

Yahtzee AI

based on searching

Rui Tang, Sue Wang | Capstone Term I | 2020-10-20

# Abstract

Yahtzee is a popular dice game in North America. Many people regard Yahtzee as a lucky game. Absolutely it is, but with strategy and optimization we can get a higher score expect value. In this essay we would see how to use searching to find the maximum expected value and select the best option in each step.

## Introduction of Yahtzee

* **Yahtzee** is a [dice game](https://en.wikipedia.org/wiki/Dice_game) made by [Milton Bradley](https://en.wikipedia.org/wiki/Milton_Bradley_Company) (now owned by [Hasbro](https://en.wikipedia.org/wiki/Hasbro)), which was first marketed as Yatzie by the National Association Service of [Toledo, Ohio](https://en.wikipedia.org/wiki/Toledo,_Ohio), in the early 1940s. It was marketed under the name of Yahtzee by game entrepreneur [Edwin S. Lowe](https://en.wikipedia.org/wiki/Edwin_S._Lowe) in 1956. The game is a development of earlier dice games such as [Poker Dice](https://en.wikipedia.org/wiki/Poker_Dice), [Yacht](https://en.wikipedia.org/wiki/Yacht_(dice_game)) and [Generala](https://en.wikipedia.org/wiki/Generala" \o "Generala). It is also similar to [Yatzy](https://en.wikipedia.org/wiki/Yatzy), which is popular in [Scandinavia](https://en.wikipedia.org/wiki/Scandinavia).
* The objective of the game is to score points by rolling five [dice](https://en.wikipedia.org/wiki/Dice) to make certain combinations. The dice can be rolled up to three times in a turn to try to make various scoring combinations and dice must remain in the box. A game consists of thirteen rounds. After each round the player chooses which scoring category is to be used for that round. Once a category has been used in the game, it cannot be used again. The scoring categories have varying point values, some of which are fixed values and others for which the score depends on the value of the dice. A Yahtzee is five-of-a-kind and scores 50 points, the highest of any category. The winner is the player who scores the most points.
* Yahtzee was marketed by the E.S. Lowe Company from 1956 until 1973. In 1973, the [Milton Bradley Company](https://en.wikipedia.org/wiki/Milton_Bradley_Company) purchased the E.S. Lowe Company and assumed the rights to produce and sell Yahtzee. During Lowe's ownership over 40 million Yahtzee games were sold worldwide. According to current owner [Hasbro](https://en.wikipedia.org/wiki/Hasbro), 50 million Yahtzee games are sold each year. A classic edition is currently being marketed by [Winning Moves](https://en.wikipedia.org/wiki/Winning_Moves).[[1]](#footnote-1)

## Rules

A classic Yahtzee game has 13 turns and 13 types of scores plus a bonus score. Each type of score must and only be selected once. In each turn, players can roll five dices and reroll up to two times. When rerolling, the player can keep any of dices at that moment and reroll the others. If players want to write the score onto the score board (or the reroll times used out), they should select one of the score type and calculate the corresponding score then record it. If the numerical score is higher than 63, a 35-score bonus will be earned.

The detail rules show following:

|  |  |  |  |
| --- | --- | --- | --- |
| **Categories** | **Descriptions** | **Scores** | **Examples** |
| **Aces** | Any combination | The sum of dice with the number 1 | scores 3 |
| **Twos** | Any combination | The sum of dice with the number 2 | scores 6 |
| **Threes** | Any combination | The sum of dice with the number 3 | scores 12 |
| **Fours** | Any combination | The sum of dice with the number 4 | scores 8 |
| **Fives** | Any combination | The sum of dice with the number 5 | scores 5 |
| **Sixes** | Any combination | The sum of dice with the number 6 | scores 18 |
| **Bonus** | If the sum of all numerical scores greater than or equals 63 | 35 |  |
| **Three Of A Kind** | At least three dice the same | Sum of all dice | scores 17 |
| **Four Of A Kind** | At least four dice the same | Sum of all dice | scores 24 |
| **Full House** | Three of one number and two of another | 25 | scores 25 |
| **Small Straight** | Four sequential dice (1-2-3-4, 2-3-4-5, or 3-4-5-6) | 30 | scores 30 |
| **Large Straight** | Five sequential dice (1-2-3-4-5 or 2-3-4-5-6) | 40 | scores 40 |
| **Yahtzee** | All five dice the same | 50 | scores 50 |
| **Chance** | Any combination | Sum of all dice | scores 13 |

## Modeling

By observing the game process, we can find that in any situation, the score available for the player’s remaining blanks depends only on which grids have been filled, and has no relationship with the scores in these filled grids (except for bonus points). Thus, we can extract this information from game states: which score types have been filled, the sum of numerical scores, the reroll times and current dices.

## state space

Before we implement the model, we need to estimate the state space, and judge whether it could be stored in storage.

* The space of which scores have been used: We have 13 types of score, each one has two status: filled or not filled. The total space is
* The range of numerical value: It could be from 0 to 126. Luckily, we only care about the values less than or equals 63. The space is 64.
* Reroll times: only three options: 0, 1, 2. The space is 3.
* The space of dices: There are 5 dices, each dice can be 1-6, and can be blank, for which is expected to reroll. Therefore, the total space is =16807

Based on Cartesian product, the total space is =26435125248, approximately 26 Trillion. If each state uses 4 bytes, the total memory usage will be 92 GB, which is too huge.

## Compress and Optimization

In the above section we can see dices use too much space, but by observing we can find some dices with different order were regard as different like (1, 2, 3, 4, 5), (5, 4, 3, 2, 1) but they are same in different combination. By a simple process we can prove the total number of combinations for N dices with M sides is C (M+N-1, M-1). Regarding the dice has 7 sides (the extra for blank), we can easily calculate for 5 dices, the number is C (11, 6) =426. After this process, indexing these dices to number 0 to 425 is possible.

After compressing, the state space will be =670040064, which is much smaller, and ideally the total memory usage will be 2.68 GB.

For computers’ hardware, it is easy to do bit calculation but harder to do divide or mod operation. Also, if the state representation method is irrelevant, the program will be also hard to maintain and debug. For these reasons, we align these status codes by bit, and the final memory usage will be 4 GB.

|  |  |  |
| --- | --- | --- |
| Type | Bits | Max range |
| Dices | 9 | 512 |
| Filled state | 13 | 8192 |
| Reroll times | 2 | 4 |
| Numerical sum | 6 | 64 |
| Total | 30 | 1073741824 |

## Algorithm and State transition

Based on the previous analysis, we can calculate the expected value which can be gained in remain steps. Firstly, we need to define border states.

### Expected value for end game: 0, because we can do nothing

* For blank dices: E=Average ({E(p)| p in all dices possibly})
* For each step needs to decide: E = max(E(next) + score in this turn| next in all possible action))

## Algorithm Impletement

See the C++ project in attachment

## Data Analysis

See the Jupyter Notebook.

## TEAMWORK

### Rui Tang: Algorithm design, code, and report

### Sue Wang: Data cleaning, analysis, and visualization

Resources:

<https://github.com/goagain/Yahtzee-AI>

1. Wikipedia <https://en.wikipedia.org/wiki/Yahtzee> [↑](#footnote-ref-1)