PyroRL: A Reinforcement Learning Environment for Wildfire Evacuation

Joseph Guman[1*], **Joseph C. O'Brien**[1*], **Christopher Pondoc**[1*], **and Mykel J. Kochenderfer**[2]

1 Department of Computer Science, Stanford University **2** Department of Aeronautics and Astronautics, Stanford University ***** These authors contributed equally.

Summary

A major effect of climate change today is the increased frequency and intensity of wildfires. This reality has led to increased research in wildfire response, particularly with reinforcement learning (RL). While much effort has centered on modeling wildfire spread or surveillance, wildfire evacuation has received less attention. We present PyroRL, a new RL environment for wildfire evacuation. The environment, which builds upon the Gymnasium API standard (Towers et al., 2023), simulates evacuating populated areas through paths from a grid world containing wildfires. This work can serve as a basis for new strategies for wildfire evacuation.

Statement of Need

There has been significant traction in the use of computational models to study wildfires. Historically, much work has focused on accurately modeling the spread of wildfires. While a lot of older methods were primarily done using physics-based methods (Andrews, 1986; Rothermel, 1972) – with Rothermel being one of the most popular, as well the one we utilize in our package – newer methods rely on machine learning and other data-driven approaches, incorporating a higher diversity of features (Diao et al., 2020; Joseph et al., 2019; Ross, 2021).

Reinforcement learning (RL), a subdomain of artificial intelligence where models learn through interaction with their environment – has also been increasingly used in the context of wildfires. In combination with other traditional statistical methods and computer vision (Ganapathi Subramanian & Crowley, 2018; Subramanian & Crowley, 2017), RL has been applied to both the surveillance and monitoring of wildfires (Altamimi et al., 2022; Julian & Kochenderfer, 2019; Viseras et al., 2021). An area where there has been little work in regards to RL is wildfire evacuation. Understanding the effective approaches for evacuating populated areas during wildfires is a key safety concern during these events (Kuligowski, 2021; McCaffrey et al., 2017), and other machine learning techniques have proven beneficial for evacuation planning (Xu et al., 2023). As a result, work has been done to better model traffic during wildfire evacuation scenarios (Intini et al., 2019; Pel et al., 2012), and agent-based evacuation simulations have been used for not only wildfires but also other natural disasters like tsunamis (Beloglazov et al., 2016; Wang et al., 2016). RL has been previously identified as an intriguing tool for evacuation operations (Rempel & Shiell, 2023) and has been used to model evacuation during electrical substation fires (Budakova et al., 2024). The application of RL techniques to the wildfire evacuation task could thus prove beneficial.

Given the growing interest in studying wildfires through a computational lens, there have been developments in simulators for wildfires. A lot of open-source software focus on providing a visualization of wildfire spread (Rezaei-Shoshtari, 2020; Soto, 2021). The most relevant piece of work to our paper are SimFire and SimHarness, which provide a system for wildland

fire spread and a way for appropriate mitigation strategy responses via RL (Tapley et al., 2023). Nonetheless, the focus is still on wildfire surveillance and mitigation, not on the task of evacuation.

Methods

Our environment for wildfire evacuation builds upon the Gymnasium API standard. This standard has functions to help the user `step` through a single discrete time step of the environment, `reset` the environment back to its original state, and `render` the environment for the user.

Wildfire Evacuation as a Markov Decision Process

Under the hood, the environment is a grid world with dimensions that the user can specify. Each cell in the grid is one of the following:

- Normal terrain
- A populated area
- A fire cell
- Part of an evacuation path

The task at hand is to evacuate all of the populated areas through paths before they are ignited. Similar to how evacuation would occur in real life, a centralized decision maker assigns paths to populated areas. Furthermore, the amount of time to evacuate a populated area is proportional to the number of grid squares in the path.

The problem is modeled as a fully observable Markov Decision Process (MDP) (Kochenderfer et al., 2022).

State Space

The state space is each possible configuration of the grid. Internally, this is represented as a 5 by m by n tensor, where m and n represent the number of rows and columns of the grid world, respectively. Each of the 5 indices corresponds to a specific attribute:

- `0 = FIRE_INDEX` – whether or not a square in the grid world is on fire or not
- `1 = FUEL_INDEX` – the amount of fuel in the square, which is used to determine if an area will be enflamed or not
- `2 = POPULATED_INDEX` – whether or not a square is a populated area or not
- `3 = EVACUATING_INDEX` – whether or not a square is evacuating or not
- `4 = PATHS_INDEX` – the number of paths a square is a part of

Action Space

The action space describes whether or not a populated area should evacuate. If evacuating, the agent must choose a specific populated area to evacuate, as well as a path from which to evacuate.

- Transition Model: determined by the stochastic nature of the wildfire implementation, which we describe below
- Reward Model: $+1$ for every populated area that has not evacuated and is not ignited, and -100 if a populated area is burned down

Modeling the Spread of Wildfires

Finally, our stochastic wildfire model is based on prior work (Julian & Kochenderfer, 2019):

Fuel

Each fire cell has an initial fuel level $\sim \mathcal{N}(\mu, (\sigma)^2)$

- A cell currently on fire has its fuel level drop by 1 after each time step until it runs out of fuel

Users can configure μ and σ. By defualt, their respective values are $\mu = 8.5$ and $\sigma = \sqrt{3}$

Fire Spread

We define the spread of the fire by the following equation: $p(s) = 1 - \Pi_{s'}(1 - \lambda \cdot d(s, s') \cdot B(s'))$

- $p(s)$ represents the probability of non-inflamed cell s alighting
- s' represents an adjacent cell
- λ is a hyperparameter that affects the probability of fire spreading from one cell to an adjacent one
- $d(s, s')$ is the Euclidean distance between cells
- $B(s)$ is a Boolean to check if cell is currently on fire

Users can configure λ. By defualt, the value is $\lambda = 0.094$.

Wind

The spread of wildfire is also influenced by wind. This bias is modeled through a linear transformation using two properties. First, the wind speed affects the probability of neighboring cells igniting the current cell. In addition, for each cell, we calculate the vectors that point in the direction of each neighboring cell. We then take the dot product between this vector and the direction of the wind.

Features

Users can control the configuration of the environment, including the dimensions of the grid world, the locations of populated areas and paths, and initial fire placement. Users can also select from our library of example environments or have the program dynamically generate environments. We also offer examples of how to use popular RL libraries with our environment, such as Stable-Baselines3 (Raffin et al., 2021).

In addition, users can generate visualizations of the environment by calling the `render` function of the Gymnasium environment object. After executing a set of actions on the environment, the user can call `generate_gif`, which stitches together all of the images rendered by the user into a single GIF.

References

Altamimi, A., Lagoa, C., Borges, J. G., McDill, M. E., Andriotis, C., & Papakonstantinou, K. (2022). Large-scale wildfire mitigation through deep reinforcement learning. *Frontiers in Forests and Global Change*, *5*, 734330. https://doi.org/10.3389/ffgc.2022.734330

Andrews, P. L. (1986). *BEHAVE: Fire behavior prediction and fuel modeling system-BURN subsystem, part 1* (pp. INT-GTR-194). U.S. Department of Agriculture, Forest Service, Intermountain Research Station. https://doi.org/10.2737/INT-GTR-194

Beloglazov, A., Almashor, M., Abebe, E., Richter, J., & Steer, K. C. B. (2016). Simulation of wildfire evacuation with dynamic factors and model composition. *Simulation Modelling Practice and Theory*, *60*, 144–159. https://doi.org/10.1016/j.simpat.2015.10.002

Budakova, D., Vasilev, V., & Dakovski, L. (2024). A reinforcement learning algorithm for the optimal evacuation route finding from an electrical substation. *AIP Conference Proceedings*, *3078*(1), 040005. https://doi.org/10.1063/5.0209018

Diao, T., Singla, S., Mukhopadhyay, A., Eldawy, A., Shachter, R., & Kochenderfer, M. J. (2020). Uncertainty aware wildfire management. *AIAA Fall Symposium*. https://doi.org/10.48550/arXiv.2010.07915

Ganapathi Subramanian, S., & Crowley, M. (2018). Using spatial reinforcement learning to build forest wildfire dynamics models from satellite images. *Frontiers in Information and Communication Technology*, *5*, 6. https://doi.org/10.3389/FICT.2018.00006

Intini, P., Ronchi, E., Gwynne, S., & Pel, A. (2019). Traffic modeling for wildland–urban interface fire evacuation. *Journal of Transportation Engineering, Part A: Systems*, *145*(3), 04019002. https://doi.org/10.1061/JTEPBS.0000221

Joseph, M. B., Rossi, M. W., Mietkiewicz, N. P., Mahood, A. L., Cattau, M. E., St. Denis, L. A., Nagy, R. C., Iglesias, V., Abatzoglou, J. T., & Balch, J. K. (2019). Spatiotemporal prediction of wildfire size extremes with bayesian finite sample maxima. *Ecological Applications*, *29*(6), e01898. https://doi.org/10.1002/eap.1898

Julian, K. D., & Kochenderfer, M. J. (2019). Distributed wildfire surveillance with autonomous aircraft using deep reinforcement learning. *Journal of Guidance, Control, and Dynamics*, *42*(8), 1768–1778. https://doi.org/10.2514/1.g004106

Kochenderfer, M. J., Wheeler, T. A., & Wray, K. H. (2022). *Algorithms for decision making*. MIT press. ISBN: 0262047012

Kuligowski, E. (2021). Evacuation decision-making and behavior in wildfires: Past research, current challenges and a future research agenda. *Fire Safety Journal*, *120*, 103129. https://doi.org/10.1016/j.firesaf.2020.103129

McCaffrey, S., Wilson, R., & Konar, A. (2017). Should I stay or should I go now? Or should I wait and see? Influences on wildfire evacuation decisions. *Risk Analysis*, *38*(7), 1390–1404. https://doi.org/10.1111/risa.12944

Pel, A., Bliemer, M., & Hoogendoorn, S. (2012). A review on travel behaviour modelling in dynamic traffic simulation models for evacuations. *Transportation*, *39*(1), 97–123. https://doi.org/10.1007/s11116-011-9320-6

Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., & Dormann, N. (2021). Stable-Baselines3: Reliable reinforcement learning implementations. *Journal of Machine Learning Research*, *22*(268), 1–8. http://jmlr.org/papers/v22/20-1364.html

Rempel, M., & Shiell, N. (2023). Using reinforcement learning to provide decision support in multi-domain mass evacuation operations. *NATO STO Review*. https://review.sto.nato.int/index.php/journal-issues/2023-fall/sas-ora-conference-2022/68-using-reinforcement-learning-to-provide-decision-support-in-multi-domain-mass-evacuation-opera

Rezaei-Shoshtari, S. (2020). *Gym Forest Fire*. https://github.com/sahandrez/gym_forestfire

Ross, W. L. (2021). Being the fire: A CNN-based reinforcement learning method to learn how fires behave beyond the limits of physics-based empirical models. *NeurIPS 2021 Workshop on Tackling Climate Change with Machine Learning*. https://www.climatechange.ai/papers/neurips2021/22

Rothermel, R. C. (1972). *A mathematical model for predicting fire spread in wildland fuels* (Vol. 115). Intermountain Forest & Range Experiment Station, Forest Service, US Department of Agriculture.

Soto, E. B. (2021). *Gym Cellular Automata*. https://github.com/elbecerrasoto/gym-cellular-automata

Subramanian, S., & Crowley, M. (2017, January). Learning forest wildfire dynamics from satellite images using reinforcement learning. *The 3rd Multi-Disciplinary Conference on Reinforcement Learning and Decision Making*. https://doi.org/10.3389/fict.2018.00006

Tapley, A., Dotter, M., Doyle, M., Fennelly, A., Gandikota, D., Smith, S., Threet, M., & Welsh, T. (2023). *Reinforcement learning for wildfire mitigation in simulated disaster environments*. https://arxiv.org/abs/2311.15925

Towers, M., Terry, J. K., Kwiatkowski, A., Balis, J. U., Cola, G. de, Deleu, T., Goulão, M., Kallinteris, A., KG, A., Krimmel, M., Perez-Vicente, R., Pierré, A., Schulhoff, S., Tai, J. J., Shen, A. T. J., & Younis, O. G. (2023). *Gymnasium*. https://doi.org/10.5281/zenodo.8127026

Viseras, A., Meissner, M., & Marchal, J. (2021). Wildfire Front Monitoring with Multiple UAVs using Deep Q-Learning. *IEEE Access*, 1–1. https://doi.org/10.1109/ACCESS.2021.3055651

Wang, H., Mostafizi, A., Cramer, L. A., Cox, D., & Park, H. (2016). An agent-based model of a multimodal near-field tsunami evacuation: Decision-making and life safety. *Transportation Research Part C: Emerging Technologies*, *64*, 86–100. https://doi.org/10.1016/j.trc.2015.11.010

Xu, N., Lovreglio, R., Kuligowski, E., Cova, T., Nilsson, D., & Zhao, X. (2023). Predicting and assessing wildfire evacuation decision-making using machine learning: Findings from the 2019 Kincade Fire. *Fire Technology*. https://doi.org/10.1007/s10694-023-01363-1