Optimal Earthquake-Prevention Renovations using Reinforcement Learning

Oumnia Chellah
Stanford University
oumnia@stanford.edu

Jean Rodmond Junior Laguerre
Stanford University
jeanlag1@stanford.edu

***Abstract*—We use reinforcement learning to determine which buildings should be renovated in an area prone to earthquakes, while minimizing the amount of renovations needed and the damage buildings will receive in case of an earthquake. We create a list of three buildings, containing three integers from 0 to 1 representing the preparedness level of each building. We start with the initial preparedness levels, and update preparedness over time: at each time step, we choose to renovate up to one building, and an earthquake could happen at each time step with a low probability. Our reward function penalizes for buildings that have a low preparedness score. Using Q-learning and SARSA, we find that a random policy outperforms Q-learning on different variations of the reward function, and that [insert results for SARSA], showing that [insert results for sarsa and q-learning].**

I. Introduction

In regions threatened by seismic activities, the resilience of infrastructure, particularly that of existing buildings, stands as a paramount concern for urban planners, architects, and policy-makers. Major earthquakes bring forth a significant amount of devastation and loss of human lives, which makes it even more important for our policy-makers to make the right decisions when it comes to retrofitting and renovating structures to withstand seismic events. However, the task of identifying and prioritizing buildings for seismic retrofitting is a challenging task.

Current approaches to this prioritization often grapple with multiple factors, including the structural characteristics of buildings, their historical significance, occupancy rates, and the urban roles that these buildings play. Additionally, budgetary constraints and the unpredictable nature of earthquakes, both in terms of magnitude and frequency, add layers of complexity to the decision-making process.

Our project aims to address this challenge by developing an innovative framework using the computational techniques that we have learned in class. By leveraging data-driven models, we want to create a more objective approach to assess the earthquake resilience of buildings, and to decide which buildings to renovate in order to minimize the damage that an earthquake would cause. In this project, we want to ensure that the most vulnerable structures receive due attention.

In this paper, we outline the development of our model, describe our work, and demonstrate its application in a real-world urban setting. Such a model is important not only for the buildings that we have looked at, but also for learning how to protect buildings in other parts of the world from earthquakes in the future.

Identify applicable funding agency here. If none, delete this.

II. Prior Work

Recent advancements in earthquake forecasting and preparedness have been heavily influenced by developments in machine learning. Key contributions in this area include:

1) **Earthquake Detection and Phase Picking**: Mousavi et al. developed an advanced deep-learning model focusing on earthquake detection and phase picking. This model significantly improves the accuracy of identifying seismic events, which is essential for timely and effective earthquake response strategies [2].
2) **Uniform California Earthquake Rupture Forecast (UCERF)**: Field et al. contributed a comprehensive model for earthquake forecasting. The UCERF model is a cornerstone in understanding seismic activities, offering a detailed view of potential earthquake occurrences in California [3].
3) **Statistical Models for Earthquake Occurrences**: Yosihiko Ogata's research into statistical models for earthquake occurrences provides invaluable insights into the patterns and likelihood of seismic activities. His work has been instrumental in developing predictive models for earthquake forecasting [4].

III. Methods

In our study, we explored three distinct reinforcement learning methods: Random Policy, SARSA (State-Action-Reward-State-Action), and Q-Learning. Each method was implemented with specific hyperparameters tailored to optimize performance.

A. Random Policy

The Random Policy approach serves as a baseline for comparison. In this method, actions are selected randomly from the action space at each step, without any learning mechanism.

B. SARSA

SARSA is an on-policy reinforcement learning algorithm that updates the Q-value based on the action taken and the subsequent state. The update rule for SARSA is given by:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma Q(s',a') - Q(s,a)\right] \quad (1)$$

where s and s' are the current and next states, a and a' are the current and next actions, r is the reward, α is the learning rate, and γ is the discount factor. The hyperparameters used in our implementation were:

- Learning Rate (α): 0.1
- Discount Factor (γ): 0.9
- Number of Episodes: 1000

C. Q-Learning

Q-Learning is an off-policy algorithm that updates the Q-value based on the maximum expected future rewards. The update rule is:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right] \quad (2)$$

The hyperparameters for Q-Learning in our study were:

- Learning Rate (α): 0.01
- Discount Factor (γ): 0.9
- Epsilon (for exploration): Initial value of 0.95, decreasing by 0.1 every 1000 episodes
- Number of Iterations: 1000

Each of these methods was evaluated based on their performance in optimizing [specific task or objective from the project].

IV. Analysis

A. Transition Function

Our transition function is integral to the state dynamics within our environment. We set the earthquake occurrence probability as $p_e = 0.05$. For each building i, the transition probabilities under renovation action a_i are defined as:

- With renovation ($a_i = 1$):

$$T(s_i + 0.2 \mid s_i, a_i = 1) = 0.9$$
$$T(s_i - 0.6 \mid s_i, a_i = 1) = p_e$$

- Without renovation ($a_i = 0$):

$$T(s_i \mid s_i, a_i = 0) = 0.7$$
$$T(s_i - 0.2 \mid s_i, a_i = 0) = 0.2$$
$$T(s_i - 0.6 \mid s_i, a_i = 0) = p_e$$

B. Reward Functions

We propose two reward function models to incentivize structural integrity:

1) The first model penalizes buildings below a preparedness threshold of 0.2 after accounting for renovation costs:

$$R = \sum_{s_i > 0.2} 1 - 20 - 50 \times \sum_{s_i \leq 0.2} 1$$

2) The second model rewards maintaining buildings above the threshold more than it penalizes collapses:

$$R = 30 \times \sum_{s_i > 0.2} 1 - 20 - 50 \times \sum_{s_i \leq 0.2} 1$$

V. Evaluation

To evaluate the efficacy of the different reinforcement learning methods, we conducted a series of experiments using the specified reward functions. The performance of the Random Policy, SARSA, and Q-Learning algorithms were compared to determine the most effective approach for our objectives.

A. Evaluation of Reward Function 1

For the first reward function, which aims to discourage buildings from collapsing by penalizing buildings below a preparedness threshold of 0.2, we observed the following results:

Reward Function 1 Evaluation:

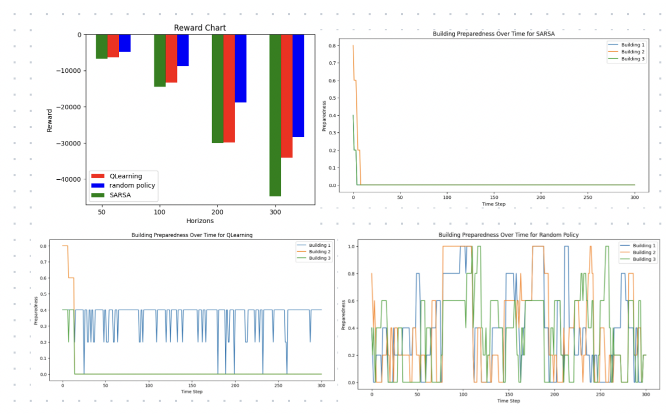

Fig. 1. Comparison of methods for Reward Function 1

B. Evaluation of Reward Function 2

For the second reward function, which focuses on maintaining buildings above the threshold more than penalizing collapses, the experimental results are as follows:

Reward Function 2 Evaluation:

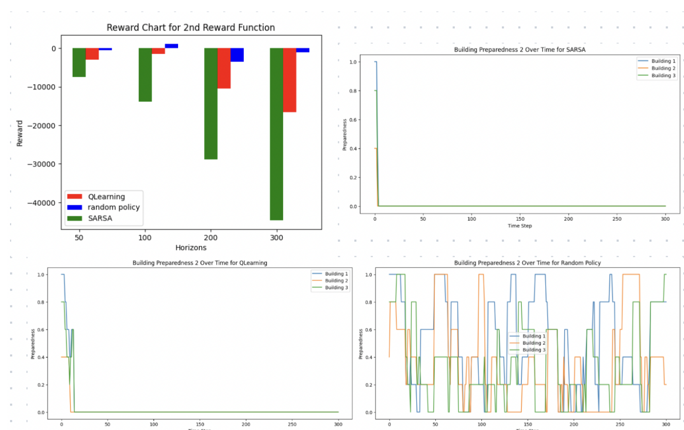

Fig. 2. Comparison of methods for Reward Function 2

VI. Discussion

Our QLearning and SARSA algorithms seem to do worse than a random policy. When some buildings are in bad shape (state = 0), our Qlearning algorithm still chooses to not renovate any buildings. This might be because the reward for buildings that are in good shape is not high enough.

The QLearning Algorithm does seem to renovate the right buildings (buildings where state ¡ 0.2) for multiple states, which helps keep these buildings in good shape in case an earthquake happens. For instance, in state [0, 0.6, 0], the optimal action we get is renovating building 3, and in state [0.4 0.2 1], the optimal action is to renovate building 2. However, our algorithm also often chooses to do nothing (in state [1, 0.8, 0], the optimal action is 0) when one of the buildings is in bad shape, which is not a behavior that we want: ideally, we would like that building to be renovated.

Random policy also outperforms QLearning for the second reward function. Let's take a look at some Q(s,a) values from the second function. Our QLearning algorithm often renovates buildings that are already in good shape in order to increase the reward, which is not behavior that we want. Ideally, we want buildings that are in terrible shape (0) to be prioritized.

VII. Conclusion

The results obtained from the evaluations of each reward function using the different methods provided insightful data on the trade-offs between immediate penalties and long-term structural integrity. These findings should be discussed in further details, focusing on the implications and potential for real-world applications.

Our evaluation indicates that each method has its strengths and weaknesses, and the choice of method may depend on specific use cases and objectives. These results will guide future research and development in the field of earthquake preparedness and response strategies.

References

[1] Kuang, W., Yuan, C., Zou, Z., Zhang, J., & Zhang, W. (2023). *Autonomous Earthquake Location via Deep Reinforcement Learning*. Seismological Research Letters.

[2] Mousavi, S. M., et al. (2020). *Earthquake Transformer - an attentive deep-learning model for simultaneous earthquake detection and phase picking*. Nature Communications, 11, 3952.

[3] Field, E. H., et al. (2017). *A synoptic view of the third Uniform California Earthquake Rupture Forecast model*. Seismological Research Letters, 88, 1259–1267.

[4] Bergen, K. J., et al. (2019). *Machine learning for data-driven discovery in solid earth geoscience*. Science, 363.