Epidemic Control with Reinforcement Learning

Qiyao Wei
Romina Abachi
Ehsan Mehralian

CONTENTS

01 SIR Model

02 Single node

03 Network

04 Further

01 SIR Model

SIR Model Diagram

Source: https://www.google.com/imgres?imgurl=https%3A%2F%2Fwww.lewuathe.com%2Fassets%2Fimg%2Fposts%2F2020-03-1

Parameter Reference

Continuous-time:

$$\frac{dS}{dt} = \frac{-\beta}{(1+v)N} SI - uS$$
$$\frac{dI}{dt} = \frac{\beta}{(1+v)N} SI - \alpha I - rI$$
$$\frac{dN}{dt} = -(1-f)\alpha I$$
$$R = N - S - I$$

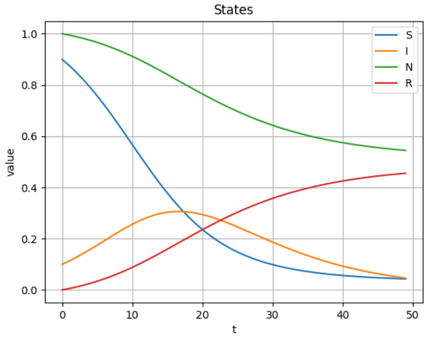
1. Beta: transmission coefficient (susceptible becoming infectious)
2. Alpha: rate of infectives leaving the infected class
3. f: proportion of infectives recovering, with the remainder dying of infection (think of this as a "subset" of alpha)
4. u, v, r: control variables/actions
5. Moment of caution: I will use the term "return" to represent the returns in training, and I will use the term "reward" and "penalty" interchangeably to represent the returns in evaluation.

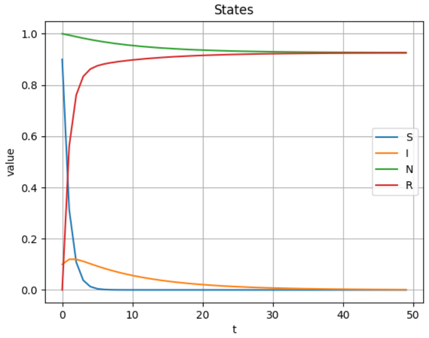
Source: https://www.maa.org/press/periodicals/loci/joma/the-sir-model-for-spread-of-disease-the-differential-equation-model

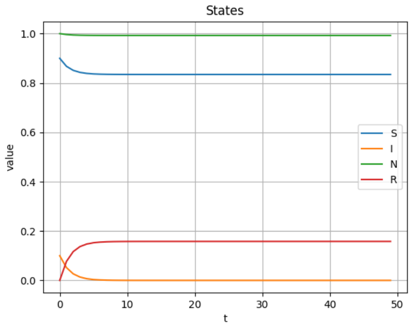
SIR Evolution

v (inform public) | u (vaccination) | r (quarantine)

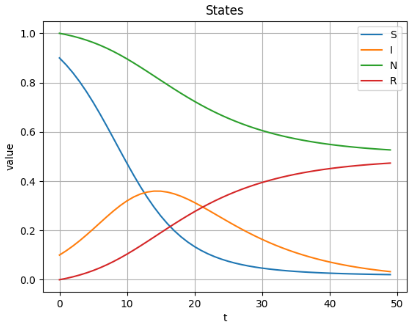
action = 1

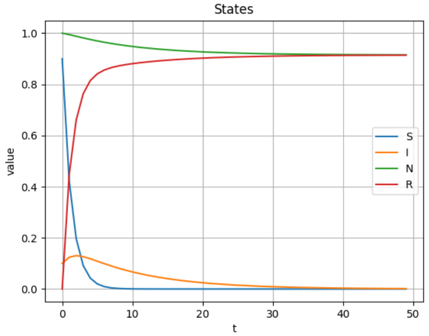


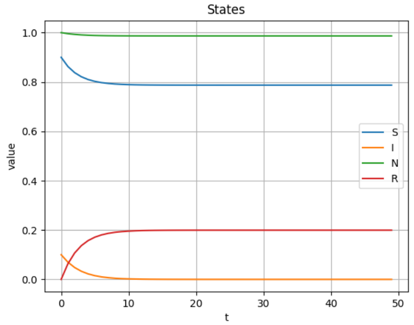




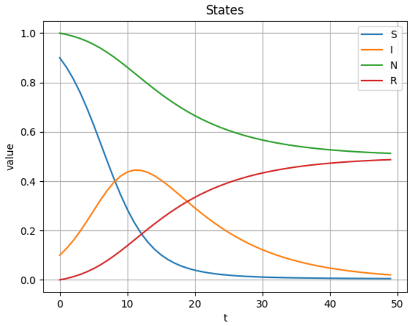
action = 0.7

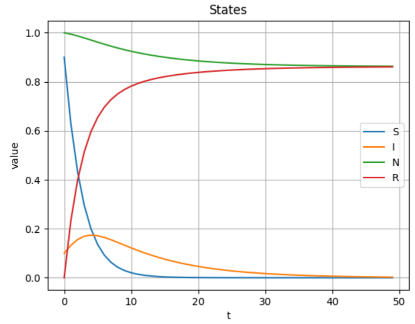


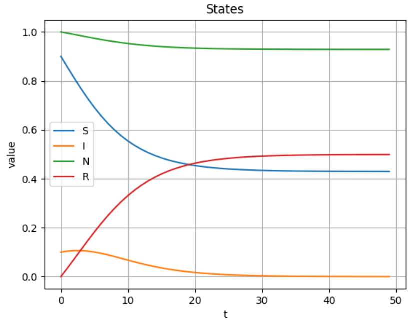




action = 0.3



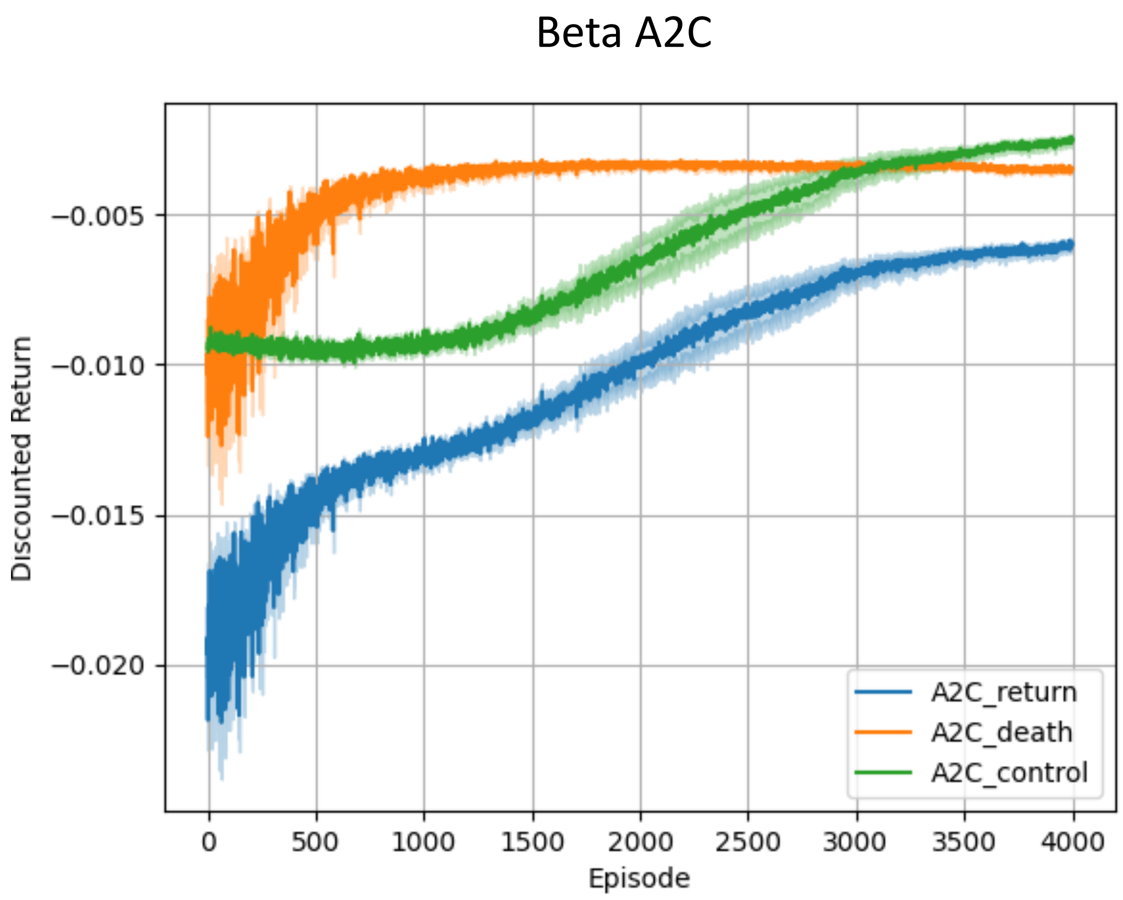


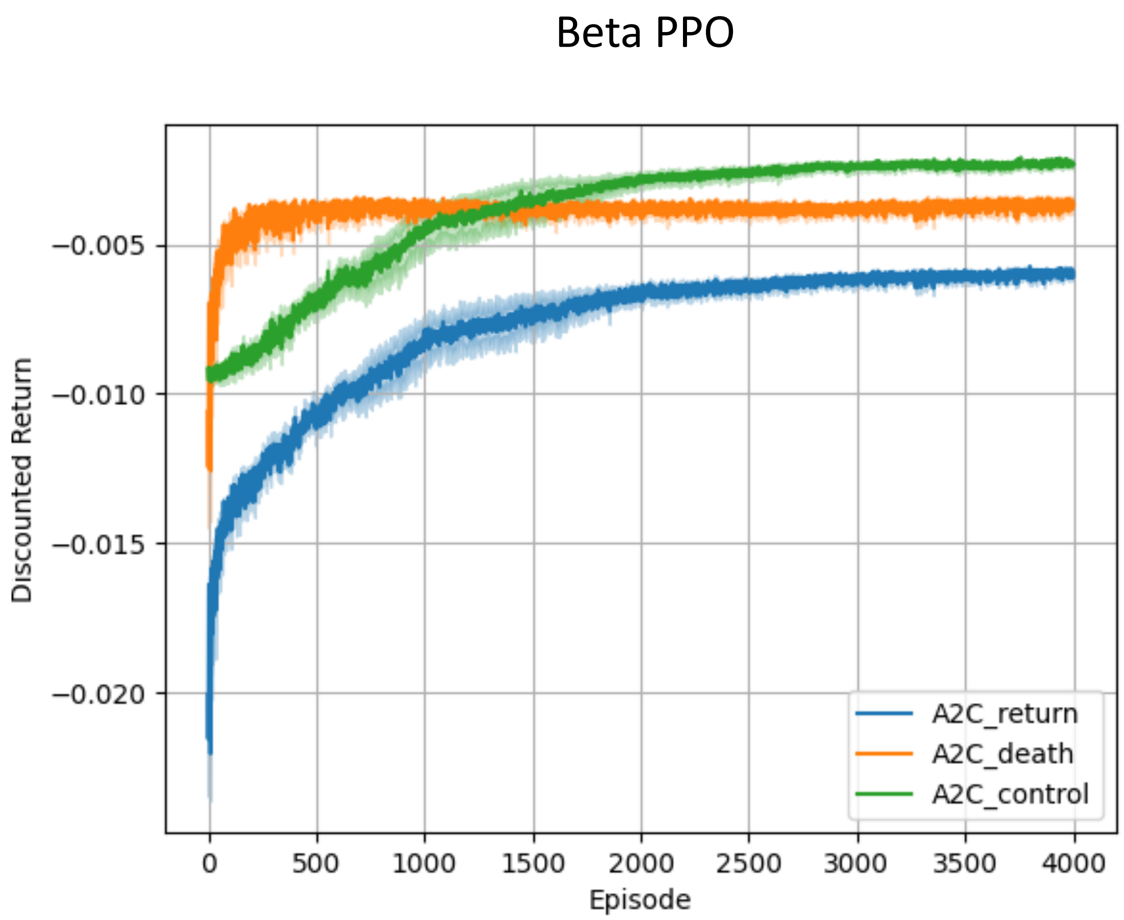


02

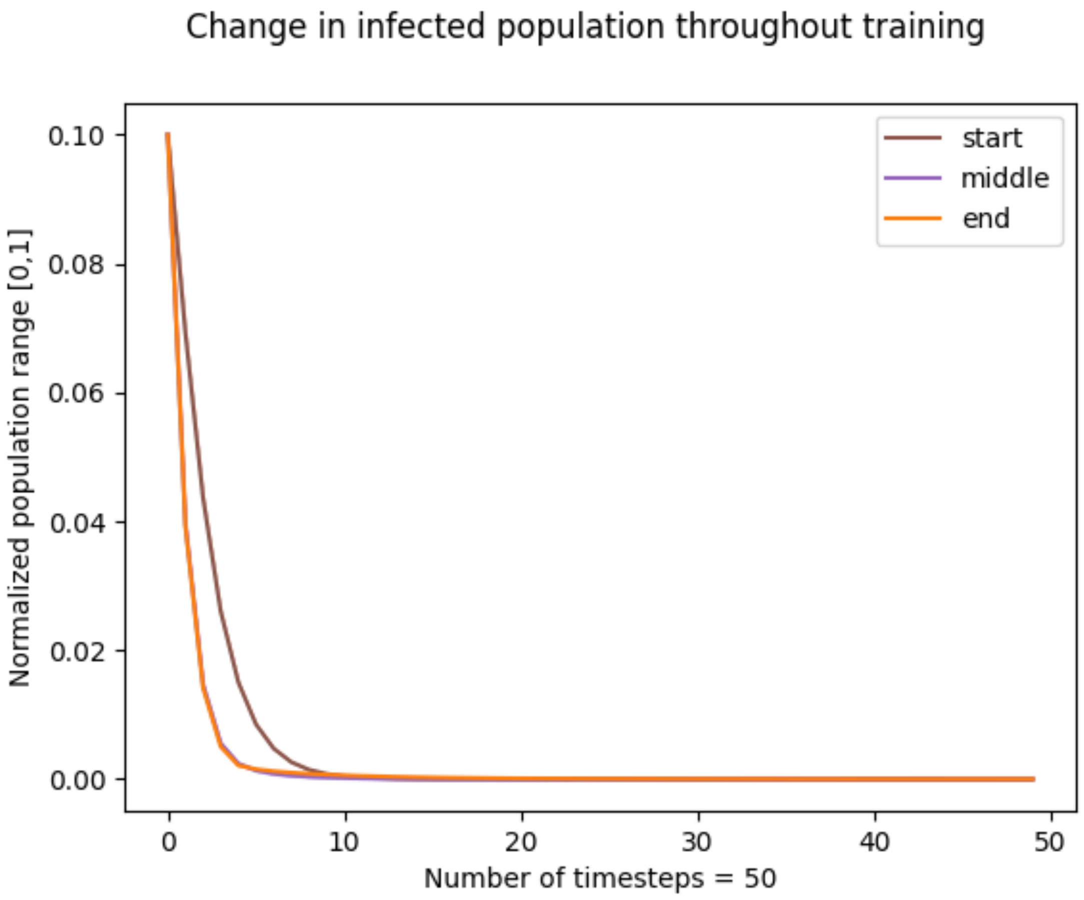
Single node

Training Returns



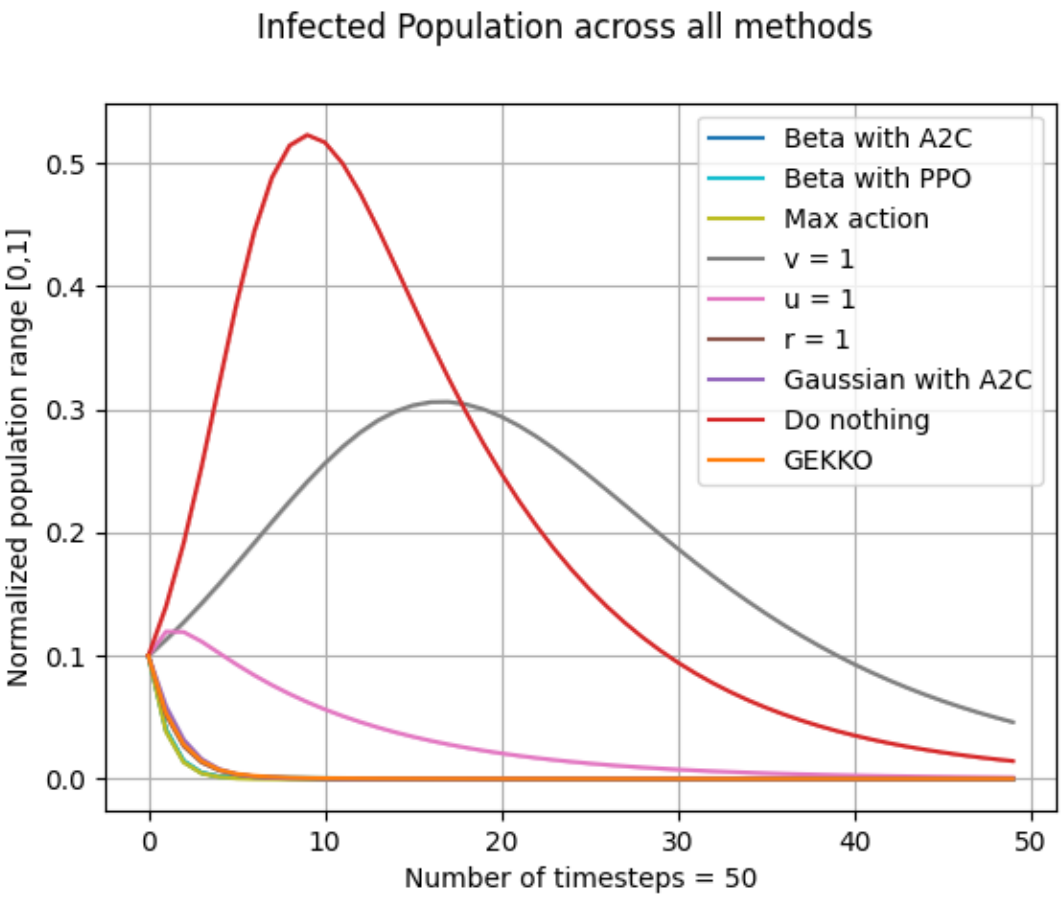


Change in infected



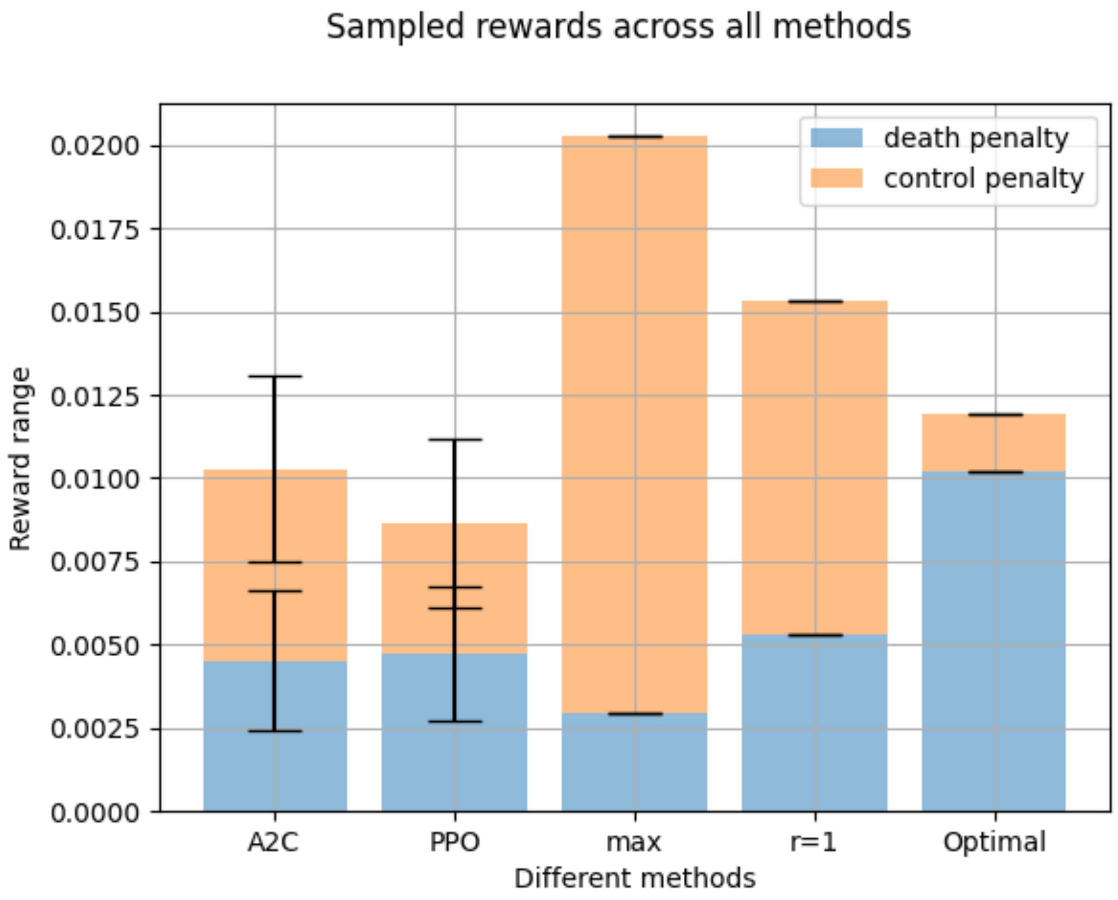
- Beta A2C
- Evaluated at timestep 0 (start), 100e3 (middle), and 200e3 (end)

Infected Population



- Sampled over 50 timesteps
- Most ”good” control curves overlap, since they control the infected population right away

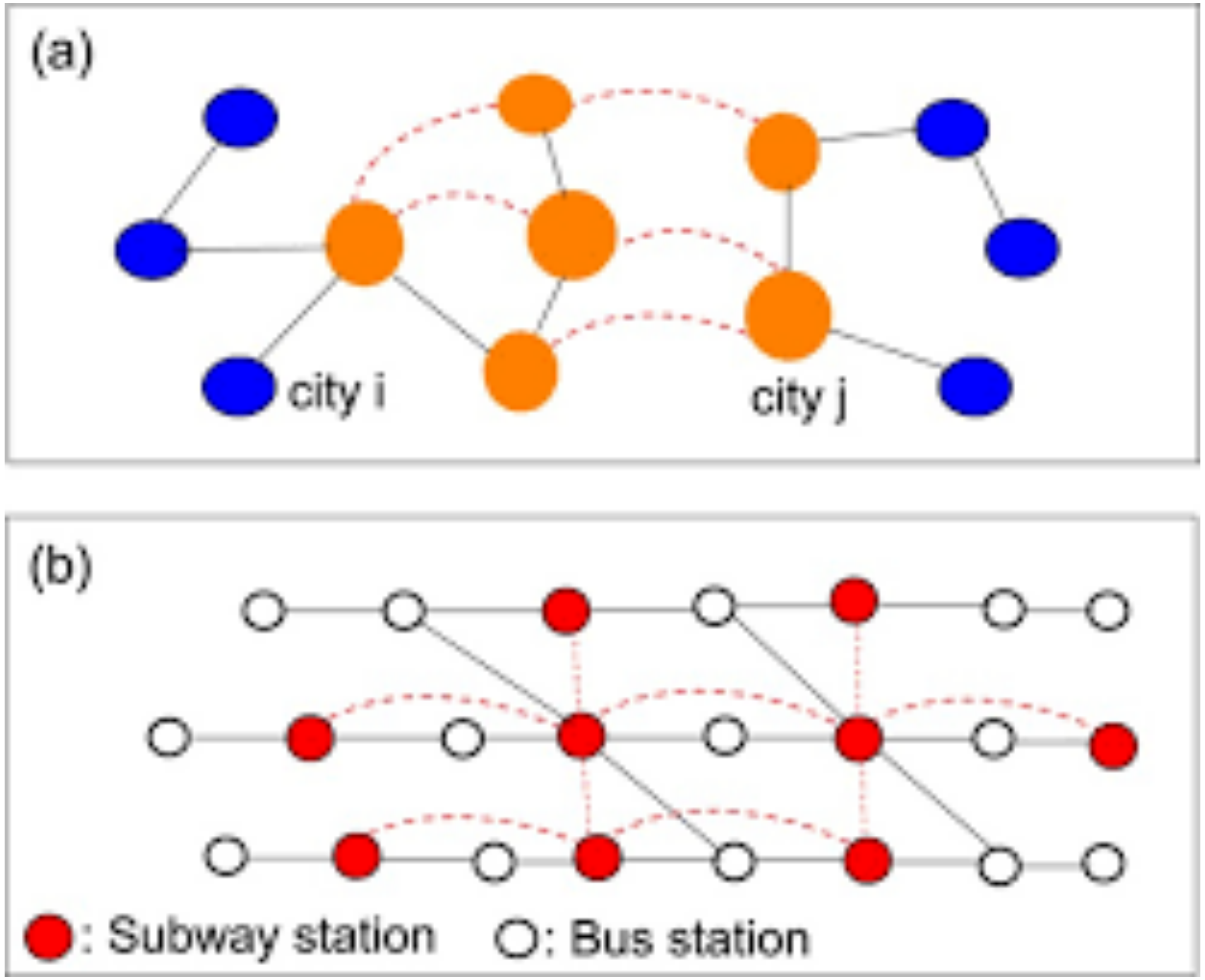
Sampled Rewards



- Our agent can outperform optimal control policy (GEKKO)
- We did not include Gaussian because it has a high variance. We did not include v = 1 and u = 1 heuristics because their penalties are too high

03 Network

SIR Network Diagram



Source: https://www.google.com/imgres?imgurl=https%3A%2F%2Fwww.researchgate.net%2Fpublication%2F317711741%2Ffigur
Source: https://www.google.com/imgres?imgurl=https%3A%2F%2Fmedia.springernature.com%2Fm685%2Fspringer-static%2Fima

SIR Network Equations

$$S'_j = -uS_j - (1-\omega)S_j \sum_k \beta m_{j,k} \frac{I_k}{N_k}$$

$$I'_j = -\alpha I_j - rI_j + (1-\omega)S_j \sum_k \beta m_{j,k} \frac{I_k}{N_k}$$

$$N'_j = -(1-f)\alpha I_j$$

$$R_j = N_j - I_j - S_j$$

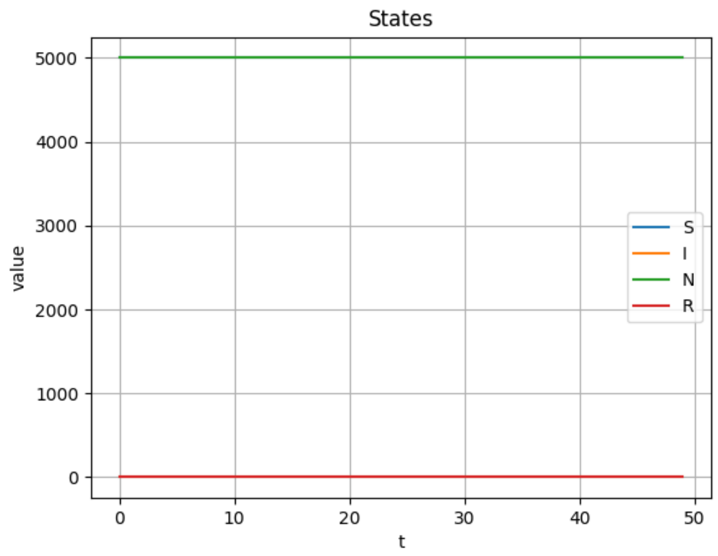
Sanity Check (1)

1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.5	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.5	0.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.1	0.7	0.2	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.2	0.0	0.8	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.8	0.2	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.1
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.9

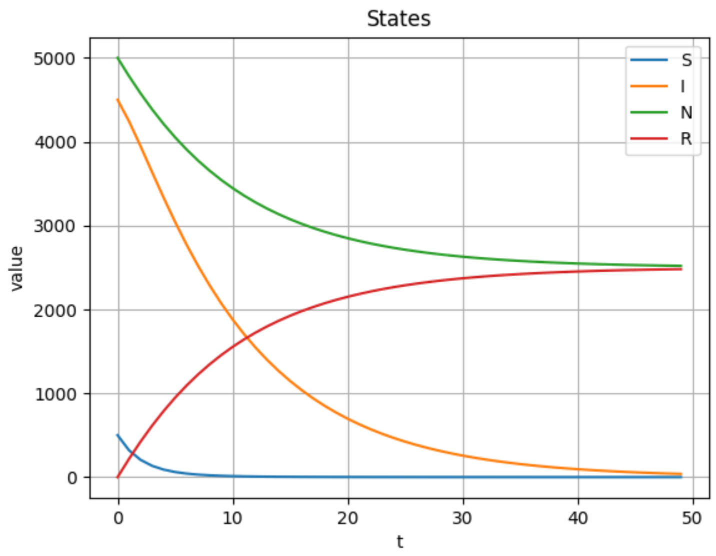
S_0	5000	4500	4500	4000	4000	3000	2000	1000	500	500
I_0	0	500	500	1000	1000	2000	3000	4000	4500	4500

Sanity Check (2)

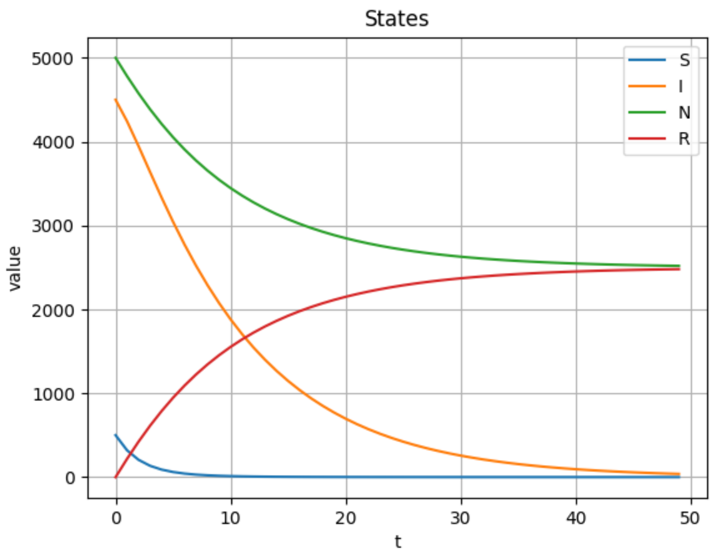
node0



node8



node9

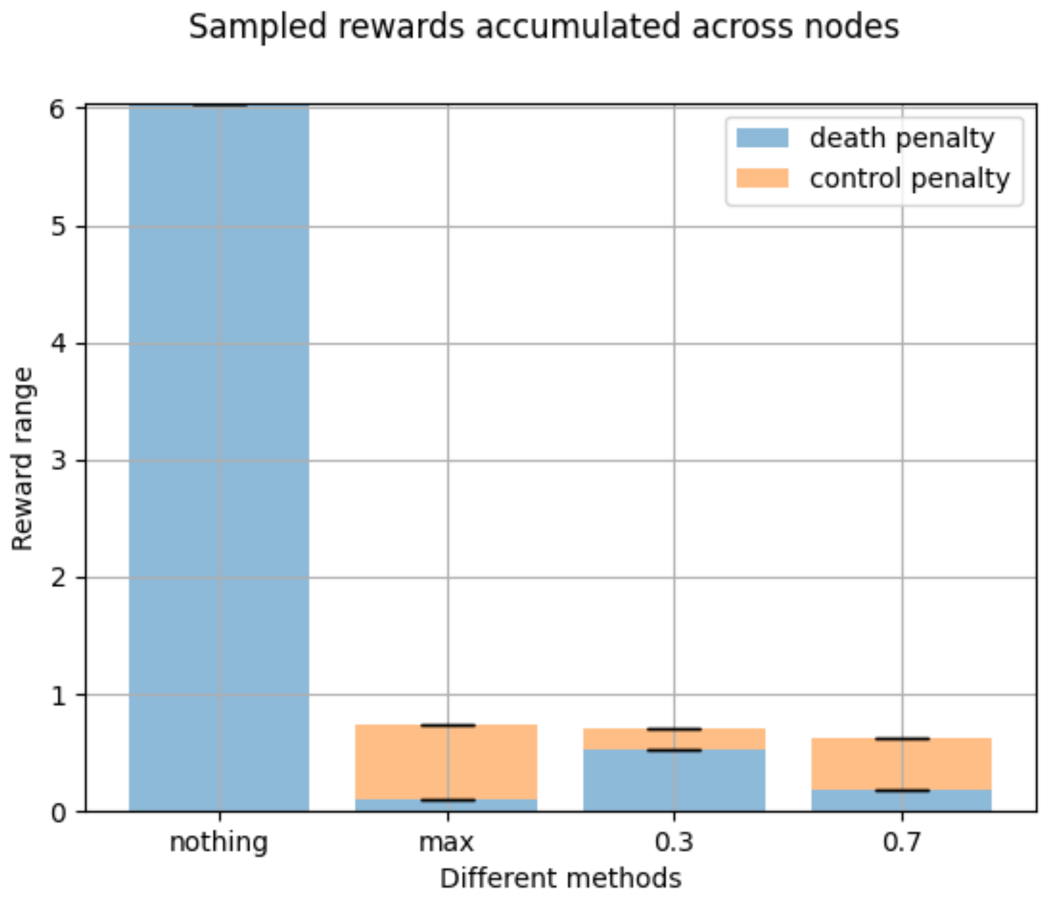


1. Node 0 does not experience the epidemic, which is expected

2. Node 8 and 9 are the most affected, since they start off with a high infected population

3. Node 8 and 9 have identical SIR curves, which is expected

Sampled Rewards



04

Future Directions

Future Directions

- 1. Design heuristic tests on the network formulation, and make sure they work with our expectations.
- 2. Test the network on real data