Deadwood in Agent's Hand

The estimated utility for the current player is based on the deadwood that is possible within the next three turns. It is given by the formula:

$$u_i(c) = \sum_{t=0}^{T_{max}} w(t)\, d_i(c,t) \quad (2)$$

where:

- $w(t)$ is a weight based on the probability that the game will end after t more turns, with adjustments based on performance in empirical trials.
- $d_i(c,t)$ is the estimate of the deadwood for the agent if we discard card c after t more turns
- T_{max} is the maximum turn considered. Setting T_{max} to 8 yielded the best performance in empirical tests.

The value for deadwood in the agent's hand if the game ends after this turn and we discard a card c, $d_i(c,0)$, can be calculated exactly by determining the deadwood remaining if card c is discarded.

The remaining values for this expression, though, are estimated. For example, in order to look one turn ahead, the agent removes the card c from the hand, and then adds, one at a time, all cards that could remain in the stock pile, based on the cards the agent has seen so far. The agent then goes ahead and tries to make melds from the 11 cards. Rather than trying all possible discards, the deadwood for a particular 11-card hand is calculated as the sum of the deadwood cards from the 11 cards, minus the most expensive deadwood card. Finally, the average of the lowest 58.8% of the deadwood values is used for the value $d_i(c, 1)$. This proportion of cards chosen is equal to 1/1.7, a parameter that worked well based on empirical trials.

A similar procedure is used to estimate the expected deadwood after two turns, $d_i(c, 2)$. All possible combinations of two cards are added to the hand, and the agent tries to make melds from the 12 cards. The expected deadwood is estimated by calculating the sum of the deadwood cards from the 12 cards, minus the most expensive deadwood cards. The average of the lowest 55.5% of the deadwood values is used as our estimate of $d_i(c, 2)$. Note that this approach deviates from the way the agent will actually play. The agent will not be able to see all 12 cards before deciding which card to discard. Moreover, it is possible that melds based on the 12 cards will not be possible to make after discarding cards, and therefore, the estimate will be overly optimistic. However, the agent is only allowed to discard one card to counter this. This approach was chosen because it was significantly faster than an approach that would evaluate the hand after each possible discard over the course of two turns. A similar approach was used to calculate $d_i(c, 3)$, in which all combinations of three cards are added to the hand, with one card discarded, and the estimate of the deadwood is based on the lowest 55.5% of values.

In order to extrapolate beyond 3 turns, the agent assumes that for a given discard in this turn, the expected value of the deadwood will continue to drop linearly. We have calculated $d_i(c, 2)$, the expected deadwood after two turns if we discard card c, and $d_i(c, 3)$, the expected deadwood after three turns. To calculate the expected deadwood if we discard card c after t turns, $t > 3$, we use formula :

$$d_i(c,t) = max(0, d_i(c,3) - \left(d_i(c,2) - d_i(c,3)\right)(t-3))$$

2. Estimating Deadwood Improvement in Opponent's Hand

The agent observes the face-up cards that the opponent has picked up and the cards that are not available to the opponent, either because they have been discarded, or because they are in the agent's hand. Gin Rummy strategy advice suggests that when evaluating the potential for a card to be useful to the opponent, a player should consider the number of cards of the same rank, cards of the same suit with an adjacent rank, or cards of the same suit with a rank that is within two of another card (Game Colony 2011). In addition, some ranks are more likely to appear in sequences than others. For example, a king can only be a part of one three-card sequence, but a seven can be a part of three.

Therefore, a range of attributes are used to estimate the expected improvement in the opponent's deadwood if we discard card c, $u_{-i}(c)$:

$$u_{-i}(c) = f(c_r, r_i, r_{-i}, r_0, s^{+1}, s^{+2}, s^{-1}, s^{-2}) \qquad (4)$$

where:

- c_r is the rank of card c.
- r_i is the number of cards of the same rank that are in the agent's hand.
- r_{-i} is the number of cards of the same rank that are known to be held by the opponent.
- r_0 it the number of cards of the same rank that are in the discard pile.
- s^{+1} indicates the state of the higher, 1-adjacent suited card, i.e., the card of the same suit with rank that is one higher than the discard. This state is chosen from the following values: unknown, discarded, in agent's hand, or in opponent's hand.
- s^{+2} indicates the state of the higher, 2-adjacent suited card, i.e., the card of the same suit with rank that is two higher than the discard. The same set of states apply as in s^{+1}.
- s^{-1} indicates the state of the lower, 1-adjacent suited card, i.e., the card of the same suit with rank that is one lower than the discard. The same set of states apply as in s^{+1}.
- s^{-2} indicates the state of the lower, 2-adjacent suited card, i.e., the card of the same suit with rank that is two lower than the discard. The same set of states apply as in s^{+1}.

To determine the values returned by the function f, a large lookup table was created using these attributes. The values in the lookup table are derived by sampling 1.4 million hands. For each hand, the potential discard was evaluated to see how much the opponent's hand would improve. The lookup table contains the average improvement in deadwood for a discard with the given attributes.

Knocking Strategy

The knocking strategy differs from the strategies used at other decision points. When it is possible to knock, the agent uses a strategy that was determined by the MCCFR Minimization algorithm. The training consisted of playing individual hands while sampling a single action for both the player being trained and the opponent. The expected value of the hand for the current player was equal either to the points earned by the training player in the hand, or the negation of the points earned by the opponent.

The abstraction used to characterize the information set uses common attributes suggested by Gin Rummy strategy guides (Brown 2019, Kotsckowski 2020). These characteristics include the amount of deadwood in the player's hand, the stage of the game, and the state of the opponent's hand.

These characteristics are captured in part with the three attributes that the agent used to characterize an infoset state: the agent's deadwood points, the number of cards remaining in the stock pile, and the lowest card that the opponent has discarded. The first attribute captures whether a knock is considered "weak" (Brown 2019). The second attribute captures the stage of the game, whether it is an "early" knock or "late-stage" knock (Kotsckowski 2020). The lowest deadwood card is intended to capture the quality of the opponent's hand, with the assumption that the opponent will start by discarding their high value deadwood cards. When an opponent is discarding low value cards, the implication is that their total deadwood is low.

As noted in the background section, the design of infoset abstraction can have a significant influence on the performance of a CFR strategy (Waugh *et al.* 2009). By choosing attributes that are suggested by Gin Rummy strategy guides, we are attempting to avoid poor abstraction approaches. Moreover, various abstraction approaches were chosen before settling on this approach as the best.

Results

Agent Performance

To evaluate the agent's performance, we played 100,000 games between it and the baseline agent provided with the EAAI competition base code (Neller 2019). The agent significantly outperformed the baseline agent, winning more than three quarters of the games that were played.

The baseline agent follows a naïve strategy (Neller 2019). It will draw the face-up card if, and only if, it can be added to a meld in its hand. Additionally, its discard strategy consists of choosing the unmelded card with the highest deadwood value and discarding it, and it will always knock as soon as its deadwood points are 10 or lower.

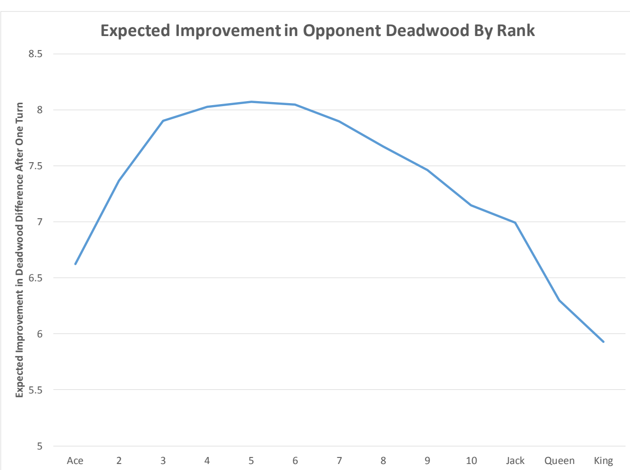

Figure 1. Expected Improvement in Opponent's Deadwood by Discard Rank.

The agent won 75.8% of the games against the baseline agent. Moreover, the margin of victory is indicative of the agent's strong performance. The agent outscored the baseline agent by 42.4 points on average, with a median difference in scores of 60 points. In 17% of the games, the agent won the game by 100 or more points.

Discard Strategy

The discard strategy suggests certain heuristics that could be used by human players to evaluate the utility of discards to their opponent. These heuristics select and evaluate potential discards based on a card's rank, discarded cards, and known cards in the opponent's hand.

Figure 1 plots the expected improvement in the opponent's deadwood after a given discard. This plot is restricted to cases where the agent knows nothing of the status of cards that could participate in three-card melds with the discard. In other words, the agent holds no other cards of the same rank as the discard, and the agent has not seen any other cards of the same rank nor cards of the same suit that have a rank that is one or two away from the potential discard.

As shown in this figure, with no information about nearby cards, the player is best off discarding a king, followed closely by a queen. These results are consistent with the fact that the king is less likely to be useful to the opponent, since it can participate in fewer melds than the queen. A jack and a ten have similar expected improvement values. Both can participate in the same number of distinct 3-card melds. Advice to new players includes the warning to be wary about discarding sevens, since they can meld in the most possible ways (Game Colony 2011). Interestingly, sevens are the highest rank for which the improvement in deadwood points is less for the agent than the expected improvement in deadwood points for the opponent. However, the expected improvement in the opponent's deadwood points is higher for fours through sixes. Below four, this deadwood improvement decreases as the value of cards that would be melded with the discard is reduced.

Figure 2 shows the expected deadwood improvement in the opponent's hand, if exactly one nearby card has been discarded, and no other information is known about cards that participate in a three-card meld with the discard. In order to limit consideration to the same set of cards in all cases, only cards with ranks of four through jack are considered. Kings, for example, are excluded because a card of the same suit with a higher rank does not exist.

Some instruction guides suggest that in situations like the ones in this figure, it is preferable to discard a card where a suited card of adjacent, or 2-adjacent, rank has been discarded (Game Colony 2011). Others suggest that these nearby cards be considered equally. However, the analysis for the agent suggests that the cards are not equal, and moreover, rank is more important. If there has been a discard of the same rank it is less likely to be useful to the opponent than if there has been a discard of an adjacent card of the same suit.

Knocking Strategy

To approximate the equilibrium knocking strategy, the MCCFR Minimization algorithm was run several times, varying the gin and undercut bonuses between each other. Each experiment ran the MCCFR minimization algorithm over 12.5 million games. In one of the optimization runs, the gin and undercut bonuses were set to 25 points each. In the other, the gin bonus was 20 points, and the undercut bonus was 10 points.

In order to speed up the training process, the drawing strategy was the one described above, but the discard strategy was limited to looking one turn ahead. The resulting knocking strategy, then, represents an approximation of the equilibrium strategy, assuming that players limit themselves to these drawing and discarding approaches.

As noted earlier, the infoset abstraction can influence the approximation of the regret minimization strategy. Of the three attributes used, deadwood remaining in the agent's hand, the size of the stock pile, and the lowest opponent discard, the deadwood remaining in the agent's hand had the largest influence on when an agent knocks.

As shown in figure 3, under either bonus configuration, the strategies indicate that a player should not knock most of the time, with the probability of knocking decreasing as the amount of deadwood decreases. In general, the player avoids knocking in order to either receive the gin bonus by melding all of their cards, or the undercut bonus should the opponent knock. However, under the smaller undercut bonus of 10, a player knocks more often. Averaged across the deadwood amounts, the player is approximately 11 times more likely to knock under this smaller undercut penalty.

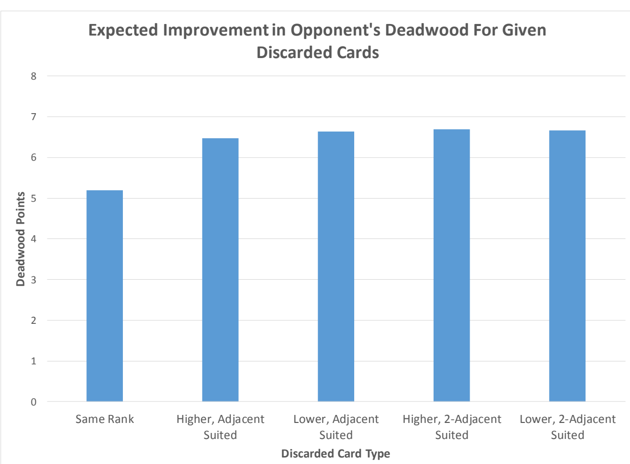

Figure 2. Expected improvement in opponent's deadwood, given different discarded card types.

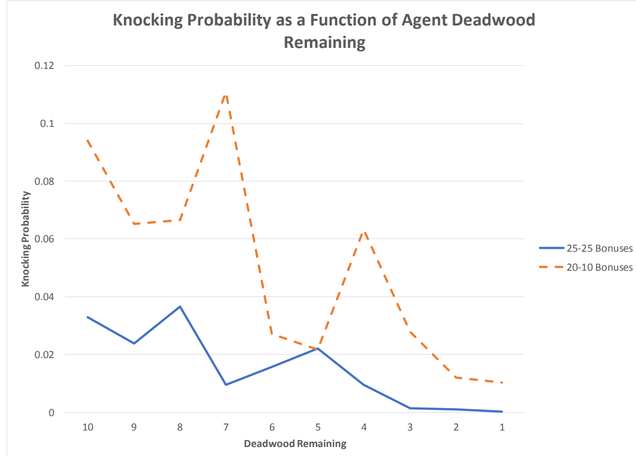

Figure 3. Knocking probability as a function of agent deadwood, under the two distinct bonus configurations.

The second most important attribute in determining whether to knock, after deadwood remaining, is the size of the stock pile. As shown in figures 4 and 5, regardless of the bonus amounts, an agent is likely to knock, when it is possible, early in the game with larger amounts of deadwood. As the size of the stock pile decreases, an agent becomes less likely to knock. Moreover, even in the early stages of the game, if an agent has smaller deadwood, they are less likely to knock. In these cases, avoiding knocking indicates that it is advantageous to try to obtain either the gin or undercut bonus.

Figures 4 and 5 also point to significant differences in strategies under the two bonus configurations. Under the 25-point bonuses, once 8 cards have been drawn from the stock pile, the knocking probability drops below 0.08. However, under the 20-point/10-point configuration, the knocking probability does not drop below this threshold until the stock pile has been reduced by 17 cards.

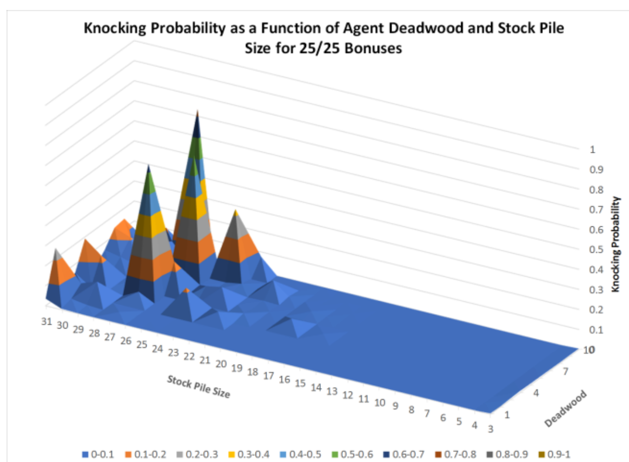

Figure 4. Knocking probability as a function of agent deadwood and stock pile size for 25-point bonuses.

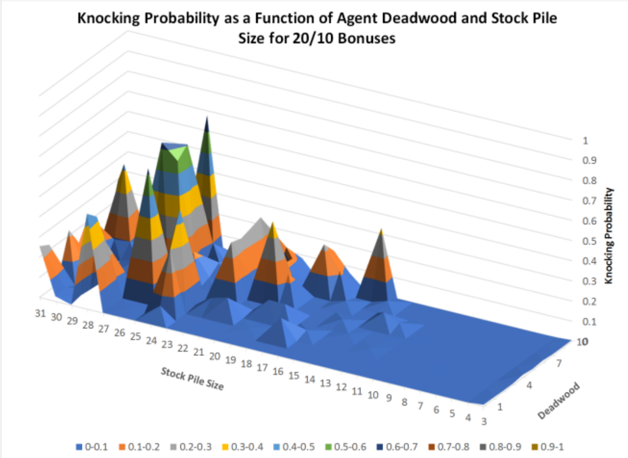

Figure 5. Knocking probability as a function of agent deadwood and the stock pile size for 20-point gin bonus and 10-point undercut bonus.

Study Limitations

The applicability of the agent's strategy to discard and knocking strategies in human play is conditional on limitations inherent in this study's approach. These limitations are inherent in the high variance nature of Gin Rummy and constraints imposed by the strategy generation process.

Outcomes in a Gin Rummy game have high variance. As a result, a very large number of games are required to accurately estimate the expected value of discards to an opponent and the regret for each infoset of the knocking strategy. While the results presented are based on millions of games, the lack of smoothness in the strategies, evident in figures 4 and 5, indicate that more games may have been needed for higher accuracy in strategy estimation.

These figures may also be indicative of important missing attributes in the infoset abstraction approach. Notably, the non-monotonicity in these figures suggests the absence of an important attribute in the abstraction. For example, this behavior may be explained by the abstraction's omission of how likely an agent was to achieve gin.

In addition to limitations due to the high variance nature of Gin Rummy and infoset abstraction, the analysis assumes that opponents will play strategies similar to the agents. An equilibrium strategy, for example, is not an optimal strategy against a fixed opponent, but rather it is a strategy that is not exploitable by opponents. The equilibrium strategy approximated in this study assumes that opponents will play drawing and discard strategies that are identical to the agents. While these strategies seem consistent with typical heuristics used by human players, this assumption, nonetheless, presents a limitation to the applicability of these results.

The analysis of the expected value of a discard to an opponent makes similar assumptions. The particular drawing and discard strategy can affect the estimates of the expected value of these potential discards. If, for example, under the agent's draw strategy, they are less likely than a typical opponent to collect kings, the analysis would underestimate the expected value of a king to an opponent. Under the assumption that the drawing and discard strategies are in line with a typical player, this limitation should not significantly affect the conclusions.

In addition, the evaluation of the agent's performance is limited by the number of agents available for comparison. The baseline agent used represents a naïve strategy, and it is not clear that the strong performance against the baseline would translate to a strong performance against an arbitrary opponent.

Conclusions

The agent described in this paper used a counterfactual regret minimization algorithm for the development of a knocking strategy, and heuristic approaches for the drawing and discarding decisions. This approach outperformed the baseline agent from the EAAI competition, winning more than 75% of the games. Moreover, the development of this agent provided insight into strategy tips for human players of the game. The tips are applicable to both knocking and discard strategies.

An equilibrium knocking strategy was approximated using a Monte Carlo Counterfactual Regret minimization algorithm. This strategy suggests that the probability of knocking depends on the rule configuration, the stage of the game and the amount of deadwood remaining. In games with 25-point gin and undercut bonuses, the strategy was to knock only early in the game. In games with 20-point gin bonuses and 10-point undercut bonuses, the strategy allowed for knocking in the middle stages of the game as well. In both bonus configurations, the agent is more likely to knock with higher deadwood points than with lower deadwood points.

The discard strategy was informed by sampling the expected value of a discard based on rank, nearby discarded cards, and known opponent cards. This analysis indicated that if no information about nearby cards is available, kings are the best cards to discard. Moreover, discarding a card with a rank of seven or below yields an expected improvement in the opponent's hand that exceeds the deadwood value of the discarded card. In addition, when a card which can meld with a potential discard is unavailable, the type of that card is important in calculating the expected benefit of a potential discard to an opponent. If a card of the same rank has been discarded, the card is less likely to be useful to the opponent than if there were a discarded card with the same suit and an adjacent rank.

References

Brown, S. 2019. 7 Simple Strategy Tips for Gin Rummy. DotDash Publishing. www.thesprucecrafts.com/gin-rummy-card-game-strategy-and-tips-412364. Accessed 8/18/2020.Game Colony. 2011. Gin Rummy – Game Strategy and Math. www.gamecolony.com/gin_rummy_game_strategy.shtml. Accessed 8/18/2020.

Hart, S. and Mas-Colell, A. 2000. A Simple Adaptive Procedure Leading to Correlated Equibrium. *Econometrica* 68(5): 1127-50.

Jacoby, O. 1959. How to win at gin rummy. New York: Henry Hold and Company.

Kotsckowski, M. 2020A. When to Knock, Play for Gin, or Underknock. Rummy Talk. rummytalk.com/when-to-knock-play-for-gin-or-underknock/. Accessed 8/18/2020.

Kotsckowski, M. 2020B. The Basics of the Safety Factor Point Count System. rummytalk.com/the-basics-of-the-safety-factor-point-count-system/. Accessed 8/18/2020.

Lanctot, M., Waugh, K., Zinkevich, M., and Bowling, M. 2009. Monte Carlo Sampling for Regret Minimization in Extensive Games. Advances in Neural Information Processing Systems (22): 1078-86.

Neller, T. W. 2019. The Gin Rummy EAAI Undergraduate

Research Challenge. cs.gettysburg.edu/~tneller/games/ginrummy/eaai/. Accessed 12/1/2020.

Neller, T.W. and Lanctot, M. 2013. An introduction to counterfactual regret minimization. In Proceedings of Model AI Assignments, The Fourth Symposium on Educational Advances in Artificial Intelligence (EAAI-2013).

Waugh, K., Schnizlein, D., Bowling, M., Szafron, D. 2009. Abstraction Pathologies in Extensive Games. In Proceedings of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009), 781-8.

Zinkevich, M., Johanson, M., Bowling, M., and Piccione, C. 2008. Regret Minimization in Games with Incomplete Information. *Advances in Neural Information Processing Systems* (20): 1729-36.