A Guide to Optimization and Optimal Play in Greed

...via markov decision processes and dynamic programming

Chance Addis

Contents

Contents (ii)

Contents (iii)

1 Abstract

The game of Greed is a probabilistic two-player game. This paper aims to derive an optimal way to play the game - based on reasonable heuristic metrics - as well as as determine the optimal chance of success, as determined by out heuristic.

In other word, this paper aims to make you the best greed player, so continue on and lord your superior greed strategy over your enemies.

2 How Do You Play Greed Again?

Game Environment & Turns

Greed is a game that could be described as "kinda like blackjack with dice".

Game parameters:
- M: max score before going bust.
- s: number of sides on each die.

Starts at $s_p = 0$ and $s_0 = 0$ respectively.

Each turn, the player up to roll will choose some $n \in \mathbb{N}^0$ number of dice to roll. Then the sum of those dice will be added to their score. Players will go back and forth like this until one of two terminal states is reached.

How To Win

1. In the first option, a player's score goes over M, i.e. and they go bust. In this case, they lose, so we'll say that their rating—the heuristic measure that we use to measure how good a position is—is 0, and thus their opponents rating is 1.

2. In the second option, a player can choose to roll 0 dice. This will initiate the last turn. The other player will then have one more chance to roll. The winner is the player who has the higher score that is not over M. In the case of a tie, rating is $\frac{1}{2}$ for each player.

3 Mathematical Framework

Consider two types of states: terminal states, which denote the last turn of the game, and normal turns, representing all other states.

Normal and terminal states are conceptualized as 2 by 2 arrays with dimension $M + 1$ by $M + 1$, representing all possible states for player and opponent.

The x-axis designates the player presently up to roll, while the y-axis designates the player who has just rolled. Therefore, any state S can be defined by the 3-tuple (s_0, s_1, l)

4 PMF, PMF, PMF

What Is The Random Variable

First, consider just a single dice. It has pmf

$$D_i^{(s)}(d) = \begin{cases} \frac{1}{s} & \text{if } d \in \{1, ..., s\} \\ 0 & \text{otherwise} \end{cases}$$

So let the random variable T denote the sum of n iid random dice, each with s sides. It can thusly be written

$$T : (\mathbb{N}_0)^n \to \mathbb{R}, T := D_1^{(s)}(d_1) + \cdots + D_n^{(s)}(d_n)$$

Therefore, our goal is to find the pmf of T, which is notated $\mathbf{p}_T^{(n,s)}(t)$, dependent on parameters n, s.

In order to find the probability mass function of T, we'll use moment generating functions. Remember that probability distributions and moment generating functions have a one-to-one correspondence. So for a discrete random variable X with probability distribution

$$f(x_i) = P(X = x_i) = p_i \text{ for } i = 1, 2, ..., k$$

then it's mgf is

$$M_X(t) = \mathbb{E}\big[e^{tX}\big] = \sum_x e^{tx} \cdot f(x)$$

$$= p_1 \cdot e^{tx_1} + \cdots + p_k \cdot e^{tx_k}.$$

and visa versa (like in *HW 6, 2.1*).

What is the MGF of T?

Recall the pmf of $D_i^{(s)}$. It's moment generating function is defined as

$$M_{D_i^{(s)}}(t) = \mathbb{E}\left[e^{tD_i^{(s)}}\right] = \frac{1}{s}(e^t + e^{2t} + \cdots + e^{st})$$

Since $D_i^{(s)}$ are all independent and identically distributed,

$$M_T(t) = \mathbb{E}[e^{tT}] = \mathbb{E}\left[e^{t\left(D_1^{(s)} + \cdots + D_n^{(s)}\right)}\right] = \prod_{i=1}^{n} \mathbb{E}\left[e^{tD_i^{(s)}}\right]$$

$$= \prod_{i=1}^{n}\left[\frac{1}{s}(e^t + e^{2t} + \cdots + e^{st})\right]$$

$$= \frac{1}{s^n}(e^t + e^{2t} + \cdots + e^{st})^n$$

We'll write the multinomial in the nicer form

$$g(x) = \frac{1}{s_n}(x + \cdots + x^s)^n$$

Our goal is to find the coefficient of the x t term.

Rewriting this with the geometric series

$$g(x) = \frac{1}{s_n}(x + \cdots + x^s)^n = \frac{1}{s_n}\left(\frac{x(x^s - 1)}{x - 1}\right)^n = \frac{1}{s_n}\frac{x^n(x^s - 1)^n}{(x - 1)^n}$$

Numerator

Looking at the numerator $(x^s - 1)^n$, we can expand this with the binomial theorem, *Fact D.2. [ASV]*

$$(x^s + 1)^n = \sum_{i=0}^{n} \binom{n}{i} (x^s)^i \cdot (-1)^{n-i}$$

multiplying by x^n, this becomes

$$x^n (x^s + 1)^n = \sum_{i=0}^{n} \binom{n}{i} x^{si+n} \cdot (-1)^{n-i}$$

Looking at the denominator $(x-1)^{-n}$. Rewrite this to $(-1)^{-n} \cdot (1-x)^{-n}$. Again this can be rewritten as the sum of binomial coefficients, but this time $(1-x)^{-n}$ is a binomial series (an expansion of the binomial theorem for complex exponents), in specific the negative binomial. Thus the denominator is

$$(x-1)^n = (-1)^{-n} \cdot (1-x)^{-n} = (-1)^n \sum_{j=0}^{\infty} \binom{n+j-1}{j} x^j$$

Combine Variables

So together, the full equation of $g(x)$ becomes the product of the numerator and denominator as such

$$g(x) = \left(\sum_{i=0}^{n} \binom{n}{i} x^{si+n} \cdot (-1)^{n-i} \right) \left((-1)^n \cdot \sum_{j=0}^{\infty} \binom{n+j-1}{j} x^j \right)$$

Moving the $(-1)^{-n}$ to the other equation and then simplifying $(-1)^{-i}$ to $(-1)^i$

$$g(x) = \left(\sum_{i=0}^{n} (-1)^i \binom{n}{i} x^{si+n} \right) \left(\sum_{j=0}^{\infty} \binom{n+j-1}{j} x^j \right)$$

The finite sum can be though of as an infinite sum that takes 0 whenever $i > n$. This allows us to use Cauchy product to get the coefficients, which generally states that:

$$\left(\sum_{i=0}^{\infty} a_i x^i \right) \left(\sum_{j=0}^{\infty} b_j x^j \right) = \sum_{k=0}^{\infty} c_k x^k$$

Where the coefficients c_k are defined as

$$c_k = \sum_{l=0}^{k} a_l b_{k-l}$$

Combinatorics Magic

21 / 39

Doing some careful accounting of coefficients, this yields the double
summation and result

$$g(x) = \frac{1}{s^n} \sum_{k=0}^{\infty} \left(\sum_{l=0}^{\lfloor \frac{k-n}{s} \rfloor} (-1)^k \binom{n}{k} \binom{k-sl-1}{n-1} \right) x^k$$

Which means that the coefficient of x^t is

$$\frac{1}{s^n} \sum_{l=0}^{\lfloor \frac{t-n}{s} \rfloor} (-1)^k \binom{n}{k} \binom{k-sl-1}{n-1}$$

A Guide to Optimization and Optimal Play in   ...via markov decision processes and dynamic

Theorem

So finally, taking a step back, the pmf of T is the coefficient of the ith term, which is defined above. Thus the pmf of T is defined

$$\mathbf{p}_T^{(n,s)}(t) = \frac{1}{s^n} \sum_{k=0}^{\left\lfloor \frac{t-n}{s} \right\rfloor} (-1)^k \binom{n}{k} \binom{t - s \cdot k - 1}{n - 1}$$

5 Terminal States

Defining a Rating Function on Terminal States

The goal is to find some n to maximize the probability of getting a new score $s_p + t$ in the range $(s_o, M]$?

Or more precisely, given a state $(s_p, s_o, 1)$, what is the optimal n such that the expected rating is maximized.

$$\text{rating}\big((s_p, s_o, 1), n\big) := \sum_{t=s_o+1}^{M} \mathbf{p}_T^{(n,s)}(t - s_p) + \frac{1}{2} \cdot \mathbf{p}_T^{(n,s)}(s_o - s_p)$$

where the summation describes the weighted sum of all next states given n, weighted according to their probability (transition matrix), hence rating is the expectation of its next possible states.

Since $n_\star$ is the optimizer of rating, it is given by

$$n_\star\big((s_p, s_o, 1)\big) := \mathrm{argmax}_n \ \mathrm{rating}\big((s_p, s_o, 1), n\big)$$

Notice that the optimal rating comes for free. It's the rating that was optimized for in finding $n_\star$, so no additional work is required.

$$\mathrm{rating}_\star\big((s_p, s_o, 1)\big) := \mathrm{rating}\big((s_p, s_o, 1), n_\star\big)$$

6 Calculating Terminal States

There are in fact certain states where the choice is immediately obvious without any need for calculations with pmfs. These states can be broken into 2 types that I'll call *forfeit* and *certain victory*.

In *forfeit states*, the opponent decided to end the game while they were behind, which guarantees that you win by doing nothing and rolling 0 dice.

In *certain victory* states, the relation between the player sp and oppoent so compared to the opponent and the maximum M is such that for some n, all possible resulting sums $s_p + t$ are in the range $s_o < s_p + t \leq M$. Thus there is some $n_\star$ with a 100% win rate.

Beginnings of the Algorithm

For most cases, a solution is not so simple. For these, the optimal $n_\star$, $\text{rating}_\star$ are found by the functions already formulated.

The algorithm for efficiently finding such maxima utilizes the properties of the distribution of T, to search minimal amount of n. The explanation for how and why is beyond the scope of this presentation, but you can read about it in the paper (though even then its a conjecture, turns out its hard to prove something when the distribution is constantly changing.)

7 Optimizing Normal States

The rating is defined in the same way as it is for terminal states: The expectation of the rating for the next possible states S_1 given some n.

Example:

Let's give an example. Imagine that the optimal $n_\star$ and rating$_\star$ for every other state is known.

Consider rolling 2 dice at state $(2, 6)$. You could end up at any of the following states: $S_1 = \{(6, 4), (6, 5), (6, 6), (6, 7)\}$. So since we know the rating of all these states, we can calculate the rating given n by taking the weighted sum over S_1, with weights given by the pmf of T, i.e. $\mathbf{p}_T^{(n,s)}(2), ..., \mathbf{p}_T^{(n,s)}(4)$.

Defining a Rating Fucntion on Normal States

Fact: Rating is complementary with respect to each player's rating for a given state.

Hence, the optimal rating for landing on a state S is $1 - \text{rating}(S, n_\star)$ (since the rating of that state will be for the opponent), where $n_\star$ is the optimal n for that state.

So the rating function is given by

$$\text{rating}\big((s_p, s_o, 0), n\big) := \begin{cases} \sum_{t=n}^{s \cdot n} 1 - \text{rating}\big((s_o, s_p + t, 0), n_\star\big) & \text{if } n > 0 \\ 1 - \text{rating}\big((s_o, s_p, 1), n_\star\big) & \text{if } n = 0 \end{cases}$$

Thus the optimal $n_\star$ given any possible state S, normal or terminal is now defined to be

$$n_\star = \text{argmax}_n \left\{ \sum_{t=n}^{s \cdot n} 1 - \text{rating}\big(s_o, s_p + t, \{0 \text{ if } n = 1 \text{ else } 0\}\big) \right\}$$

with the rating function being either equation (1) or (4) depending on whether its a terminal state or normal state.

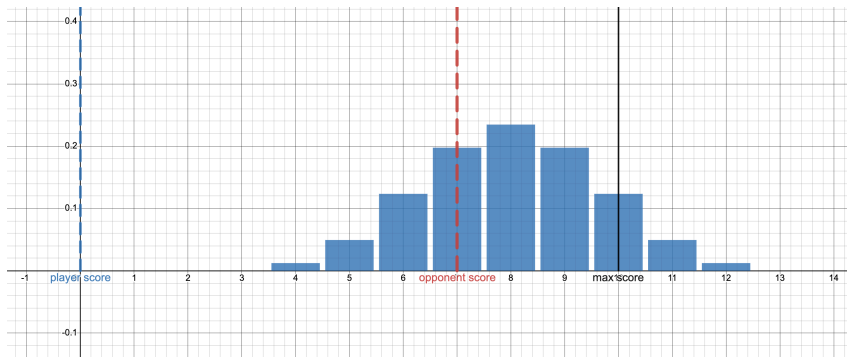With $\text{rating}_\star$ defined the same as for terminal states.

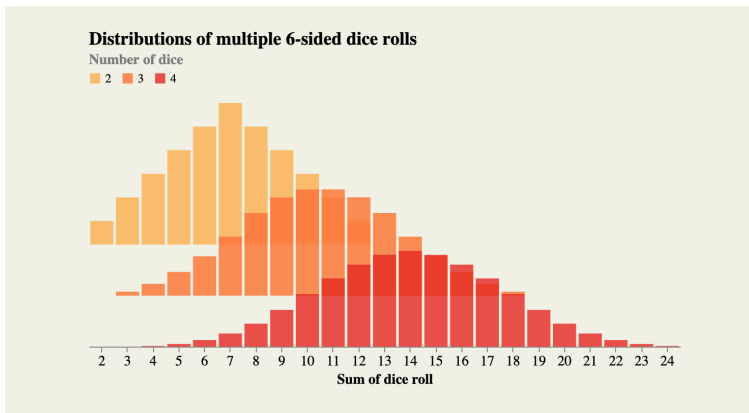# 8 Plots

Figure 1: Visual demonstration of terminal states

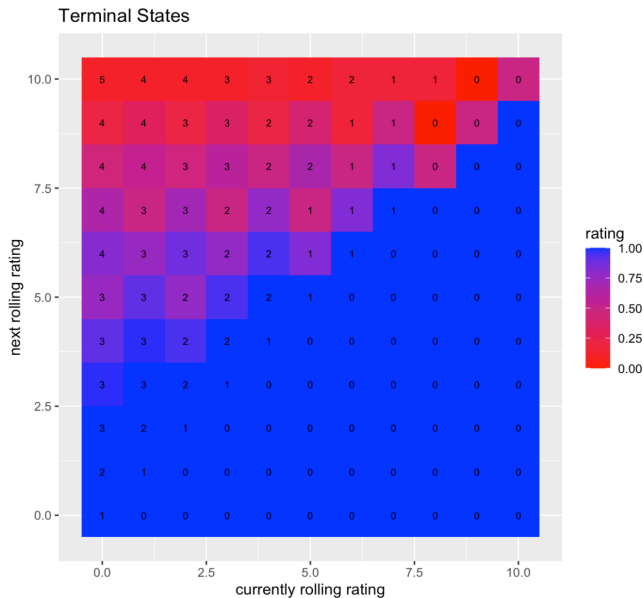Figure 2: Distributions of $n = 2, 3, 4$ with $s = 6$

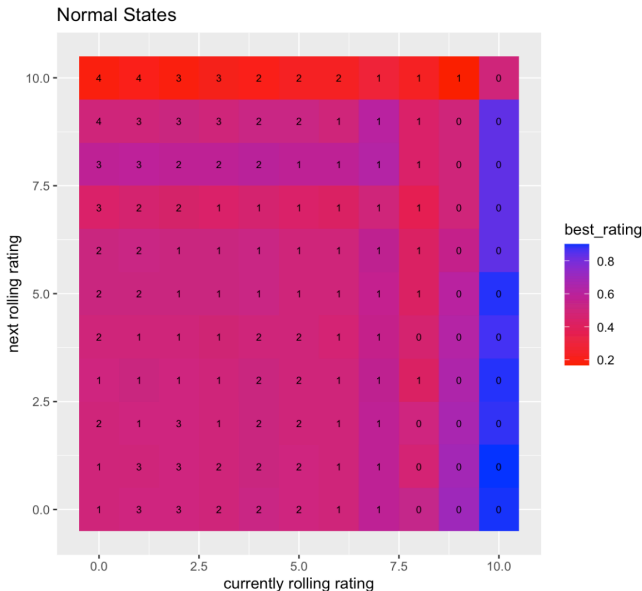Figure 3: Optimal actions and ratings for terminal states

Figure 4: Optimal actions and ratings for normal states

9 References

References

[1] D. F. Anderson, T. Seppäläinen, and B. Valkó, *Introduction to Probability*. Cambridge: Cambridge University Press, 2018.

[2] A. K. Singh, R. J. Dalpatadu, and A. F. Lucas, "The Probability Distribution of the Sum of Several Dice: Slot Applications," *Gaming Research & Review Journal*.

[3] Analytics Check, "Deriving the Probability Distribution for the Sum of Many Dice Rolls." [Online]. Available: https://www.analyticscheck.net/posts/sums-dice-rolls