# Reinforcement Learning

## Trong

November 1, 2018—December 20, 2018

# Contents

# 1 DOOM: Health Gathering in Acid Lake

Agent seems to converge to a specific behavior, e.g. if in the beginning it tends to turn left, it will gradually turn tighter and tighter corners until it's spinning in one place. I suspect this is a property of the loss function

$$L = -\frac{1}{n} \sum_i \langle A, \log s(\pi) \rangle \cdot G$$
$$G = \frac{[G_0 \cdots G_n] - \mu}{\sigma}$$

and not the network architecture. This is because the loss function's main goal via the $\langle A, \log s(\pi) \rangle$ factor is to create actions that conform with previous behavior.

Adding randomness in the loss function, e.g. in the way we generate the discounted rewards, seems to help, but that only delays the pathological behavior.