School of Electrical Engineering and Computer Science

# Fast Simulation for Computational Sustainability Sequential Decision Making Problems

Sean McGregor, Rachel Houtman, Hailey Buckingham, Claire Montgomery, Ronald Metoyer, and Thomas G. Dietterich


## Abstract

Solving sequential decision making problems in computational sustainability often requires simulators of ecology, weather, fire, or other complex phenomena. The extreme computational expense of these simulators stymie optimization and interactive visualization of decision rules (policies). This work presents our results in creating an interactive visualization for a wildfire management problem whose simulator normally takes several hours to run. We successfully generate visualizations for a landscape's development over 100 year time spans within 3 seconds, when the original simulator took several hours.


## Markov Decision Processes

- A theoretical formulation for *sequential decision making subject to uncertainty*
  - ✧ Wildfire management[1]
  - ✧ Timber harvest planning
  - ✧ River flow management
  - ✧ Invasive species eradication[2]

More formally, a Markov Decision Process is

| | |
|---|---|
| *S* | All States of the World |
| $P_0$ | Starting State Distribution |
| *A* | Available Actions |
| *R(s, a)* | Rewards |
| $\gamma \in (0, 1)$ | Discount |
| *P* | State Transition Probability |
| $\pi(s) \rightarrow a$ | Policy |

A wildfire suppression Markov Decision Process is

| | |
|---|---|
| *S* | All tree and weather configurations |
| $P_0$ | A snapshot of the current forest, with a random fire |
| *A* | Suppress or let-burn |
| *R(s, a)* | Timber harvest, Suppression Expense |
| $\gamma \in (0, 1)$ | 0.96 (Forest Service Standard) |
| *P* | Several Simulators |
| $\pi(s) \rightarrow a$ | Suppress all fires |

## Goal: Visualize and Optimize Decision Rules

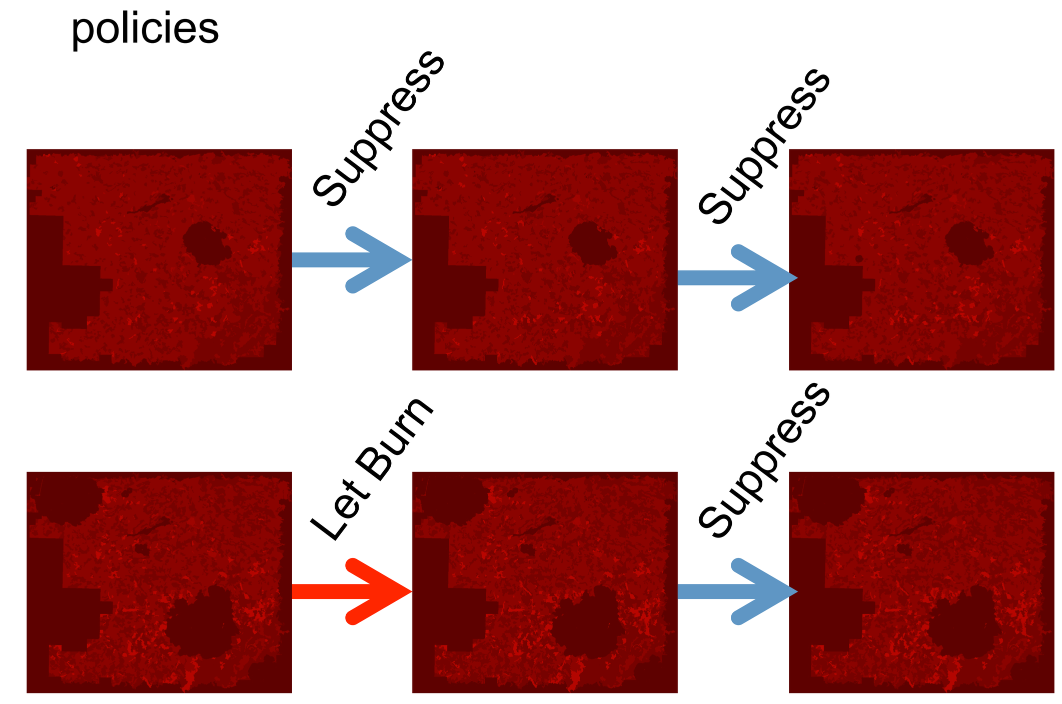
- Wildfire: given a wildfire on timber producing lands, how do we balance suppression costs, timber revenues, and ecological services when deciding to suppress a fire or let it burn?

## Problem: Simulating Nature is Computationally Expensive

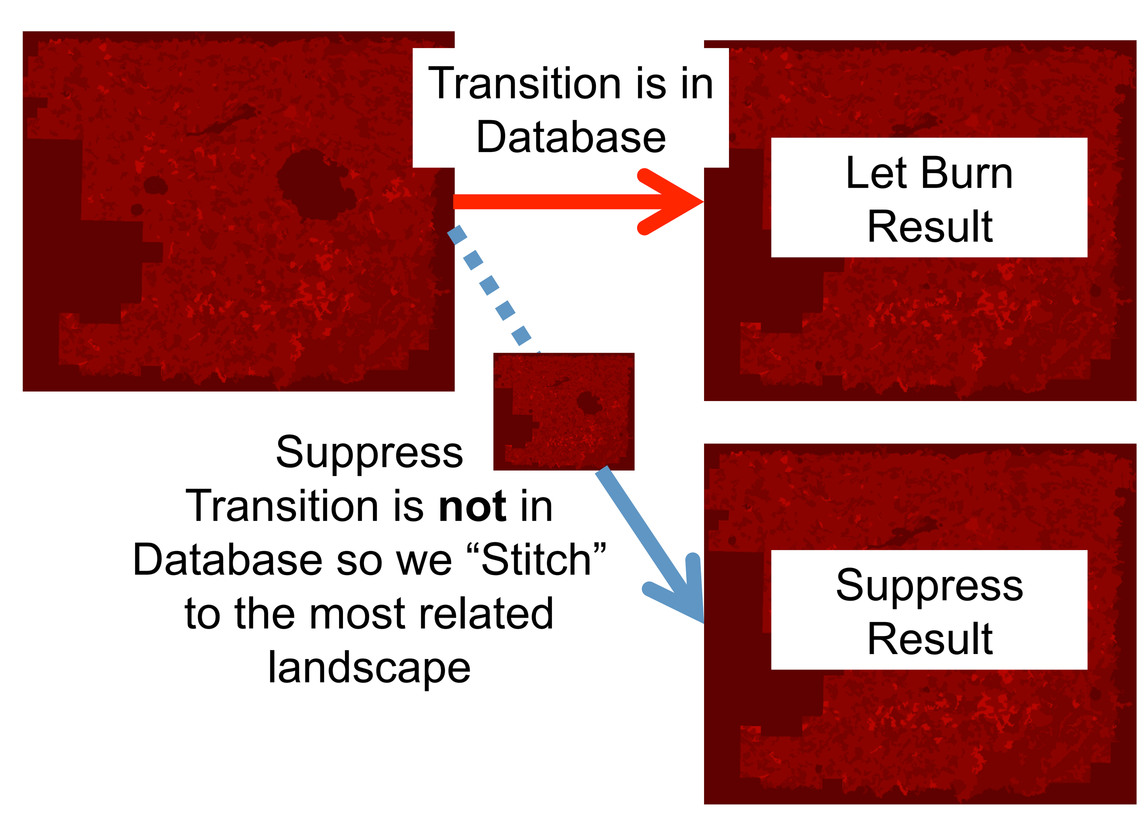
Simulating ecological processes over many decades requires models for weather, climate, fire spread, human encroachment, succession, and more. These models can take hours or days to complete a single scenario!

## Solution: Synthesize Trajectories from a Database

1. Generate a dataset of simulations from many different policies

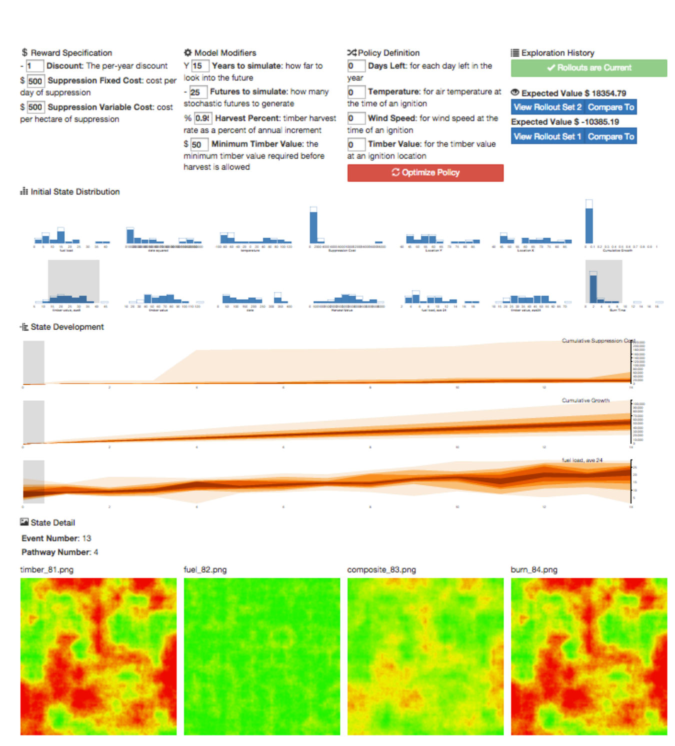



2. When visualizing or optimizing a policy that has not been sampled, use state similarity to "stitch" states together into complete trajectories




3. Visualize[4,5] or optimize based on the generated trajectories

MDPVis.github.io




## State Variables in Computational Sustainability Domains

We model variables in the database differently based on whether they are **persistent** or **exogenous**. Persistent variables are highly correlated from one time step to another, but exogenous variables are independent and identically distributed within every time step.

### Persistent

- Plant cover
- Fuel levels
- Species presence/ absence
- Elevation, latitude, and longitude

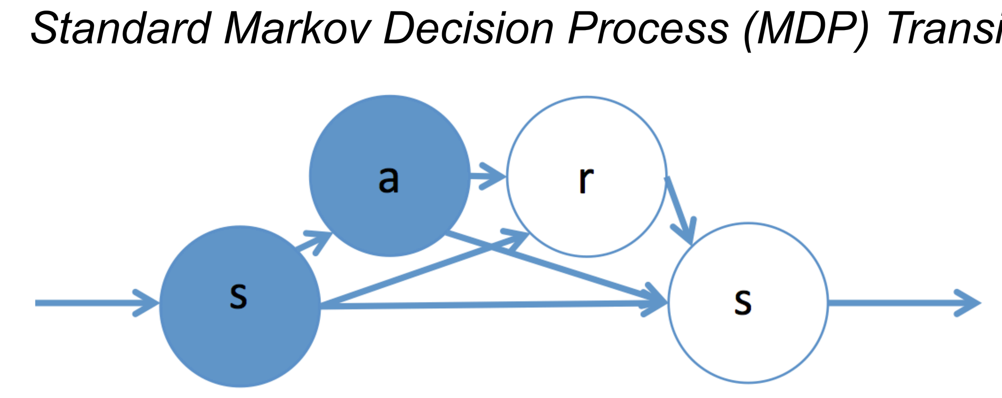
### Exogenous

- Weather events
- Wildfire Ignitions
- Invasive species introduction
- Timber prices

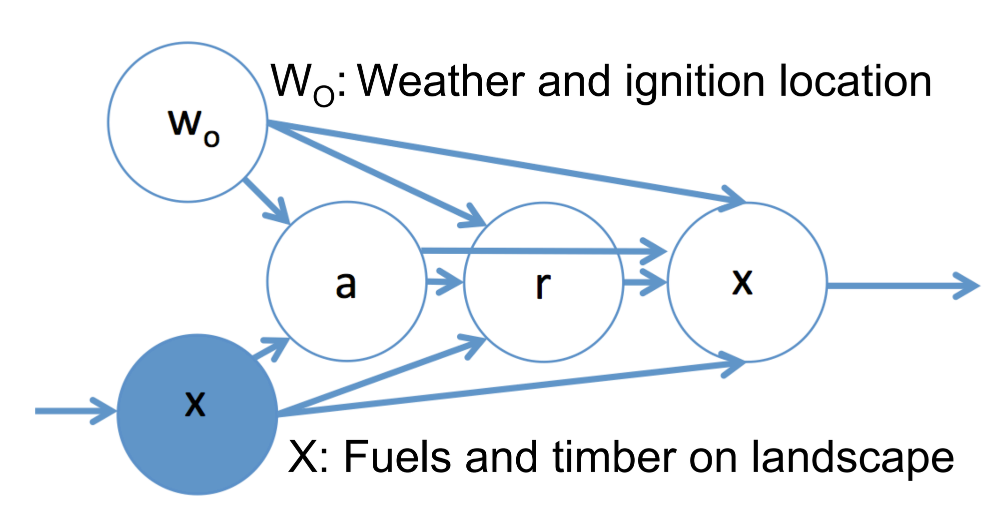
## How do We Model Persistent and Exogenous Variables?

We stitch to a state in the database if it is similar to the state we are in. **Similarity does not need to include exogenous variables!** Traditionally, similarity is measured against the complete state and action as highlighted below. In our version of trajectory synthesis, we separate the action and exogenous variables and stitch based solely on the persistent state.

*Standard Markov Decision Process (MDP) Transition*




*Our Trajectory Synthesis Transition*




The probabilistic graphical models shown above factorizes the state such that we can stitch states based solely on the configuration of a **persistent state**. We don't need to consider similarity of exogenous variables like weather!

## Conclusion

We can visualize and optimize policies for computationally expensive sustainability domains with a database of state transitions whose computational cost is independent of the computational cost of the modeled phenomena.


References

1: Houtman, R. M., Montgomery, C. A., Gagnon, A. R., Calkin, D. E., Dietterich, T. G., McGregor, S., & Crowley, M. (2013). Allowing a Wildfire to Burn: Estimating the Effect on Future Fire Suppression Costs. International Journal of Wildland Fire, 22(7), 871–882.
2: Dietterich, T., Taleghan, M., & Crowley, M. (2013). PAC Optimal Planning for Invasive Species Management: Improved Exploration for Reinforcement Learning from Simulator-Defined MDPs. Twenty-Seventh AAAI Conference on Artificial Intelligence.
3: Fonteneau, R., Murphy, S. a, Wehenkel, L., & Ernst, D. (2013). Batch Mode Reinforcement Learning based on the Synthesis of Artificial Trajectories. Annals of Operations Research, 208(1), 383–416.
4: McGregor, S., Buckingham, H., Dietterich, T. G., Houtman, R., Montgomery, C., & Metoyer, R. (2015). Facilitating Testing and Debugging of Markov Decision Processes with Interactive Visualization. In IEEE Symposium on Visual Languages and Human-Centric Computing. Atlanta.
5: McGregor, S., Buckingham, H., Houtman, R., Montgomery, C., Metoyer, R., & Dietterich, T. G. (2015). MDPvis: An Interactive Visualization for Testing Markov Decision Processes. In AAAI Fall Symposium on Sequential Decision Making for Intelligent Agents.


Oregon State UNIVERSITY