

Rummy OpenAI Gym: Scope Document

Luke Bhan

November 1, 2021

1 Introduction

The goal of this project is to implement the game rummy for an OpenAI gym and attempt to train agents to play better at this game. It is a partial information game and thus may be challenging for agents to defeat. In this document, we will define both the game rules as well as the initial design for the gym.

2 Game Rules

Our form of rummy resonates closer with the popular game Rummy 500 in which players attempt to rid their cards while accumulating large scores. The game will be played with the following rules.

2.1 Initial Setup

The initial setup of the game will consist of n players in which each player is dealt an initial hand of 7 cards that the other players cannot see. Furthermore, the remaining cards will be placed faced down in a deck such that no players can see the cards. The first card will be flipped face up and placed next to the deck for all players to see. This initial card will be the basis for what we denote our *discard deck* (although it is not fully a discard deck). From here, the player to go first will be chosen randomly.

2.2 Turns

Each player will take turns in clockwise formation until one player has 0 cards. On a turn, a player has the option to grab one card from the facedown deck, or they can grab cards from the discard deck (See section 2.2). Once they have obtained their cards, they can choose to place cards in front of them in two forms. The first form consists of three or more cards of the same number. The second form consists of three numbers of the same suite in continuous order with ace being high. For example, 3, 4, 5 of diamonds is a legal placement while 3, 4, 6 of diamonds is not. Furthermore, 3 of diamonds, 4 of spades, and 5 of diamonds is not a legal play. These initial sets must consist of at least 3 cards each. In addition to these sets, players are able to place cards in front of them based on the other players sets. For example, if player 1 played 3, 4, 5 of diamonds. Player 2 can play the 6 of diamonds and player 3 can play the 7 of diamonds on their respective turns. When playing off another opponent's set, a player does *not* need to play at least 3 cards. After, a player has placed all their cards, they must contribute one card from their hand to the discard pile.

2.3 Discard Pile

The discard pile consists of a set of cards in which players must contribute to at the end of their turn. Furthermore, the order of the cards placed in the discard pile are maintained such that every player can see the order in which discards are played. Now, at the beginning of each turn, a player has the option to pick cards up from the discard pile. If a player chooses to do this, they must play a set for the card in which they pick up and must place all cards that were discarded after this initial card into their hand. For example, if players have discarded 4Spades, 5Diamonds, and JClubs in that order, and the player currently has the 4Hearts and 4Diamonds in their hand. They can pick up the 4Spades and play a triple set of 4s under the condition they also add the 5Diamonds and JClubs into their hand.

2.4 Ending the Game

The round ends when a player runs out of cards in their hand by playing sets. Note, a player must discard one card at the end of their turn and as such if a player has 3 cards and they form a set, they cannot play this set as they would not have enough cards to discard. As such they must discard one of their cards from their set. Furthermore, all the cards left in players hands when one player ends the game will count negatively towards the other players scores. (See Section 2.5)

2.5 Scoring

Scoring is as follows. For each card played with values between 2 and 9, the points for those cards are at face value. So a 9 is worth 9 points and a 5 is worth 5 points. Furthermore, any card that is 10, J, Q, K is worth 10 points. Lastly, an Ace can be played as high or low and is worth either 1 point or 15 points respectively. A set of 3 Aces is worth 45 points. At the end of a round, a players score is determined by counting the points they obtained by playing cards off of other players as well as their own sets minus the cards left in their hand. When an ace is left in hand, it is minus 15.

2.6 Nuanced Rules

We introduce a list of nuanced rules here specific to our game:

- A player must discard on their turn and cannot go out(end the game) without discarding
- Aces can be played as high and as low. They are 1 point as low and 15 points as high. They count as -15 when left in hand if the game ends.
- A player must grab all cards before a given card if they want to pull from the discard pile. Furthermore, the card they grab must be played instantly in the form of a set. It cannot be played directly off another players card. It can, however, be combined as such. Say a player has the 4, 5, 6 of diamonds down. The next player can pick up from the 8 of diamonds and play the 7 and 8 of diamonds legally. However, they cannot pick up the 7 of diamonds and play that by itself.
- Sets can exist in the discard pile. However, they cannot be played as a last card unless the player can add to the set with a card from their hand

2.7 Example

An example of Rummy with three players:

The initial hands are delt.

- Player 1: 2Dia, 3Dia, 4Dia, AceSpades, 4Clubs, QHearts, 5Clubs.
- Player 2: 4Spades, 7Spades, 8Spades, 2Hearts, KHearts, JClubs, AceDia
- Player 3: 5Spades, 5Hearts, 5Diamonds, 3Clubs, 9Hearts, 6Hearts, AceClubs

Player 1 goes first. The discard deck has a 2 of clubs as its first card.

- Player 1 cannot use the 2 of clubs and thus draws randomly. They obtained the AceHearts. From here, they play their 2, 3, 4 of diamonds as a set. They choose to discard the 5Clubs.

This discard deck contains 2 clubs, 5 clubs in that order.

- Player 2 cannot use either to make a set so they draw randomly. They obtain the KClubs. They cannot play sets. They choose to discard the 8Spades.

The discard deck contains 2 clubs, 5 clubs, 8 spades in that order.

- Player 3 can use the 2 of clubs to play Ace, 2, 3 of clubs with the ace as low and so they pickup the discard deck. They now have the 5 of clubs and 8 of Spades in thier hand as well. They have to play the Ace, 2, 3 of clubs as as a set. Furthermore, they choose to play all of their 5s leaving them with the 9Hearts, 6Hearts, and 8Spades. They discard the 6Hearts and move on.

At this point the game board looks like:

- Discard pile has a single card - the 6Hearts.
- Player 1 has played a set of 2, 3, 4 of Diamonds (9 pts) and has 4 cards in their hand.
- Player 2 has not played anything and has 7 cards in their hand.
- Player 3 has played a set of A, 2, 3 of Clubs (6 pts) and 5 of Diamonds, Spades, Hearts, and Clubs (20 pts) and has 2 cards in their hand

It is now player 1's turn.

- Player 1 cannot use the 6Hearts. They draw randomly and obtain a 2 of Spades. They cannot play any sets. However, they can play off the A, 2, 3 of Clubs set using their 4 of clubs. They play this card and obtain 4 pts for it. They choose to discard their 2 of spades.

We continue this sequence until one player has no cards left. Then, we count the scoring for that round and that person is the winner.

3 Modeling as an OpenAI Gym

We will now model this game above as an OpenAI gym. It will contain a series of classes each described here:

3.1 Deck Class

This class will represent any deck of cards. It consists of a dictionary that will hold keys in the form of NumberSuite with 10, Jack, Queen, King, Ace represented as T,J,Q,K,A respectively. An example is 2Clubs or ASpades. The dictionary is chosen as the deletion times are slightly faster than arrays in Python. We will overload the contains, repr, and len special methods for the deck.

3.2 Hand Class

This will hold each players hands. It will inherit from a deck class and behave similarly except with identifiers for sets and runs. We will have a method returning the 3-tuple identifying the number of sets, type of set, and points of each set for a specific hand.

3.3 Player Class

This will represent a player class for the game. It will contain a hand and provide the action and observation spaces for each player. The observation space will consist of a hand, the discard deck, and the three-tuple aformentioned in 3.2. The action space will consist of three parts. Picking up cards, Playing sets/cards, and discarding.

3.4 Game Class

This will initialize a game with n players and randomly choose who goes first. It will calculate the scores of each player and identify when the game ends.

4 Reinforcement Learning Design

4.1 Reward Function

The reward function will be a simple linear combination of both minimizing cards and maximizing score for an individual player.

$$R = -\omega_n * (\text{Number Cards}) + \omega_p * (\text{Points played}) + \omega_w * (1 \text{ if Win} | 0 \text{ if lose}) \quad (1)$$

where ω are weight parameters chosen by the game designer.