

Reinforcement Learning 4: 101 And You're Done

Eckel, TJHSST AI2, Spring 2024

Background & Explanation

I was at a math conference once, in the late 2010s, and to demonstrate the idea of a problem that was simple to state but complex to solve, one of the teachers running a session – Marilyn Burns – showed a game she'd created called "101 And You're Done".

It goes like this. You get 6 turns. At each turn, you roll a single die (a normal one, with the sides 1-6 equally weighted). Your choice is whether to *keep* the value you got, or to *multiply it by 10*. Add each result to your total. The goal is to get as close to 100 as possible without going over. If you don't go over, your score is whatever your sum is. If you do go over, your score is 0. So, the highest score possible is 100.

One example:

- 1) You roll a 5. You choose to x10. Current total: 50.
- 2) You roll a 6. You choose to keep. Current total: 56.
- 3) You roll a 1. You choose to x10. Current total: 66.
- 4) You roll a 3. You choose to x10. Current total: 96.
- 5) You roll a 3. You choose to keep. Current total: 99.
- 6) You roll a 4. You choose to keep. Your total exceeds 100, so your final score is 0.

Marilyn invented this game to help elementary schoolers practice place value. But she asked us – how good can an ideal strategy get? What average score will you get if you play perfectly? She said she hadn't solved it. I didn't solve it either, and neither did anyone else at my table in that session.

It wouldn't quite be accurate to say this question has haunted me ever since, but it has *bugged* me. I find myself thinking about it every so often. I haven't ever answered it.

UNTIL NOW.

Reinforcement learning for the win! I solved this problem on the evening of Sunday, 6/2, and I'm *extremely* pleased.

Your turn! Your task is to find the best strategy for this game that you can.

Specification

Submit a **single python script** to the link on the course website and **send a message on Mattermost** (see below).

It takes no command line arguments.

Run some kind of training process, train a strategy, then play 100,000 games with your fully trained strategy (no further modifications). Print out your total score across all 100,000 games.

This assignment is **complete** if:

- You follow the instructions on the submission form to format your submission properly.
- Your script works as described.
- Your strategy is very close to the ideal strategy.
- You send a message on Mattermost explaining briefly but clearly the training process you ran.