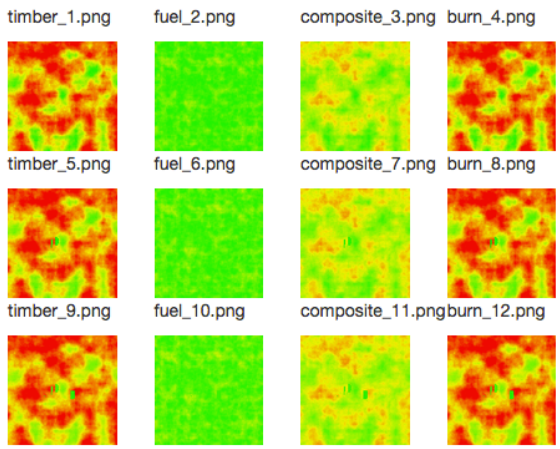
COLLEGE OF ENGINEERING

Facilitating Testing and Debugging of Markov Decision Processes with Interactive Visualization

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

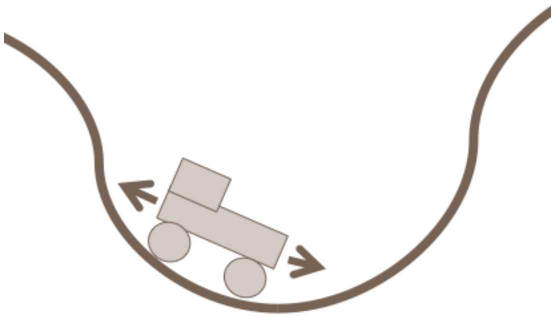
What are Markov Decision Processes (MDPs)?

Sequential Decision Making Under Uncertainty



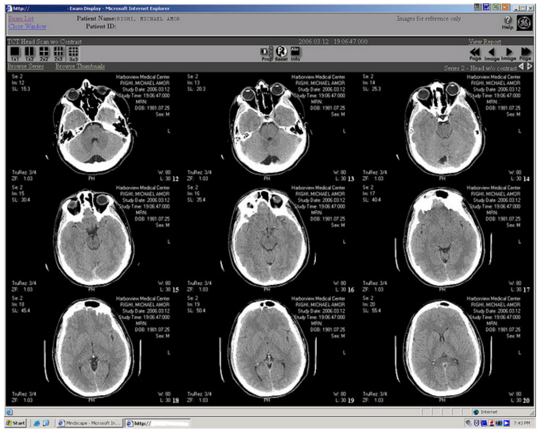
Wildfire Suppression

Autonomous Helicopter[0]



Mountain Car

Logistics[1]



Medical Diagnosis[2]

Outline

1. **Markov Decision Processes (MDPs)**
   - Basic Introduction
   - Testing
2. MDPvis
   - Design
   - Testing Examples
   - MDPvis Use Case Study
3. Concluding

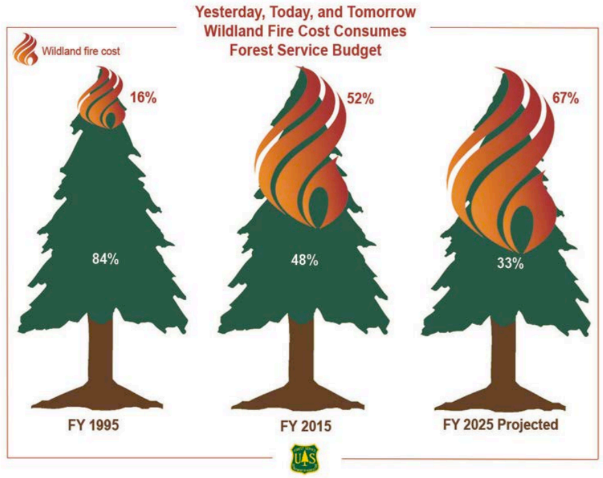
Notation, M = $\langle S, A, P, R, \gamma, P_0 \rangle$

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 Probabilities (Simulators)
$\pi(s) \rightarrow a$	Policy

Puterman, M. (1994). Markov Decision Processes: Discrete Stochastic Dynamic Programming (1st ed.). Wiley-Interscience.

Motivating Domain of Wildfire

Starting in 1935, the United States adopted the "**10 AM policy**"



We need a more nuanced approach.

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.

http://www.fs.fed.us/sites/default/files/2015-Fire-Budget-Report.pdf

Modeling Wildfire

S	All the possible configurations of trees/ignitions
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

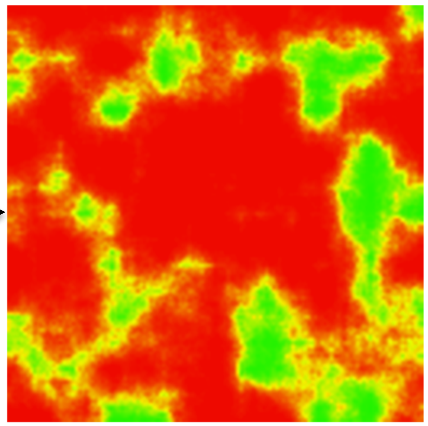
Represents a challenging and more general class of MDPs

- High Dimensional States
- Large State Space
- Integrates Several Simulators

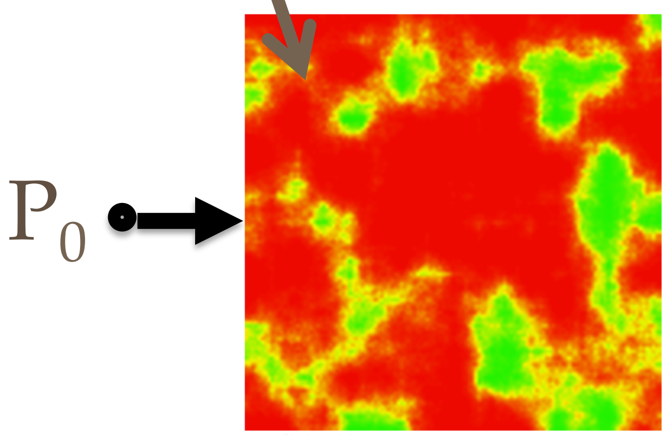
P_0

Start with Today's Landscape

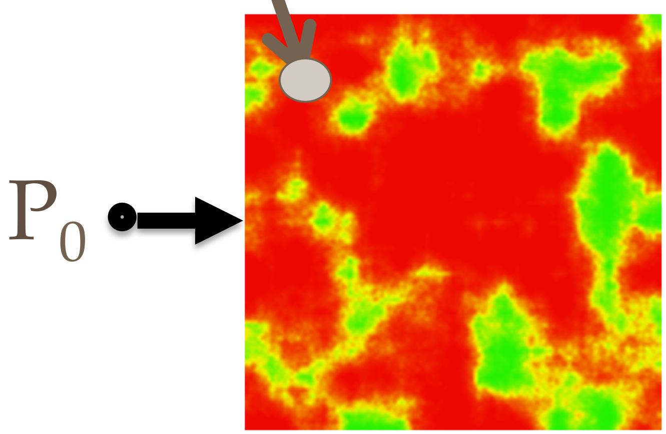
P_0



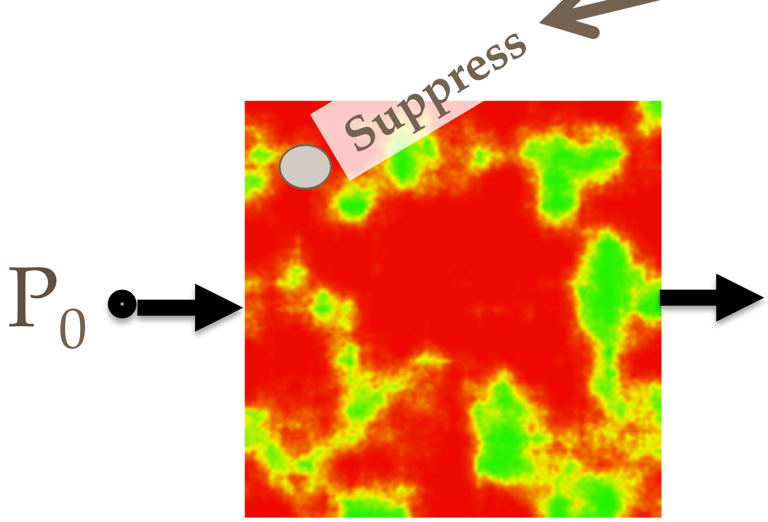
Generate an ignition and weather



Generate an ignition and weather



Select an Action

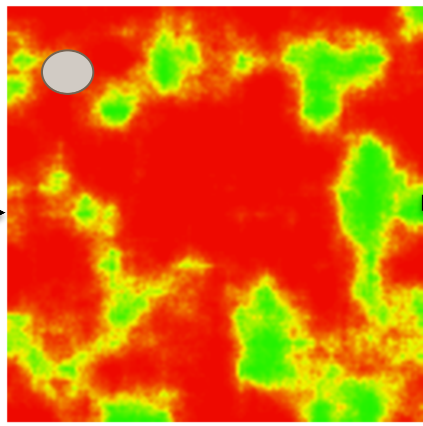


Fire Suppression Effort

Fire Suppression Costs

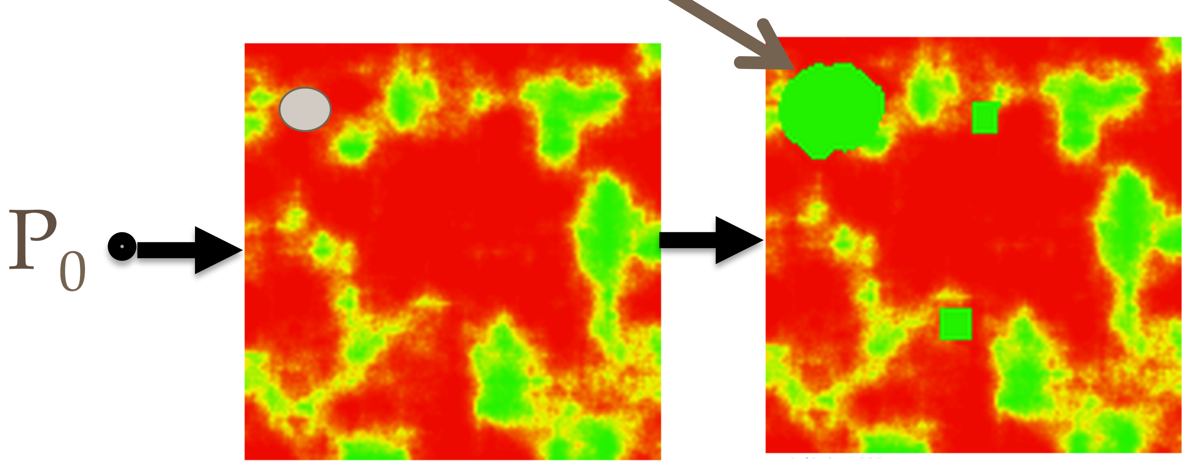
$(95,000)

P_0



OSU Oregon State UNIVERSITY

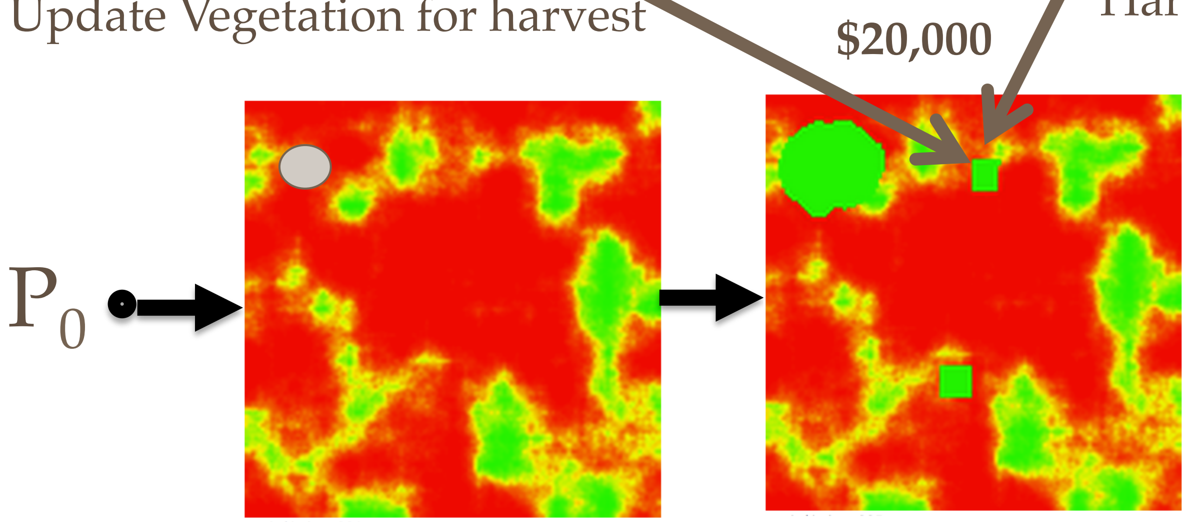
Update Vegetation for Wildfire



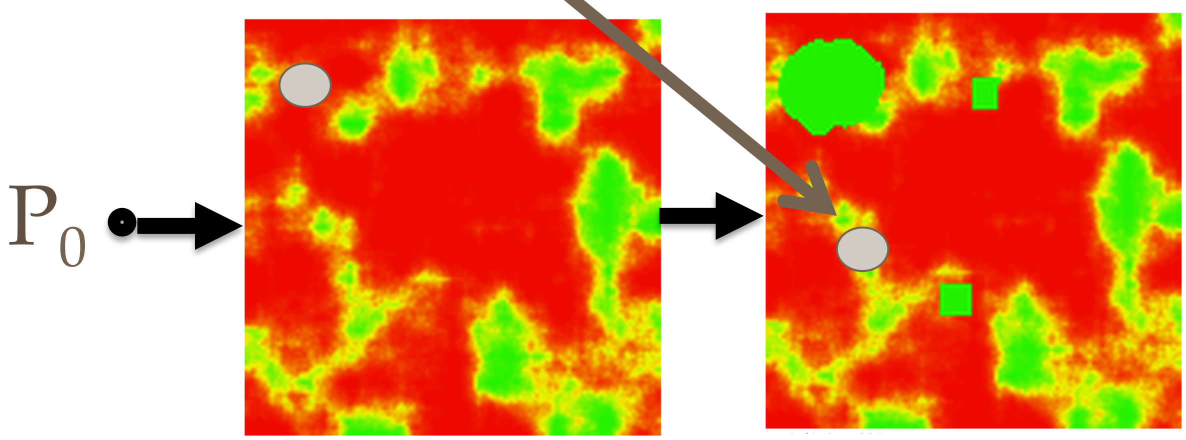
Oregon State University

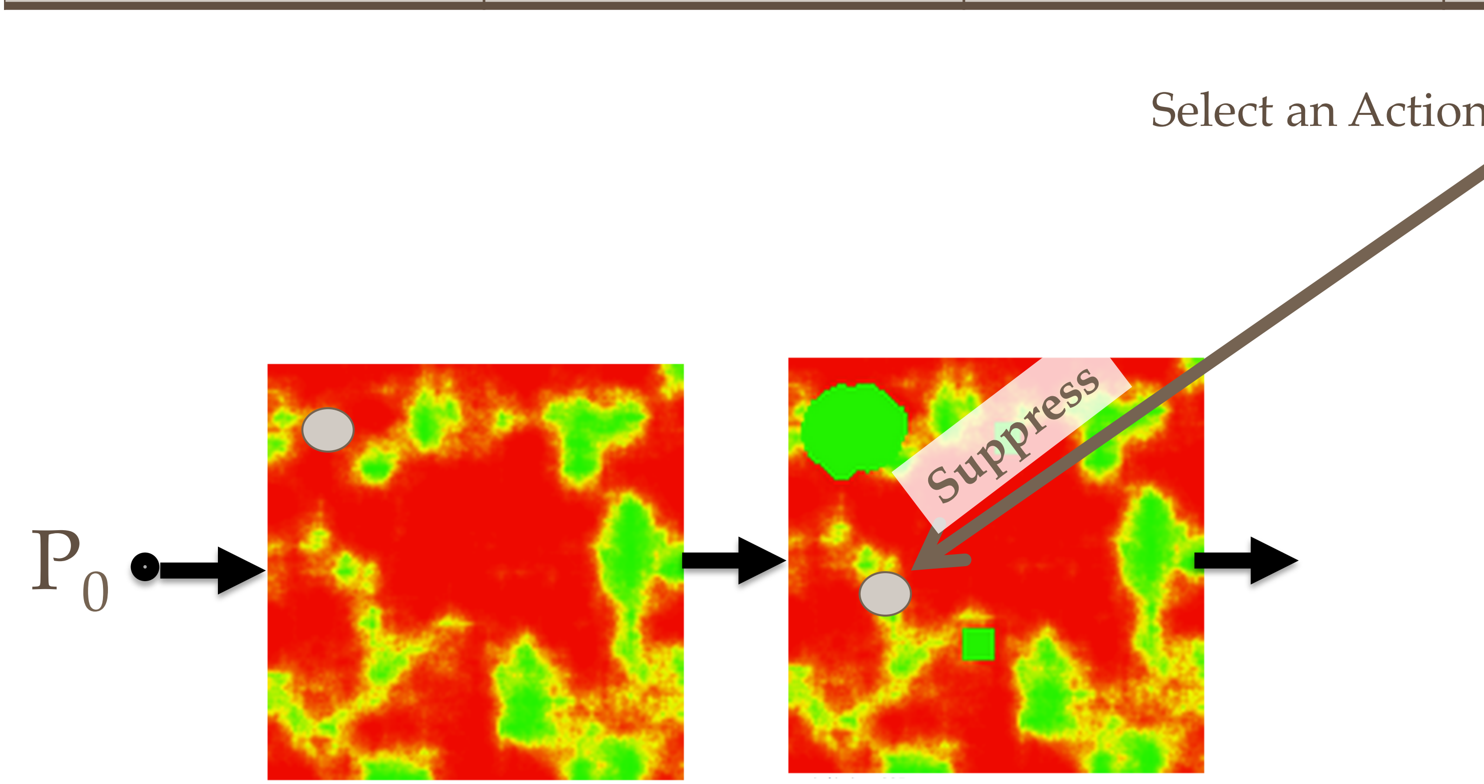
Update Vegetation for harvest

Harvest Revenue



Generate an ignition and weather

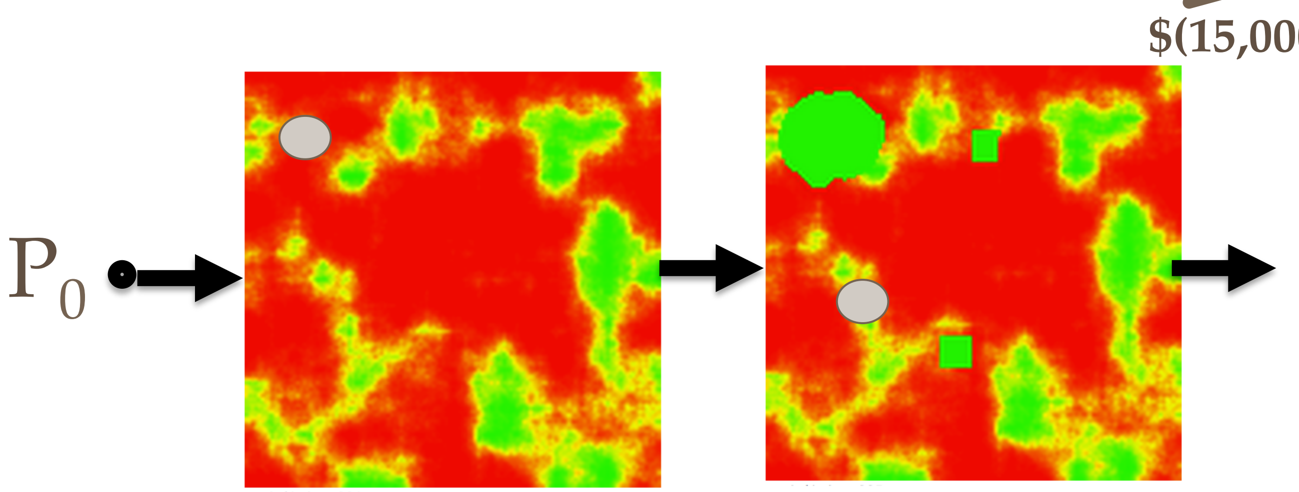




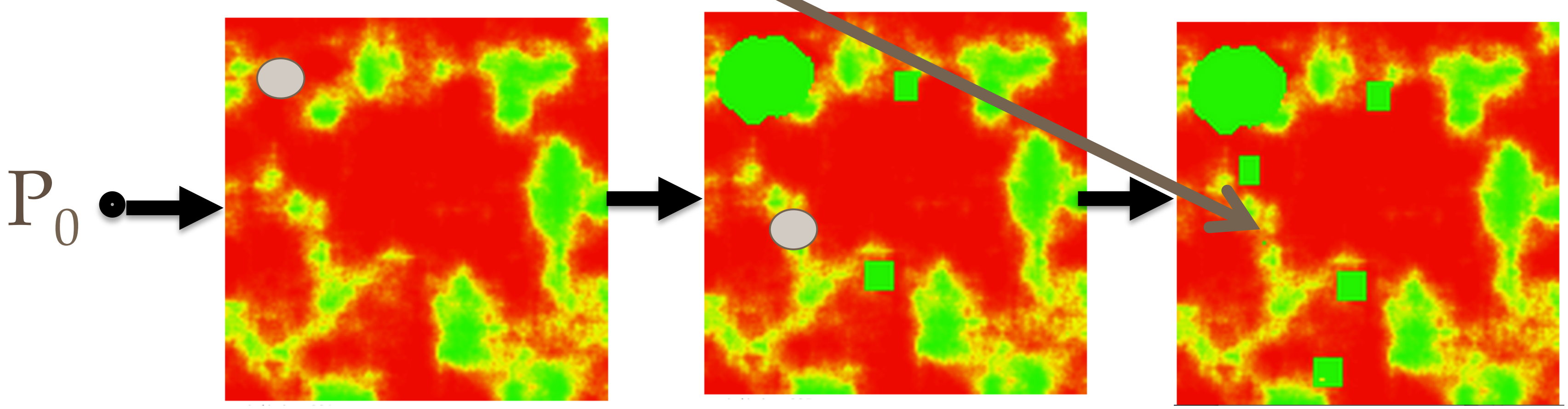
Oregon State University

Fire Suppression Effort

Fire Suppression Costs

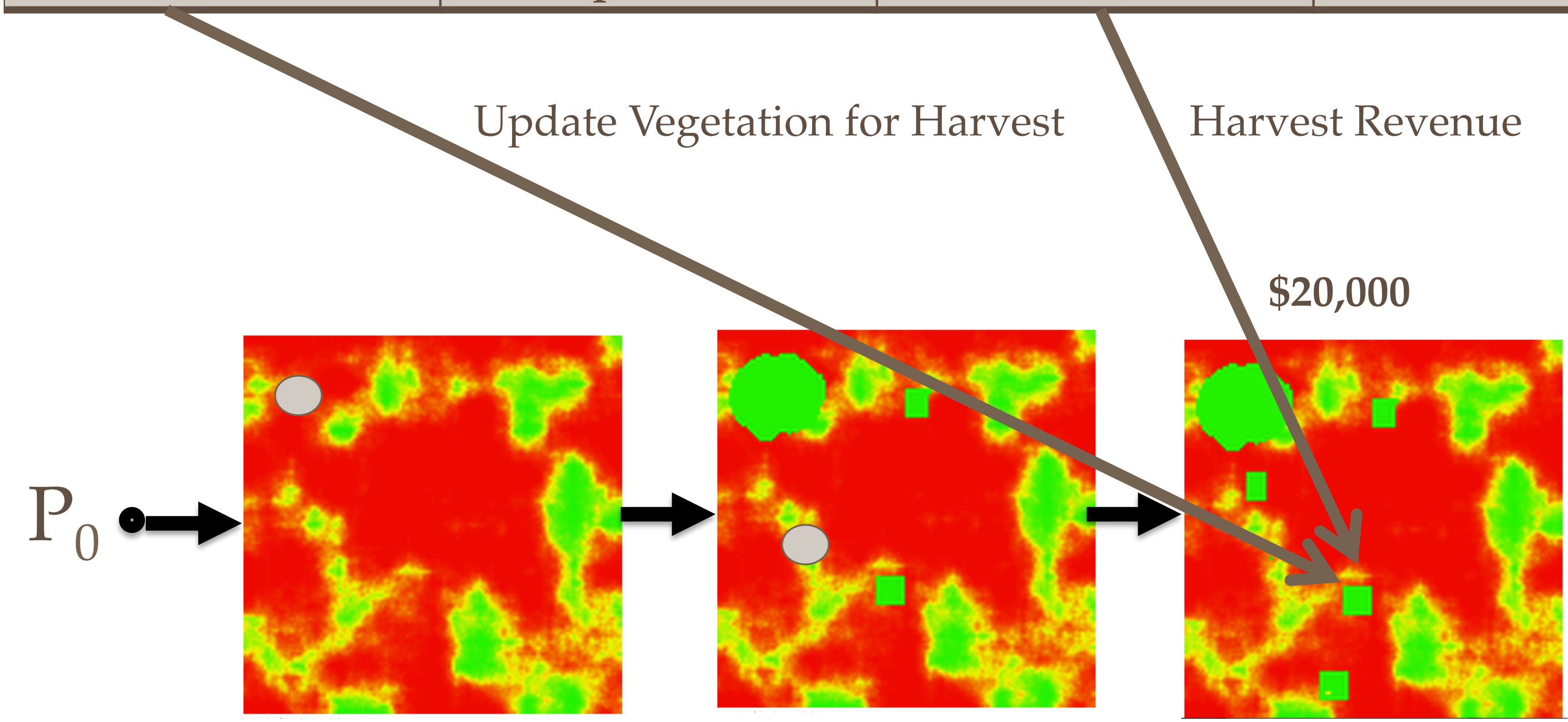


Update Vegetation for Wildfire

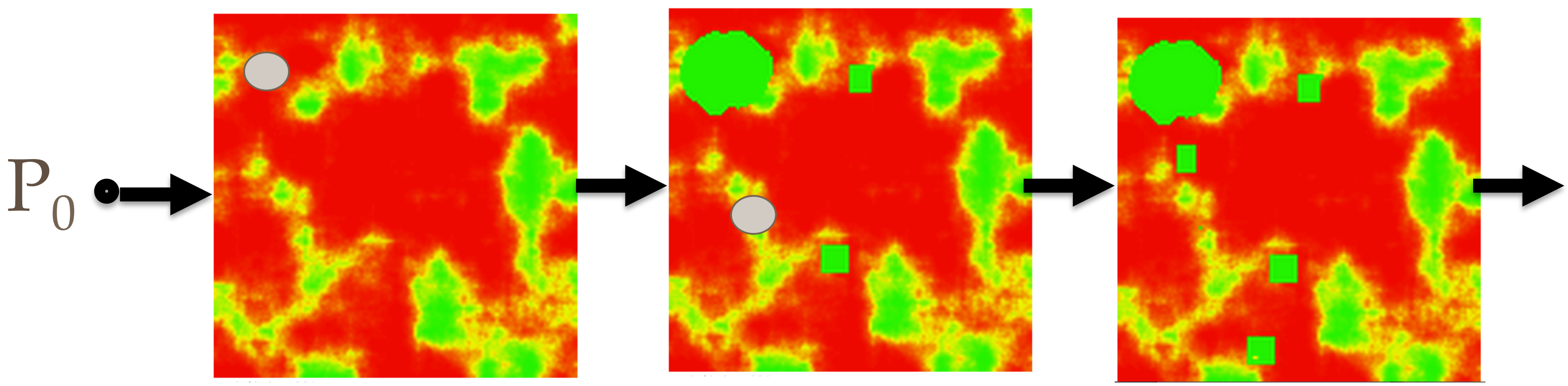


Oregon State University

Simulators	Optimizer	Rewards	Policy

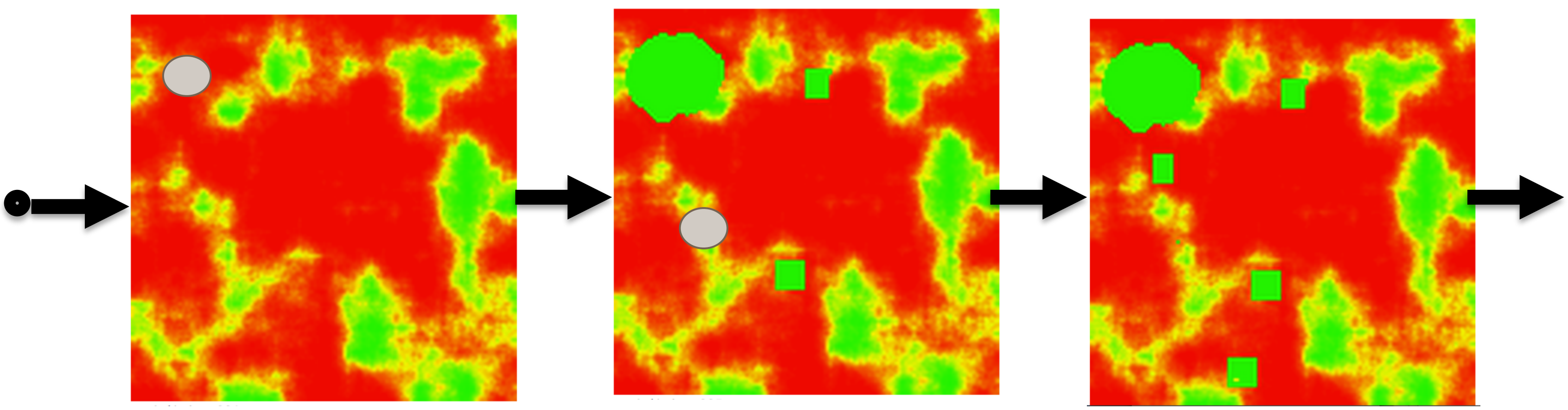


(Continue Until Reaching the Horizon)



A High Dimensional Probabilistic Time Series

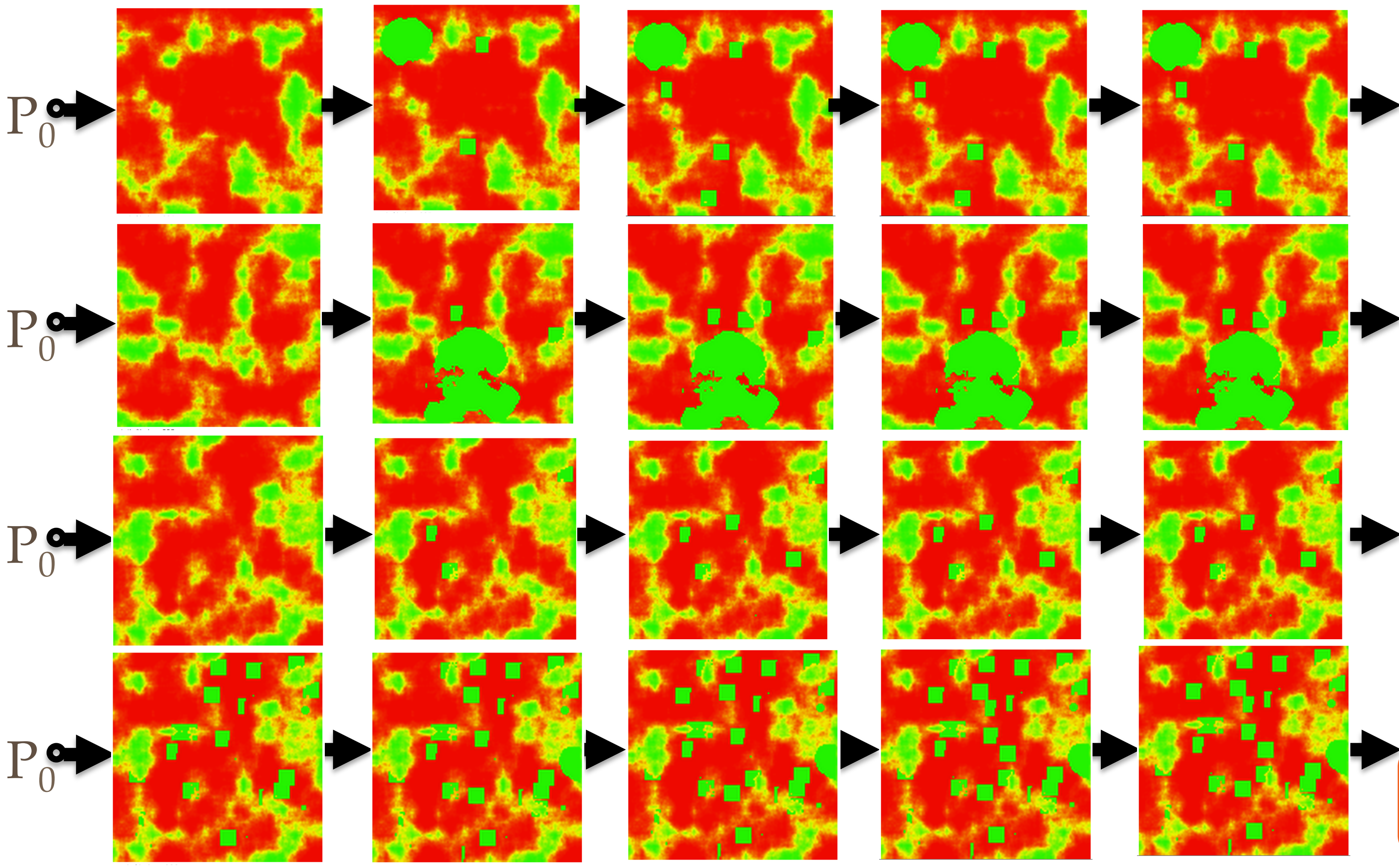
P_0



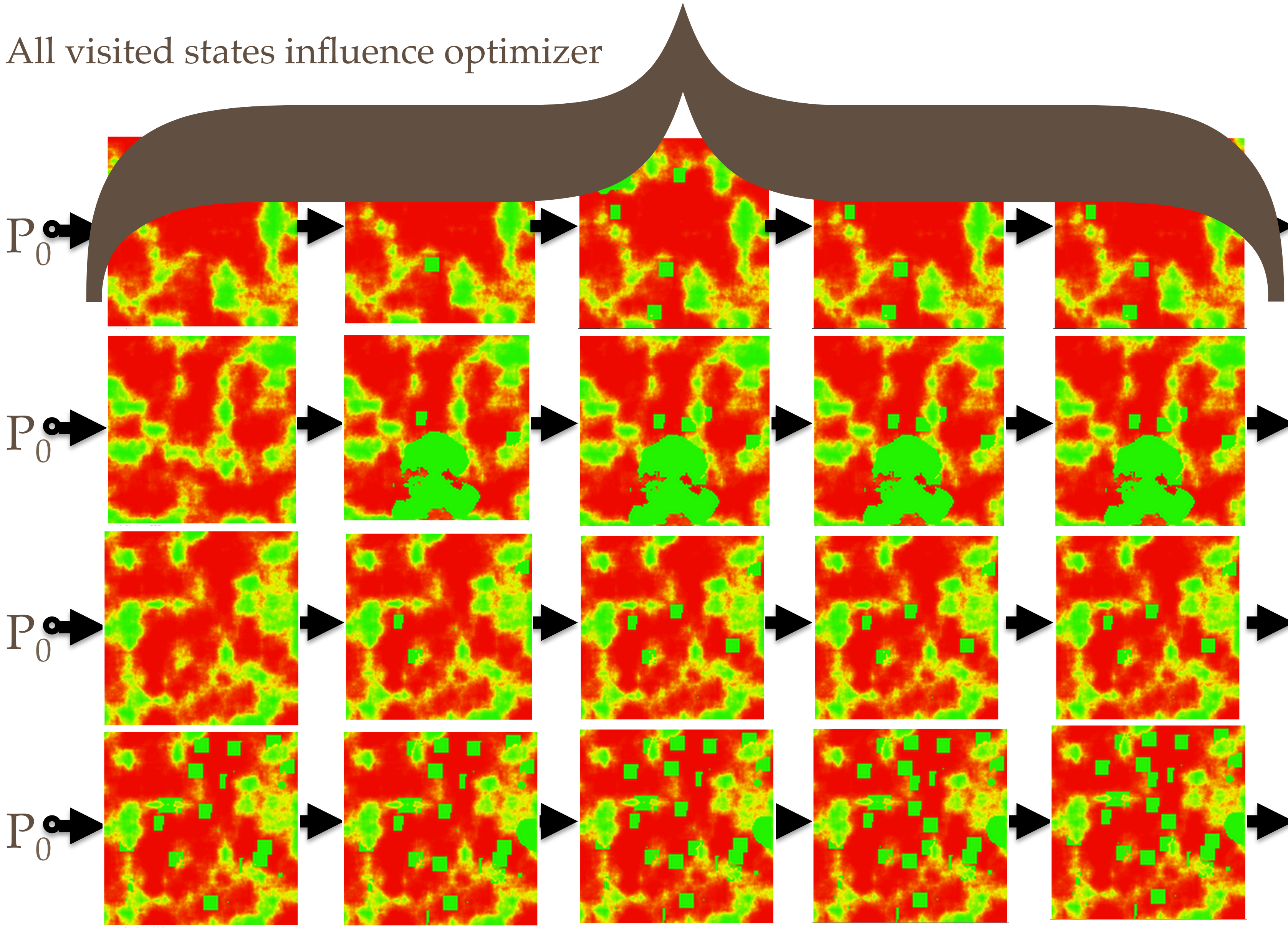
…And this is just one of many!

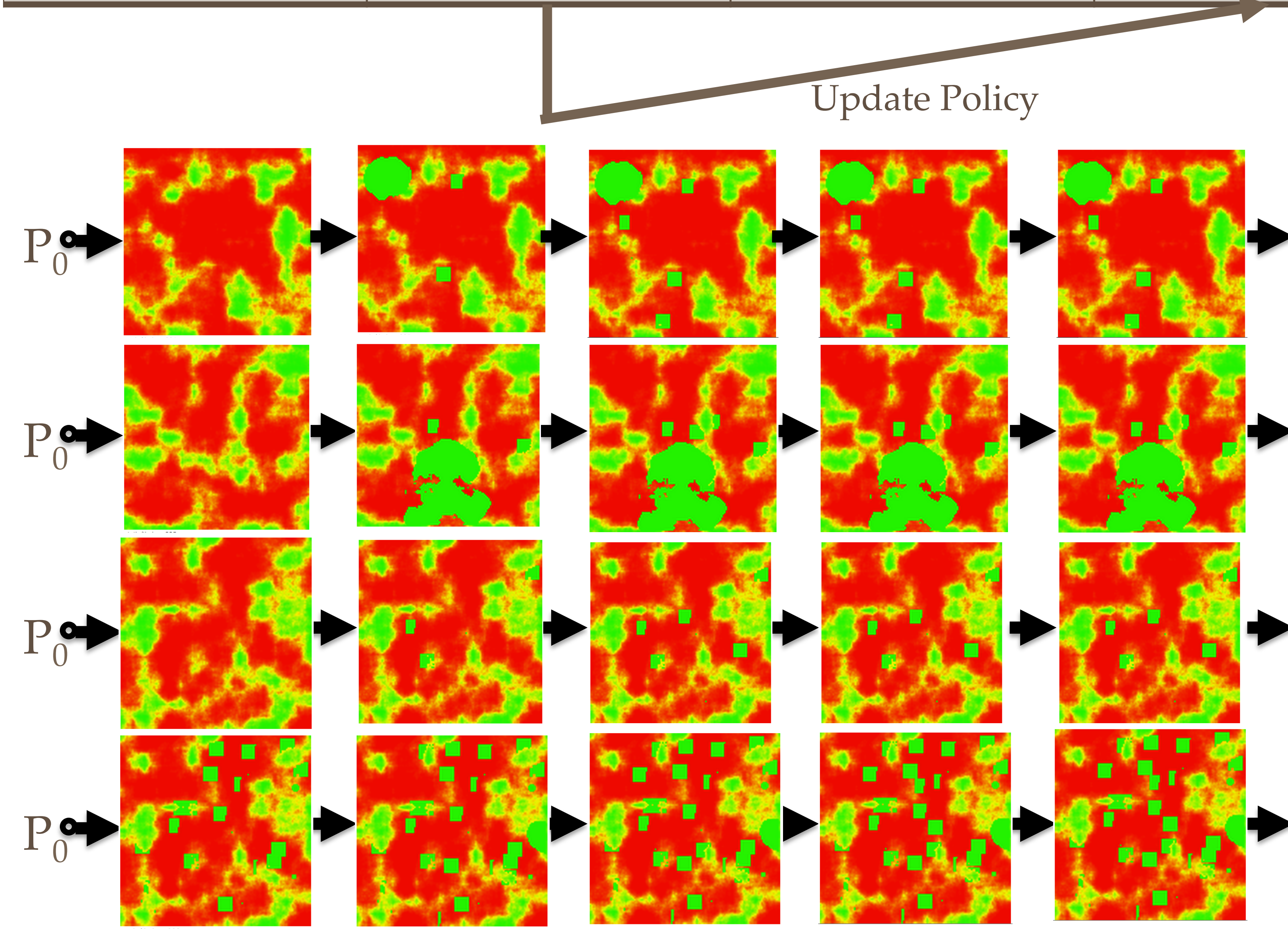
Oregon State University

Monte Carlo Rollouts

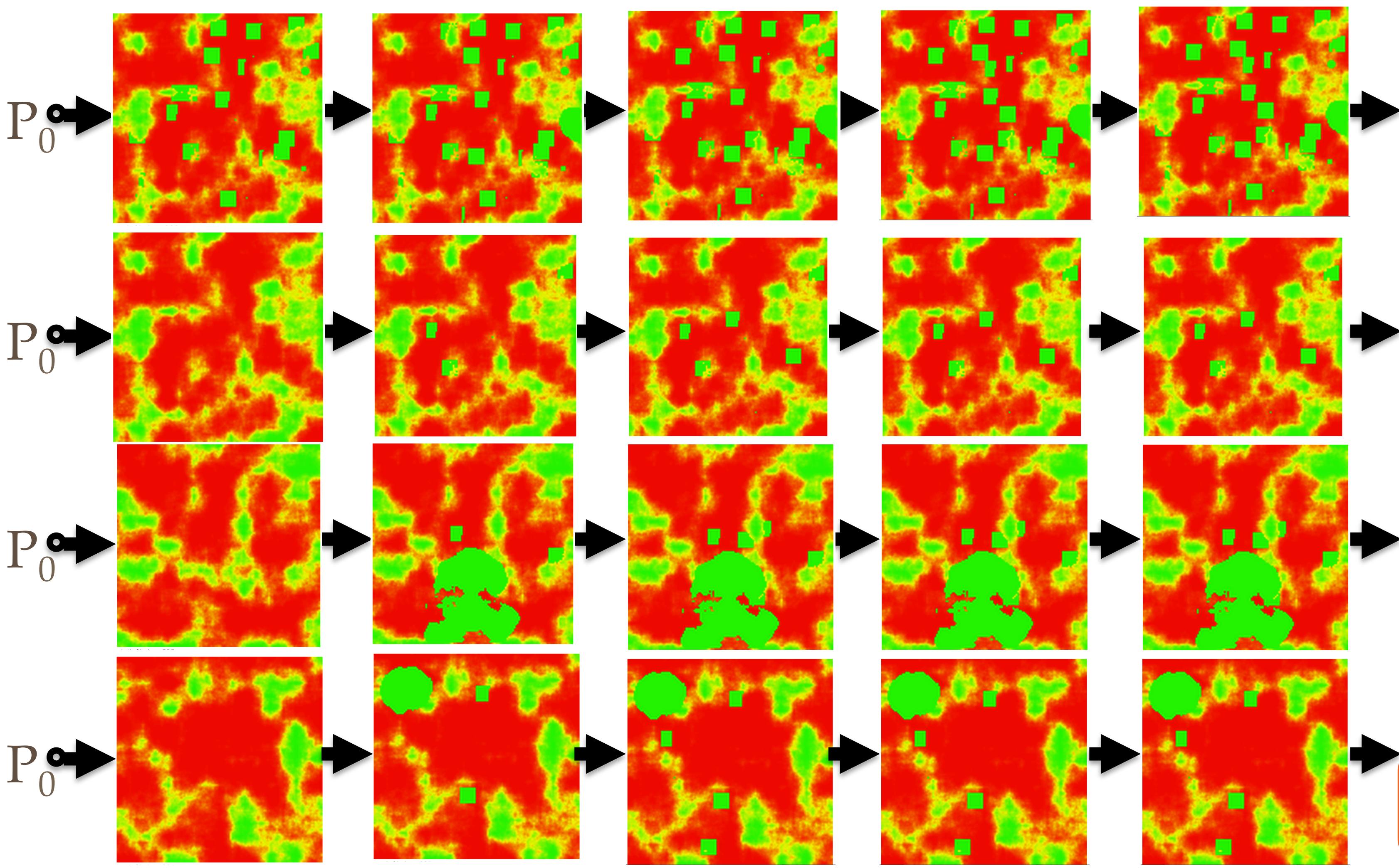


All visited states influence optimizer




Update Policy
P_0
P_0
P_0
P_0

The Rollout Distribution Changes!



MDP Testing Challenges

- Bugs are **probabilistically** expressed in a **high dimensional temporal dataset.**
- The **dataset changes** with changes to parameters.
- The **optimizer sees more of the state and policy space** than the user.

Testing requires exploring rollouts and parameters

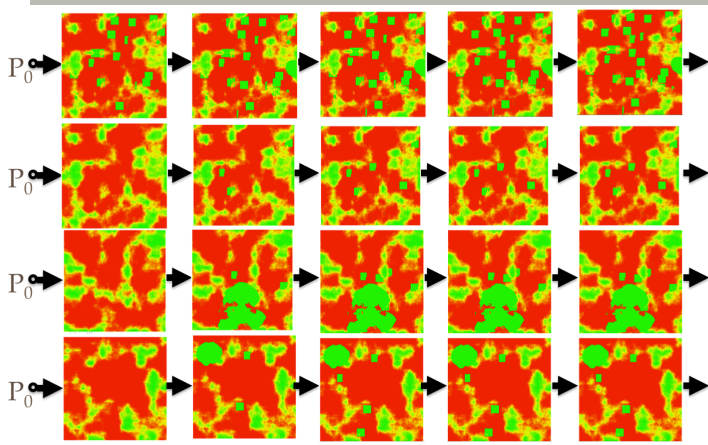
MDP Debugging and Fault Isolation

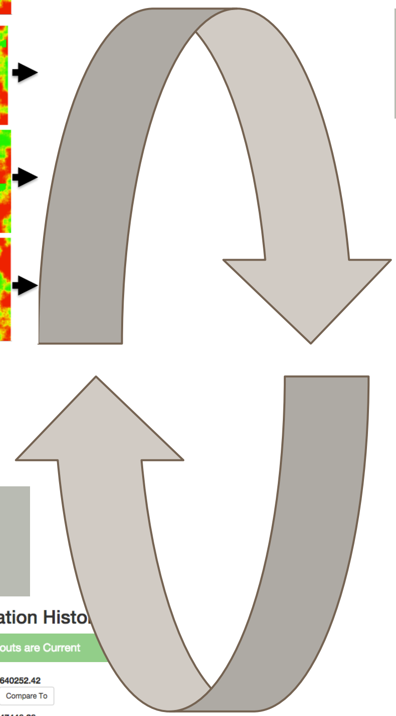
- Deactivate/modify components to isolate fault
  - e.g. Balance reward magnitude and frequency

Debug MDP specification and integration with **parameter changes**

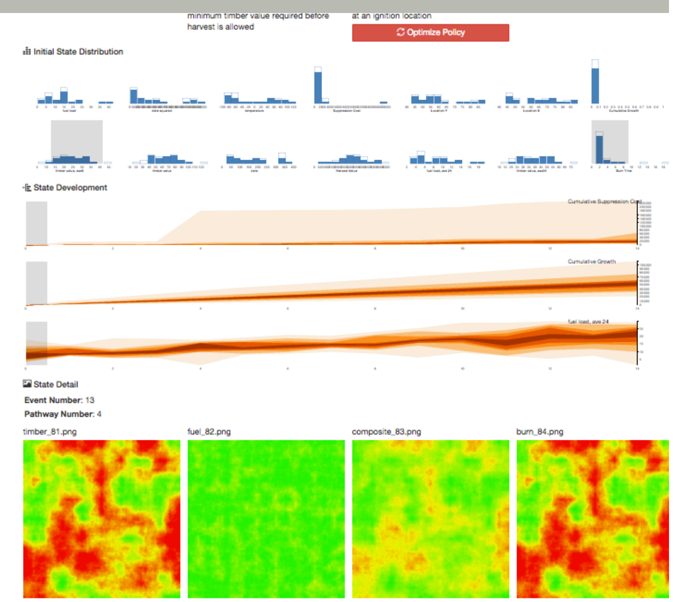
Testing and Debugging Process

1. Generate Rollouts

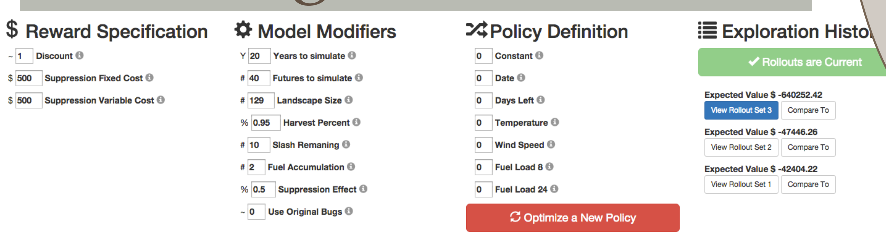




2. Visualize the data



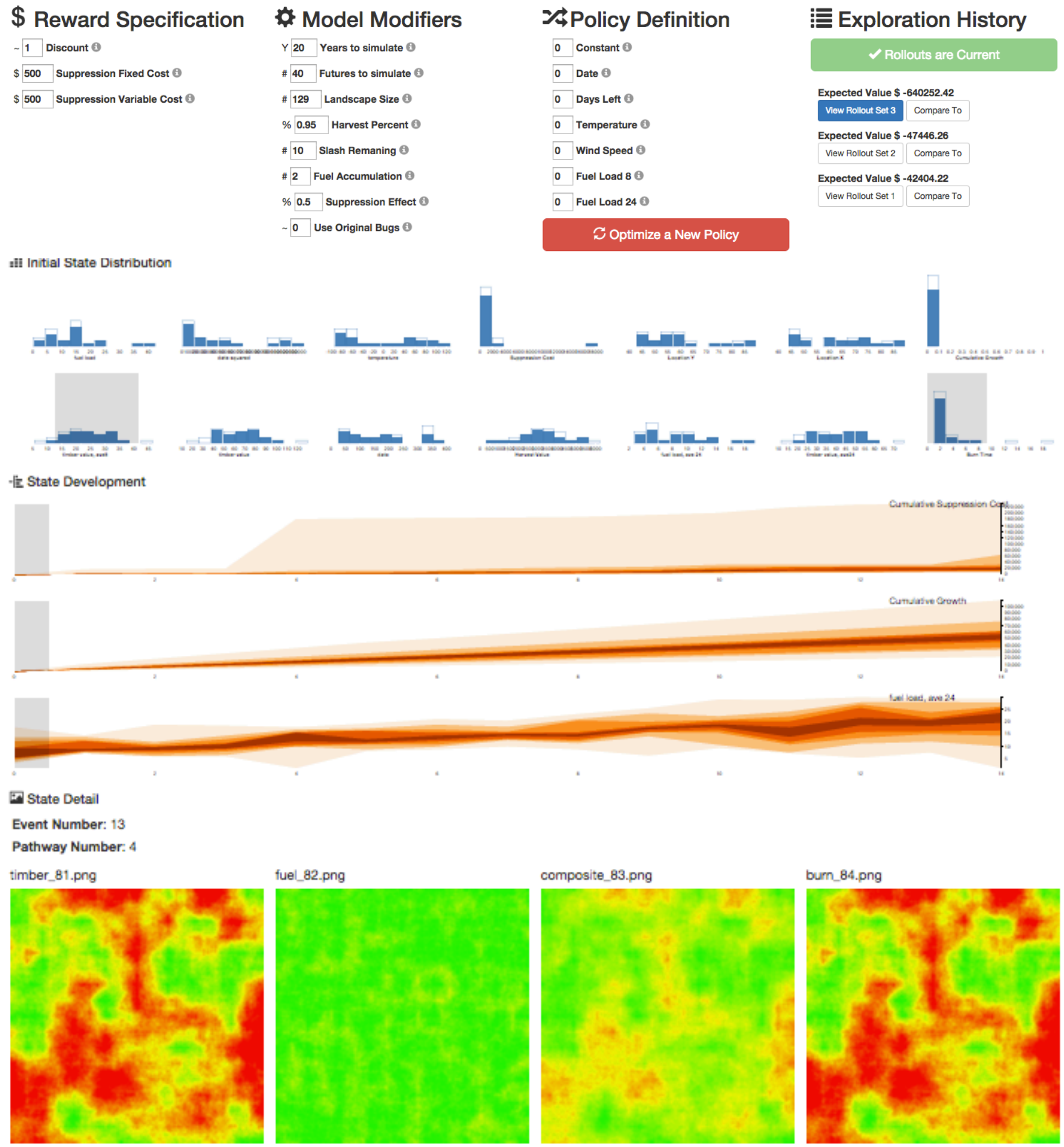
3. Change Parameters



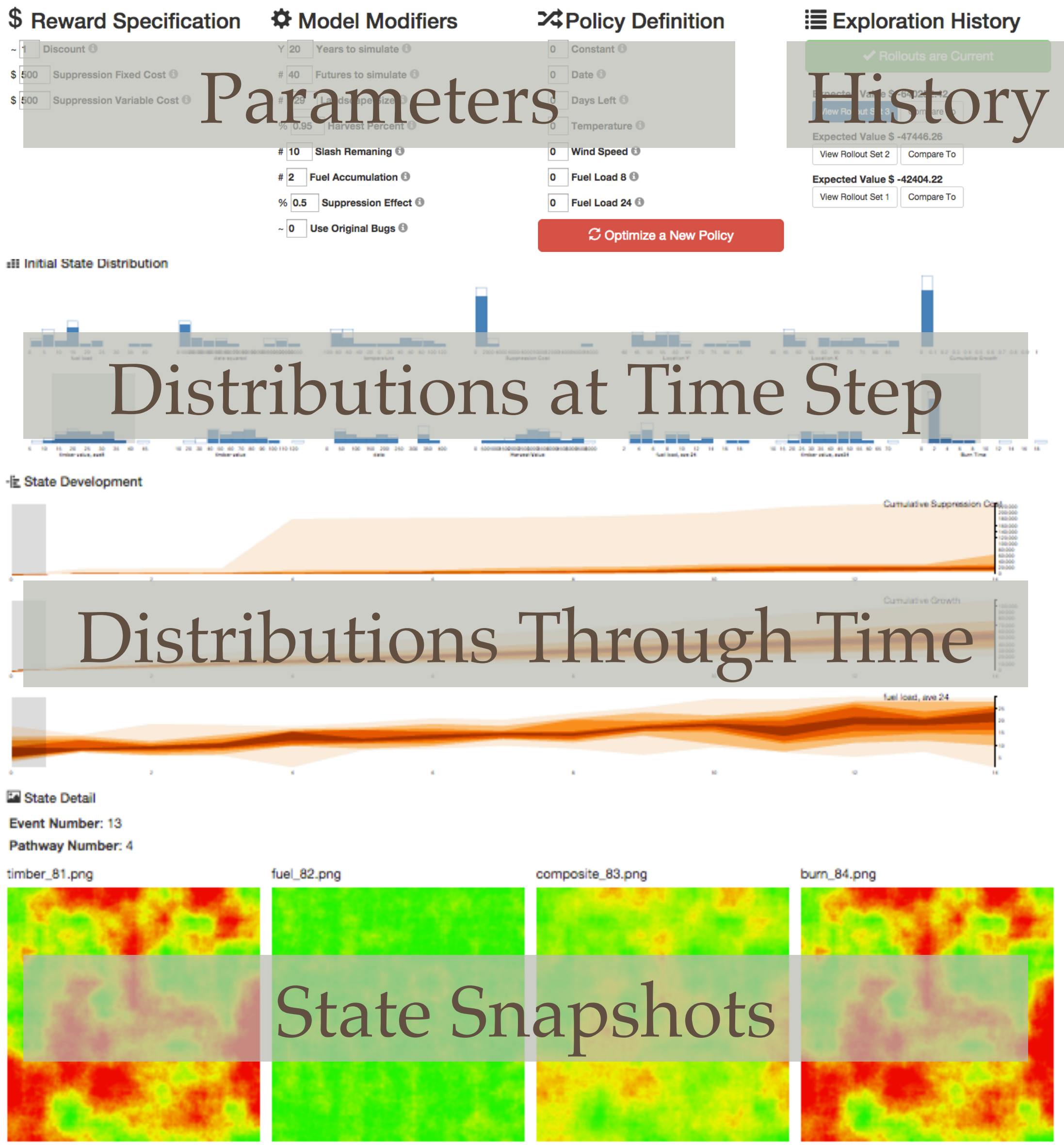
Outline

1. Markov Decision Processes (MDPs)
   - Basic Introduction
   - Testing
2. **MDPvis**
   - Design
   - Testing Examples
   - MDPvis Use Case Study
3. Concluding

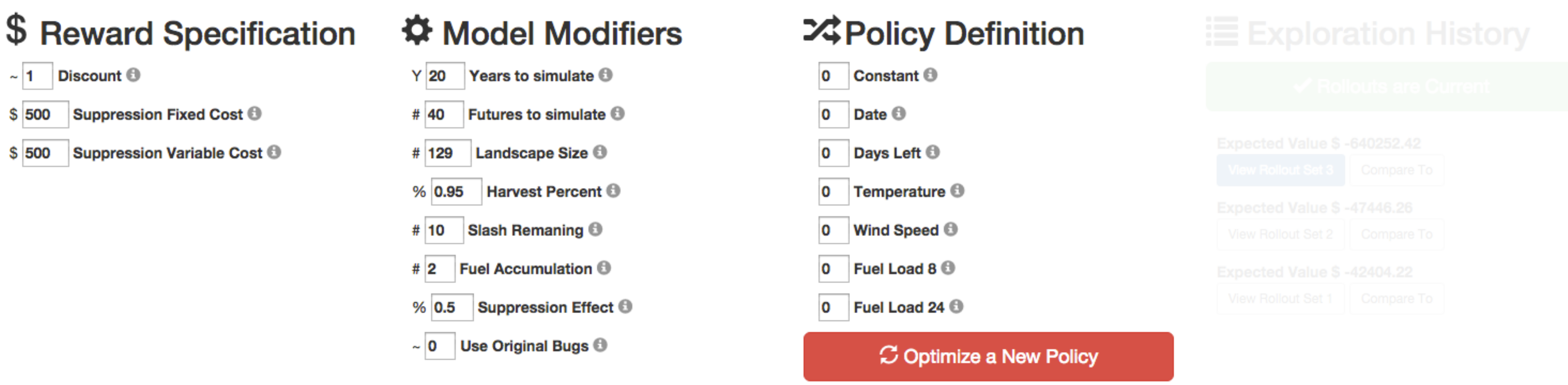
Introducing MDPvis



What are the elements of the MDPvis design?

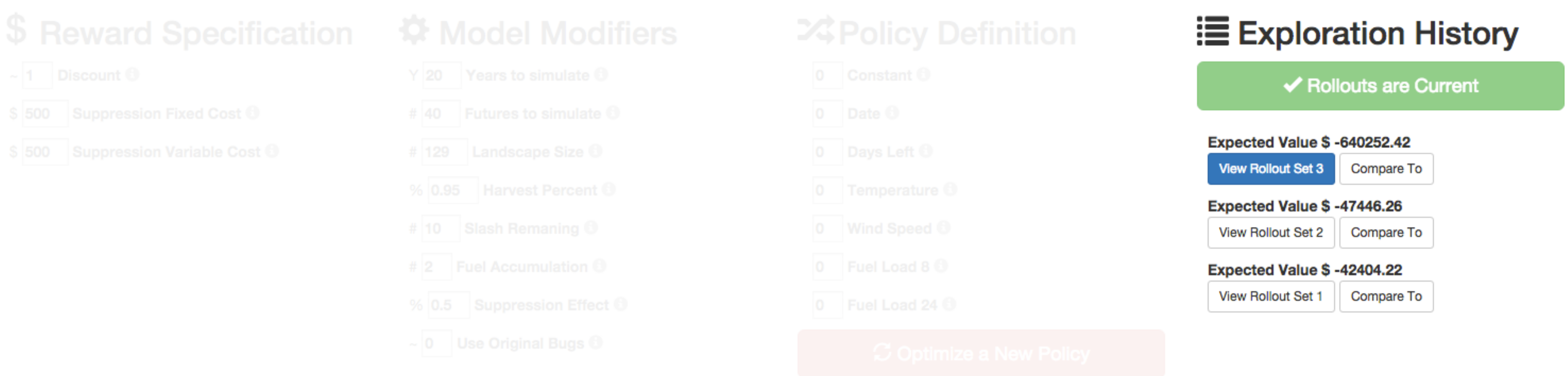


Parameter Areas



Oregon State UNIVERSITY

History Area

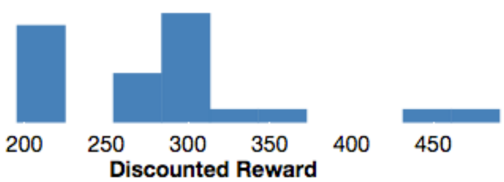


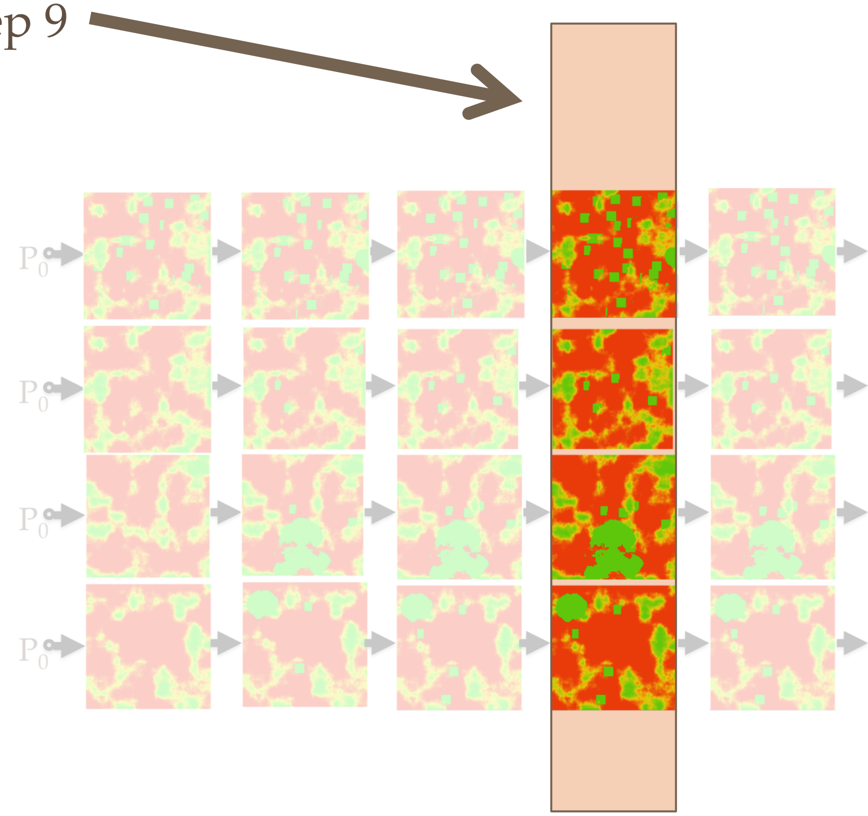
Oregon State UNIVERSITY

Visualization Areas

State Variable Distributions for a Fixed Time Step

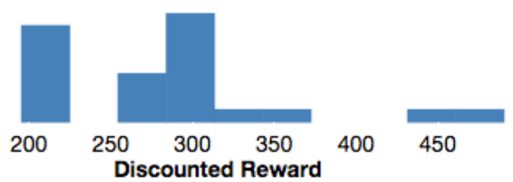
Time step 9



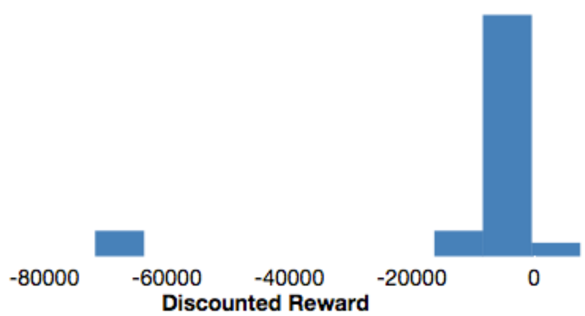


Oregon State University

State Variable Distributions for a Fixed Time Step



π_1: Let-Burn

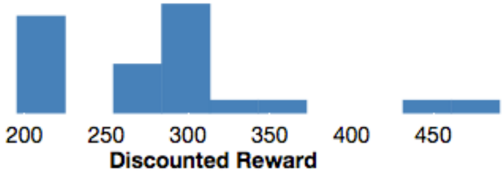


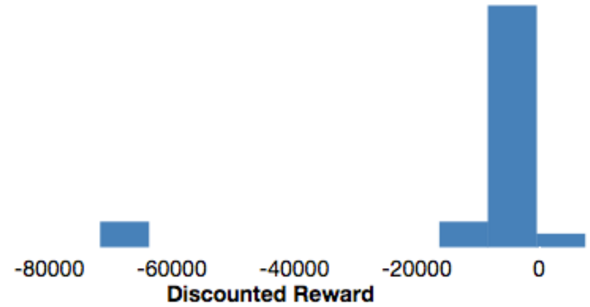
π_2: Suppress-All

Comparison

$\pi_1 - \pi_2$

State Variable Distributions for a Fixed Time Step

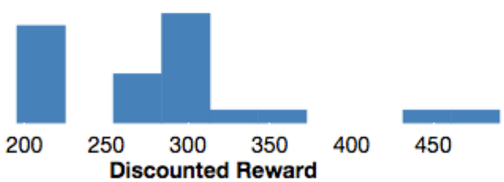


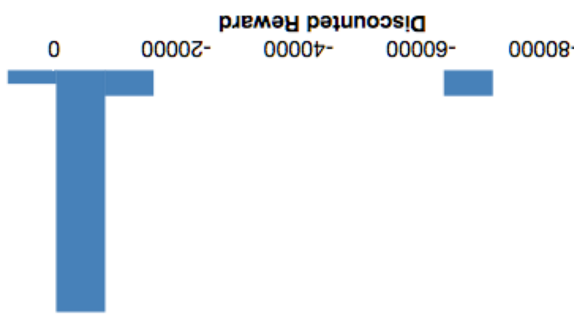


Comparison

$\pi_1 - \pi_2$

State Variable Distributions for a Fixed Time Step



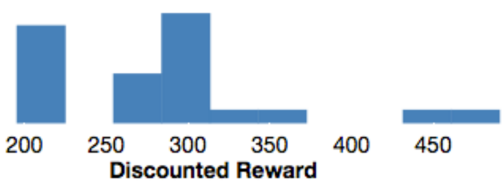


Comparison

$\pi_1 - \pi_2$

Oregon State University

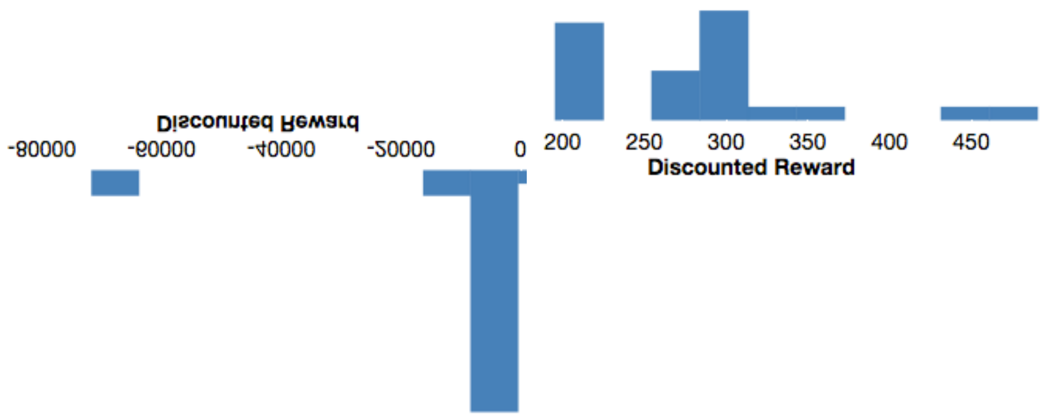
State Variable Distributions for a Fixed Time Step



Comparison

$\pi_1 - \pi_2$

State Variable Distributions for a Fixed Time Step

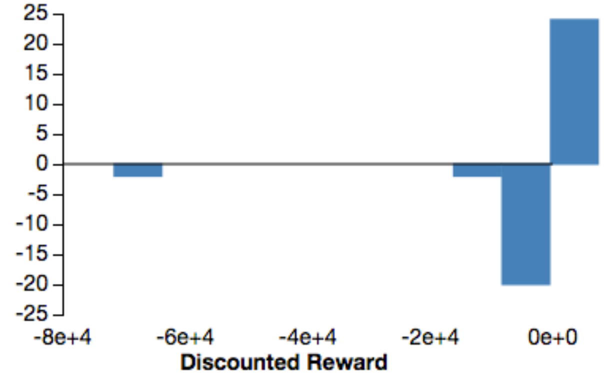


Comparison

$\pi_1 - \pi_2$

State Variable Distributions for a Fixed Time Step

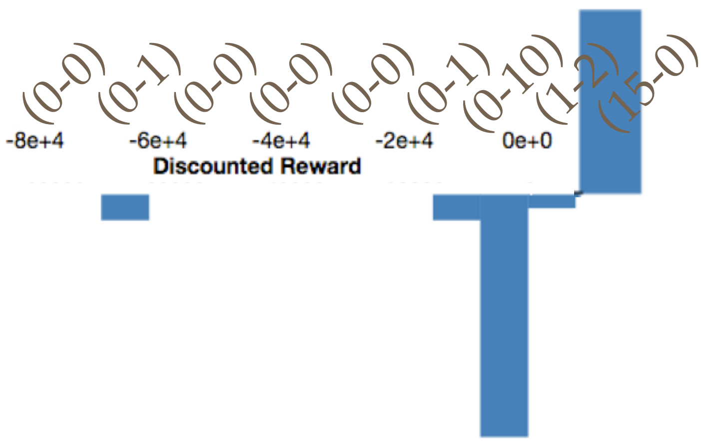
Rescale



Comparison

$\pi_1 - \pi_2$

State Variable Distributions for a Fixed Time Step

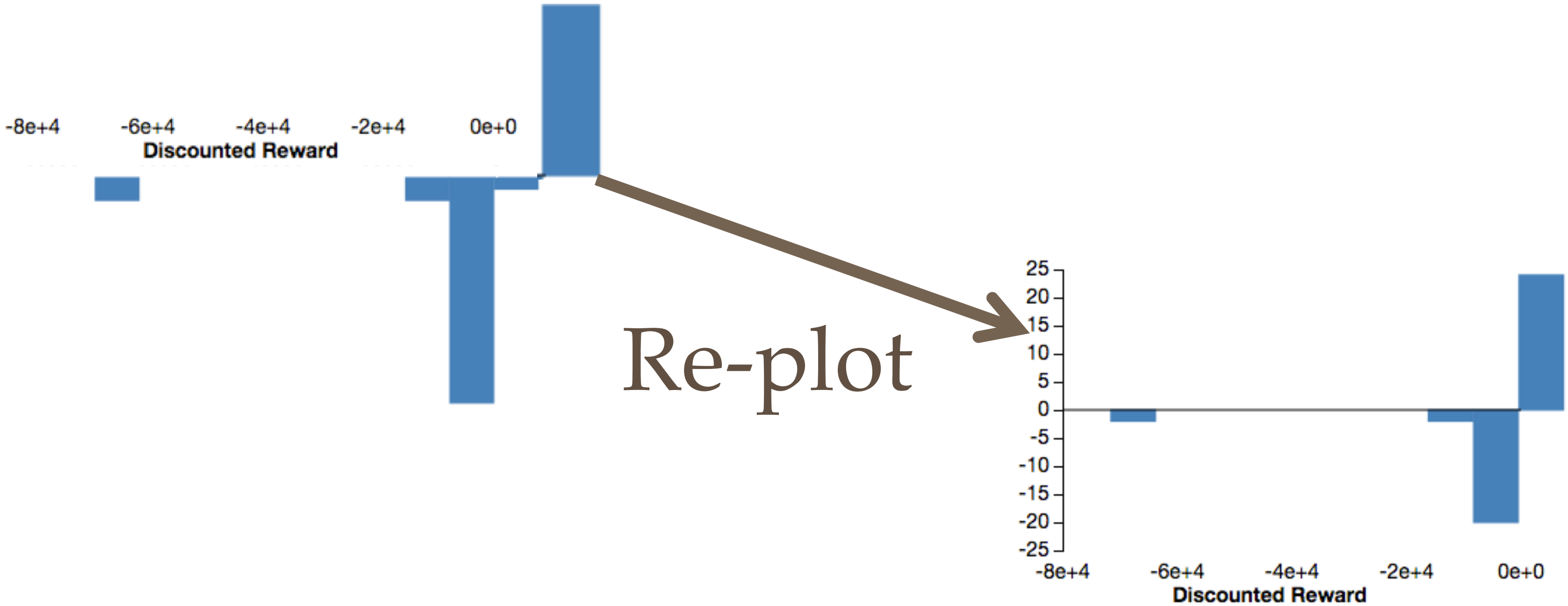


Take Difference in Counts

Comparison

$\pi_1 - \pi_2$

State Variable Distributions for a Fixed Time Step



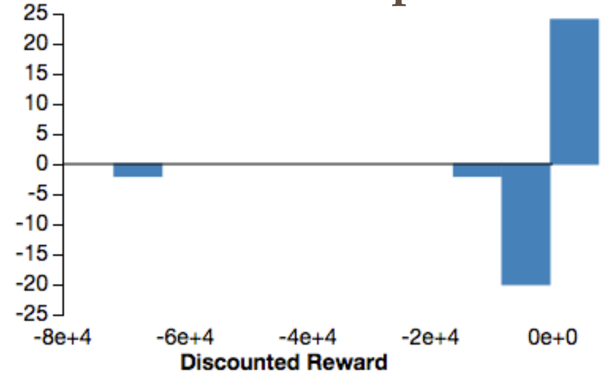
Comparison

$\pi_1 - \pi_2$

Oregon State University

State Variable Distributions for a Fixed Time Step

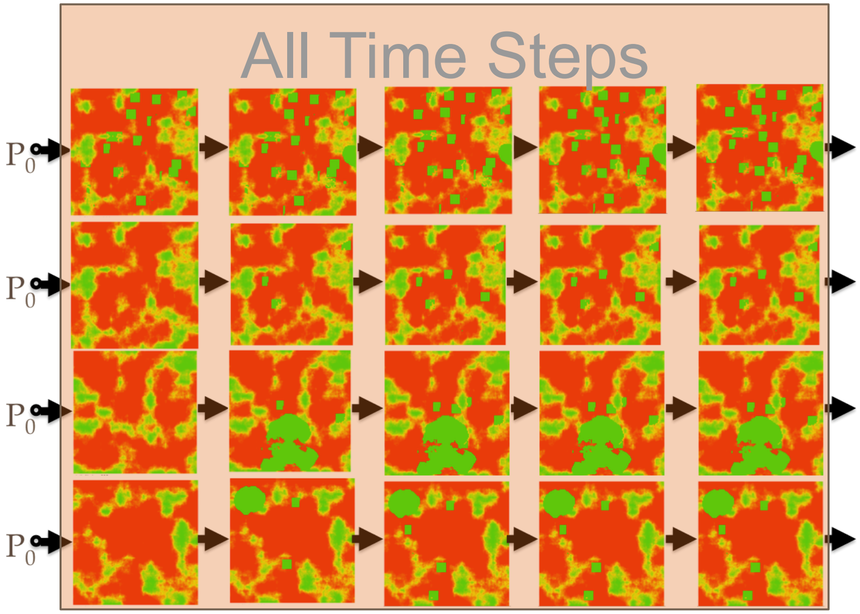
Let-Burn Dominates Suppress-All in this time step



Comparison

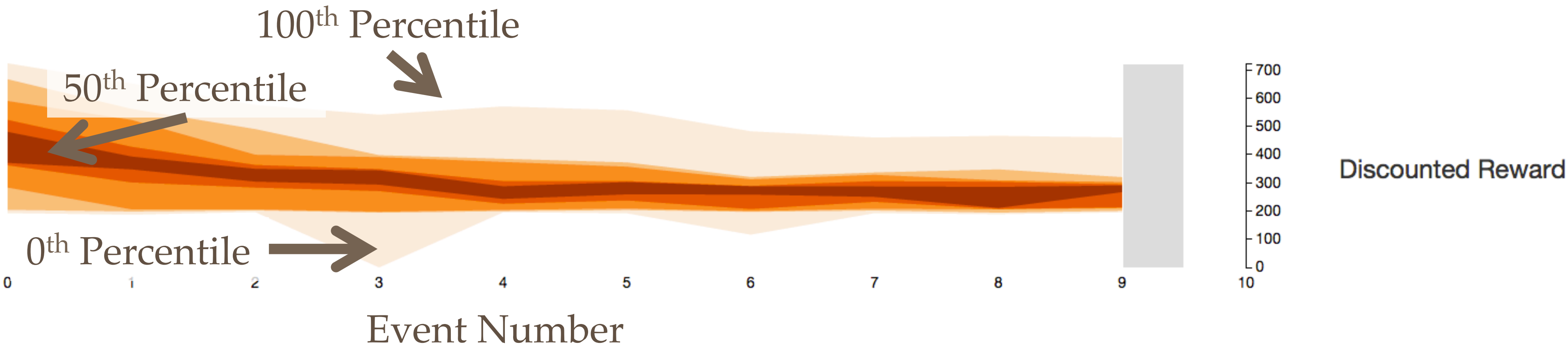
$\pi_1 - \pi_2$

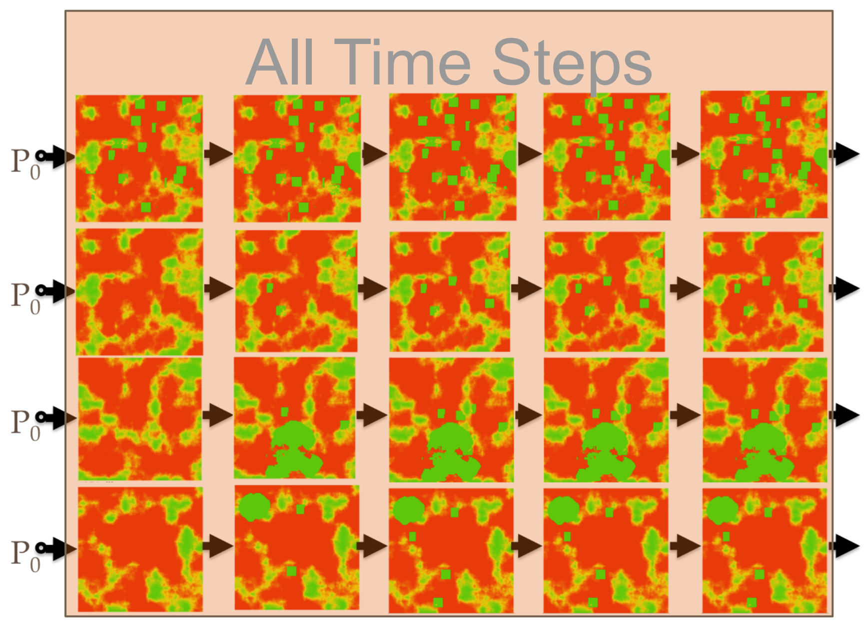
State Variable Distributions through Time



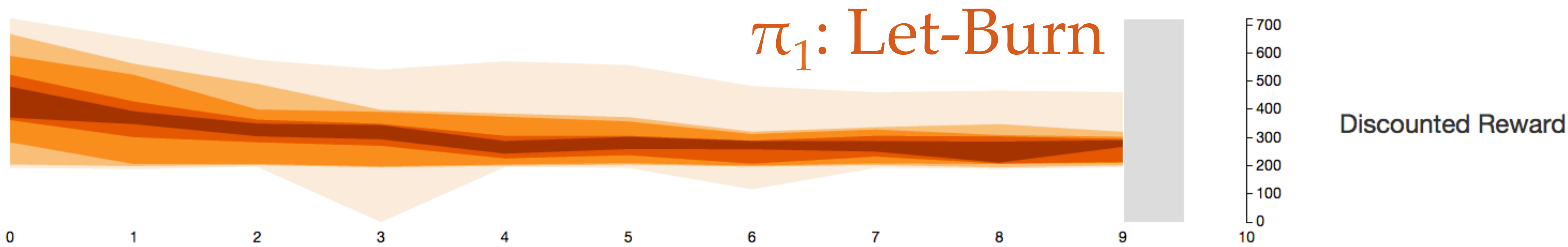
Oregon State UNIVERSITY

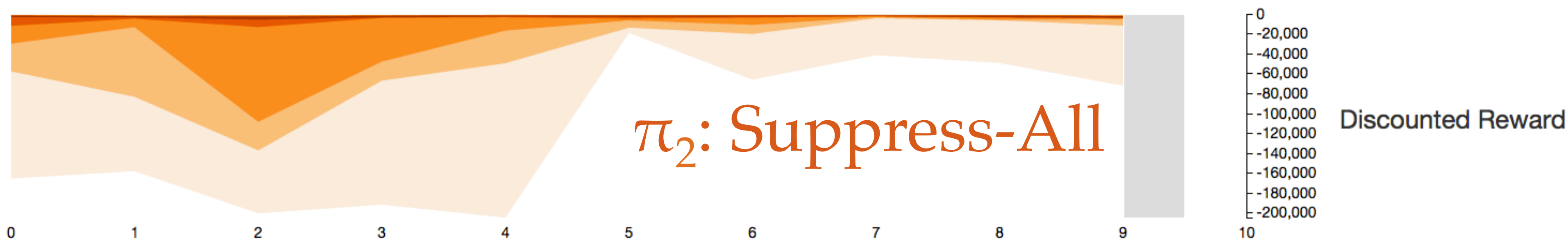
State Variable Distributions through Time

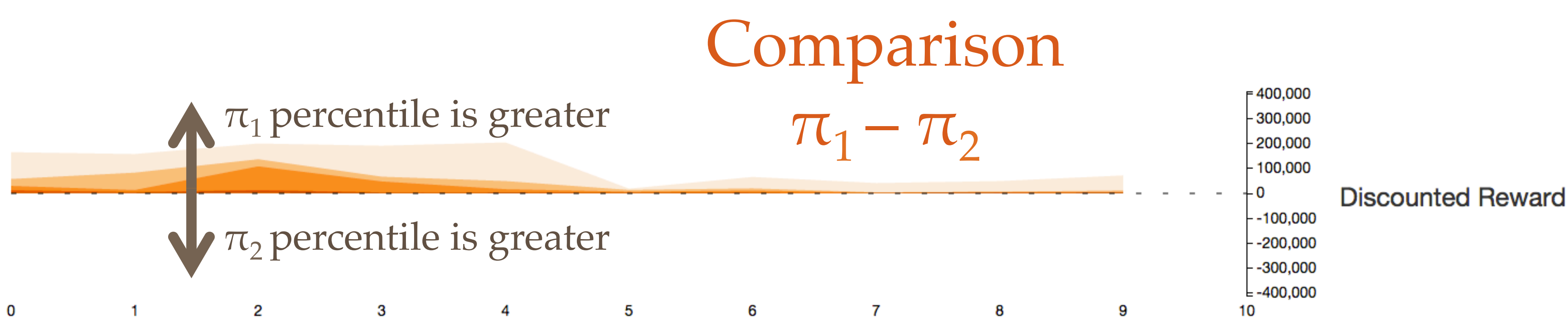




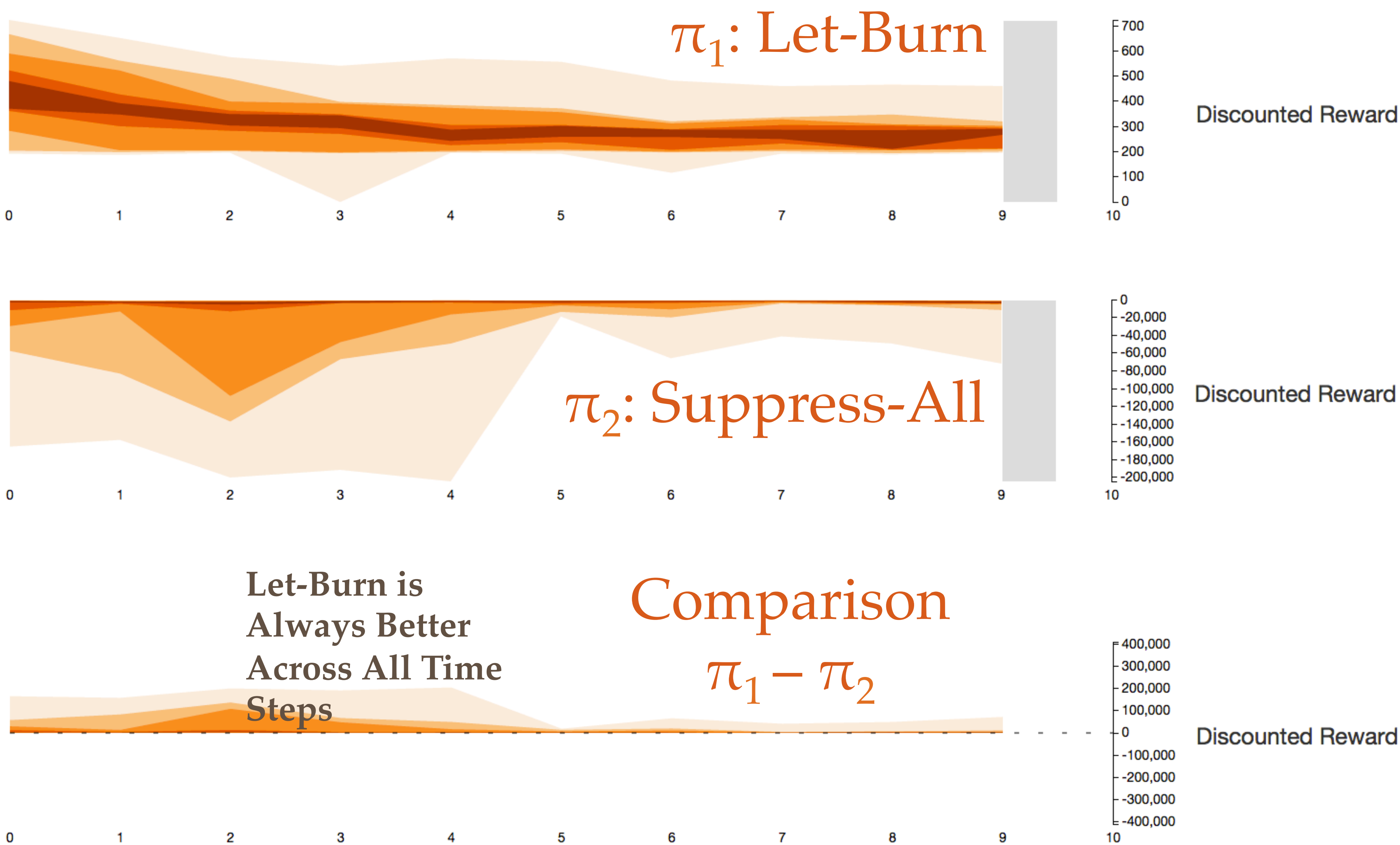
State Variable Distributions through Time





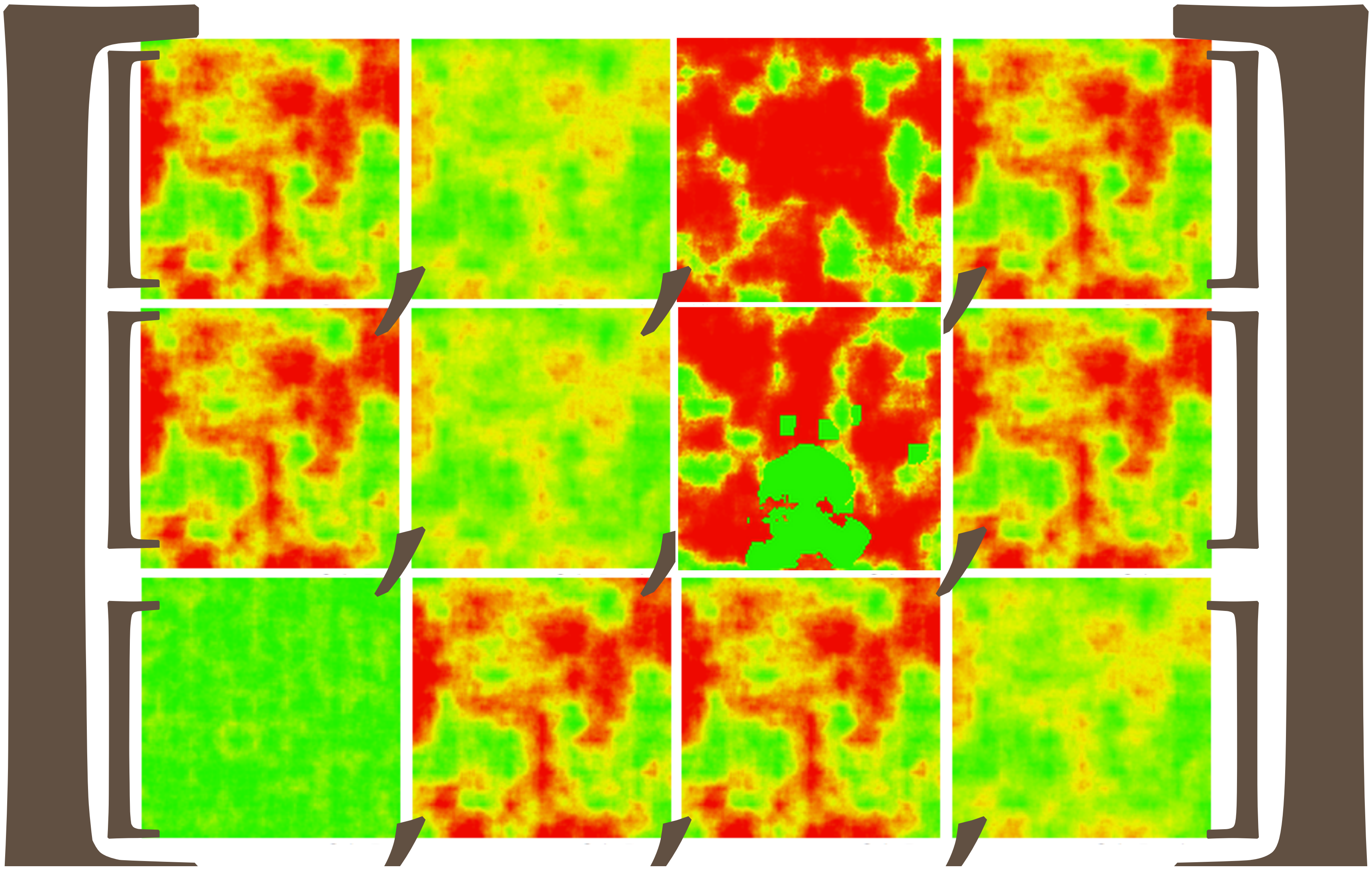


State Variable Distributions through Time



State details

Allow MDP Simulator to Generate State Visualizations



Parameter Space Analysis (PSA)

"[PSA] *is the systematic variation of model input parameters, generating outputs for each combination of parameters, and investigating the relation between parameter settings and corresponding outputs.*"

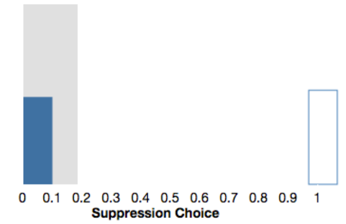
Categories
Sensitivity
Optimization
Outliers
Partition
Uncertainty
Fitting

Sedlmair, M., Heinzl, C., Bruckner, S., Piringer, H., & Möller, T. (2014). Visual parameter space analysis: A conceptual framework. Visualization and Computer Graphics, IEEE Transactions on, 20(12).

Is the policy Sensitive to the state?

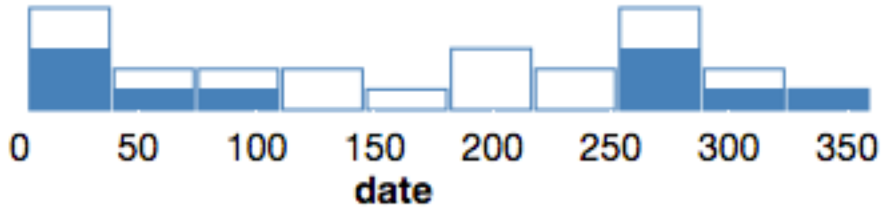
Interaction

1. Brush suppression choice to select Let-Burn



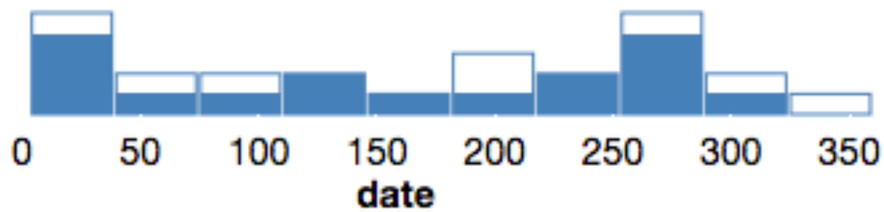
Expectation

2. Date is a determinant of suppression choice



Buggy Result

3. Date does not determine suppression choice



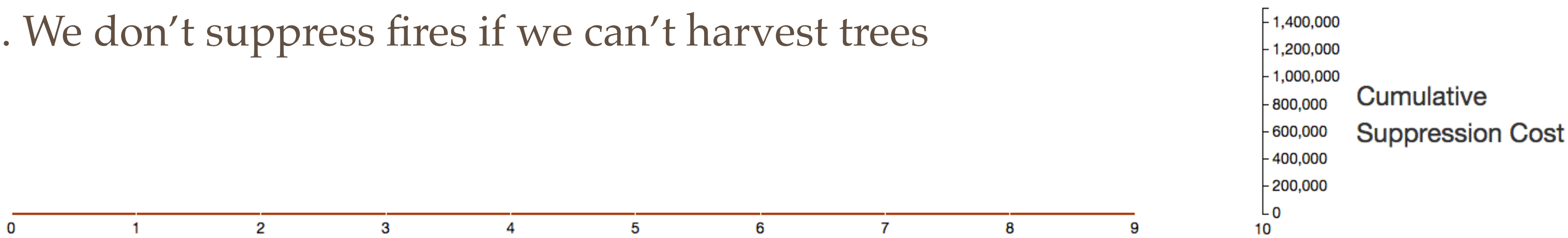
Is the optimization sensitive to the reward signal?

Interaction

1. Zero-out harvest rewards % 0 **Harvest Percent**
2. Re-optimize and generate rollouts **Optimize a New Policy**

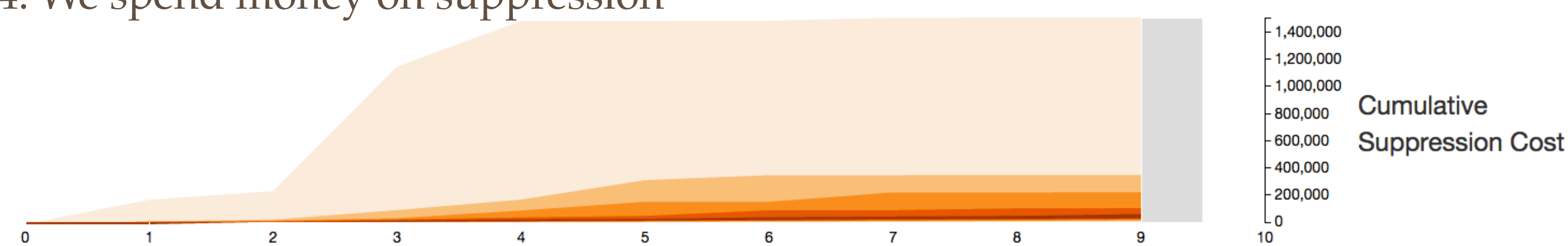
Expectation

3. We don't suppress fires if we can't harvest trees



Buggy Result

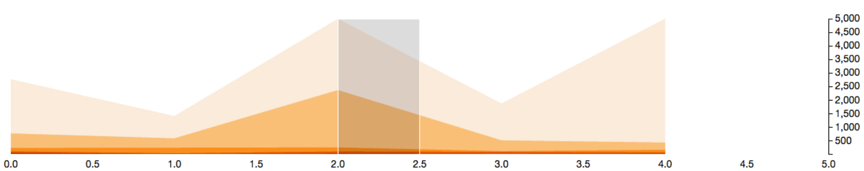
4. We spend money on suppression



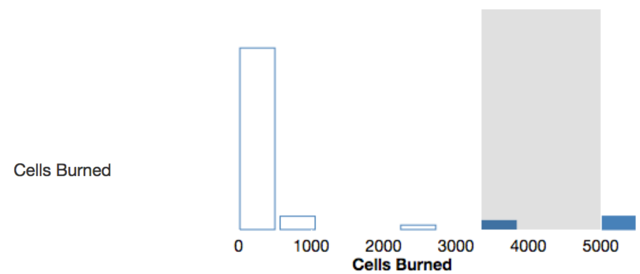
Are the largest fires realistic?

Interaction

1. Change the year to the one with the largest fire

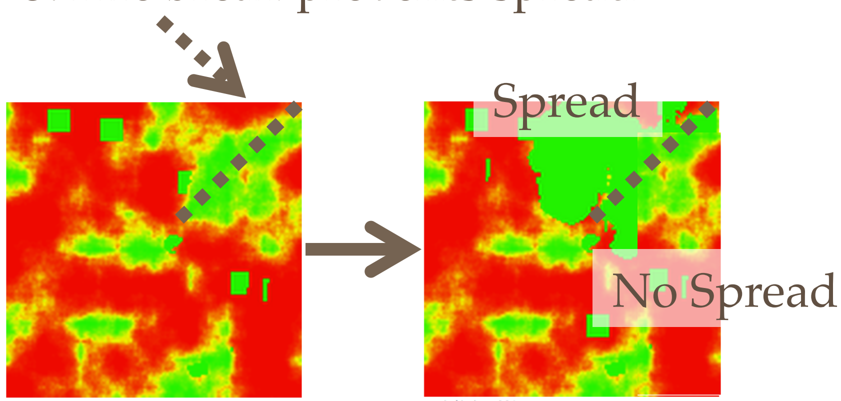


2. Brush histogram to view largest fires



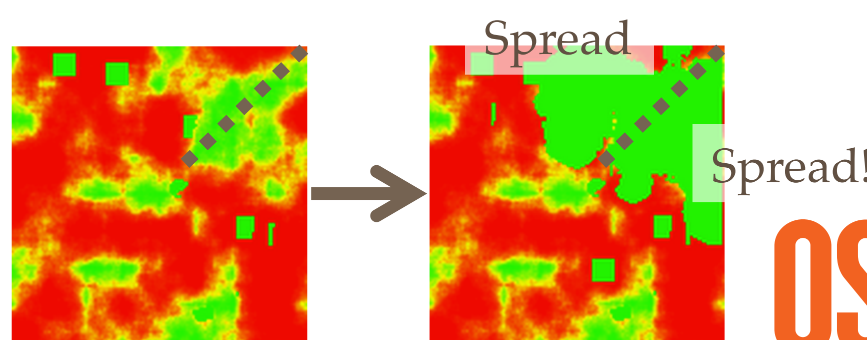
Expectation

3. Fire break prevents spread



Buggy Result

4. Fire break doesn't prevent spread



Oregon State UNIVERSITY

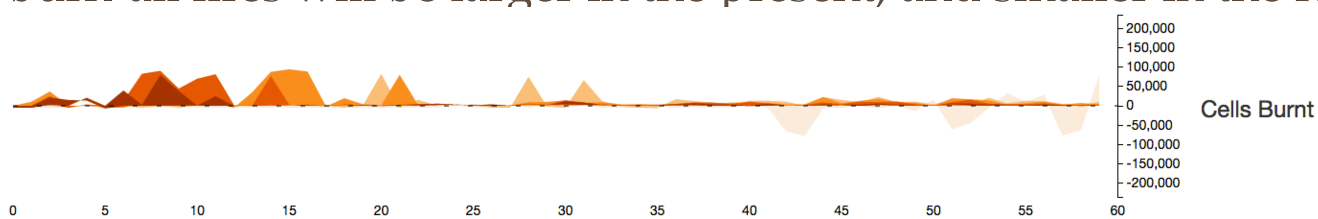
Do policies partition the state space?

Interaction

1. Generate suppress-all rollouts
2. Generate let-burn-all rollouts
3. Click the "compare rollouts" button

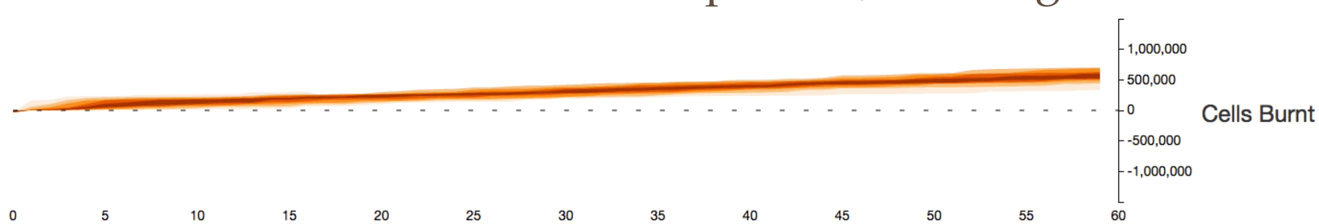
Expectation

4. Let-burn-all fires will be larger in the present, and smaller in the future



Buggy Result

5. Let-burn-all fires are the same in the present, and larger in the future



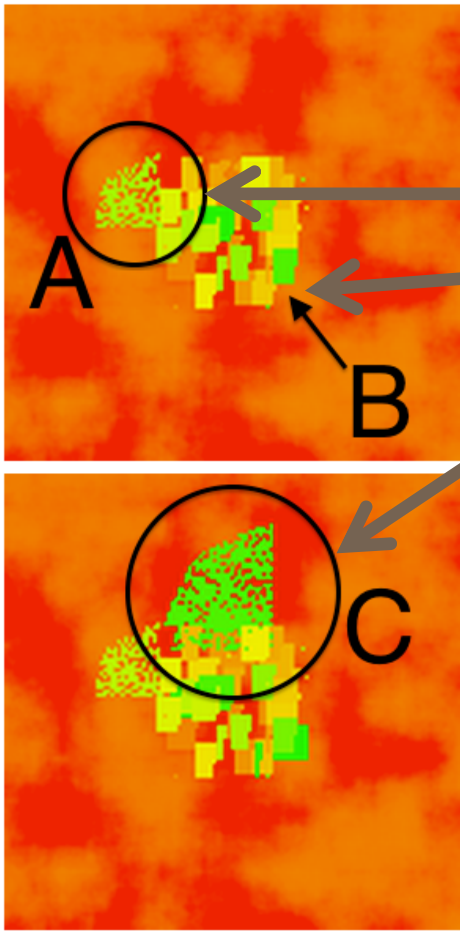
Use Case Study of MDPvis

We tested a new wildfire policy domain

- Visualization Developer: 1 Ph.D. Student in Computer Science
- New Fire Domain Developer: 1 Ph.D. Student in Forestry
- Wildfire Optimization Expert: 1 faculty research assistant

We found numerous bugs

Evaluation of MDPvis



Viewed Largest Fires in Rollouts

Second Largest Fire

Harvest Areas

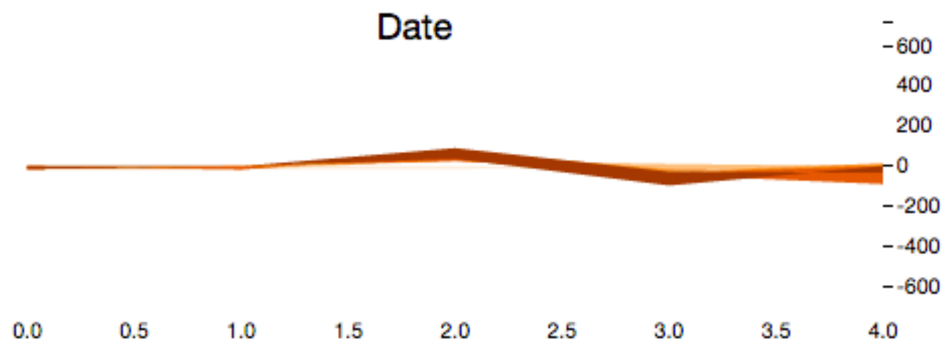
Largest Fire

Fires are not spreading east!

Hidden except in most extreme fire by harvests

Evaluation of MDPvis

1. **Compare:** Same model with different policies
2. **Expect:** Same ignition date in both rollout sets.
3. **Actual:** Policies change the weather.



Conclusion

Summary

We need visualization IDEs for MDPs!

Interactive Demo

MDPVis.github.io

* Not robust to many *simultaneous* requests

Thanks

- **Reviewers:** <you know who you are>
- **Advisor:** Thomas Dietterich
- **Research Group:** Ronald Metoyer, Claire Montgomery, Rachel Houtman, Mark Crowley, Hailey Buckingham
- **Funder:** National Science Foundation

MDPVis.github.io

This material is based upon work supported by the National Science Foundation under Grant No. 1331932.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Oregon State
UNIVERSITY

End.

Questions?

MDPVis.github.io

Contact
Email: VLHCC@SeanBMcGregor.com
Twitter: @SeanMcGregor

Oregon State
UNIVERSITY

End.

Come to the Full Demo!

MDPVis.github.io

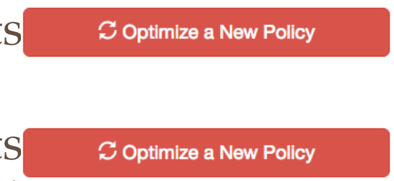
Contact
Email: VLHCC@SeanBMcGregor.com
Twitter: @SeanMcGregor

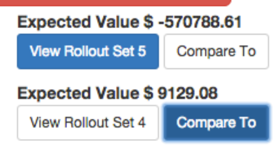
Oregon State UNIVERSITY

How consistent is the policy for small changes to the model?

Interaction

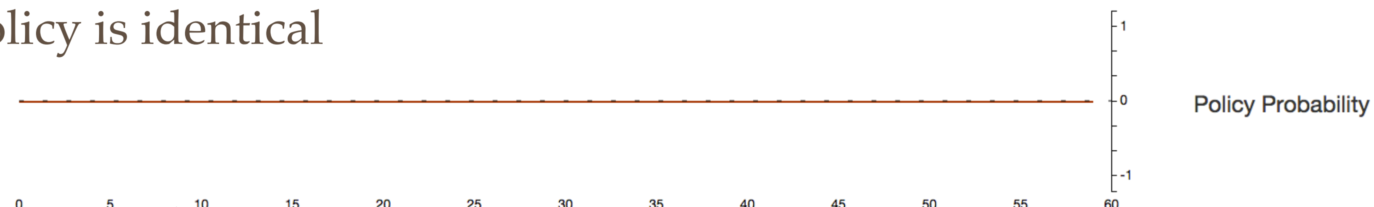
1. Optimize and generate rollouts
2. Add air tankers to the model
3. Optimize and generate rollouts
4. Click the "Compare Rollouts" button





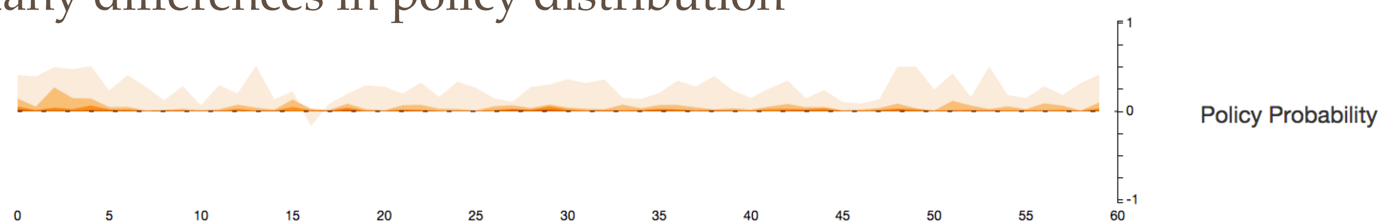
Expectation

5. Policy is identical



Buggy Result

6. Many differences in policy distribution



Oregon State University

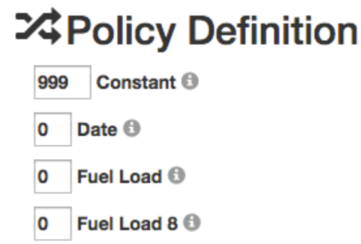
Does the growth rate match the historical dataset?

Pre-Process

1. Add a variable for the growth percentile within the historic data

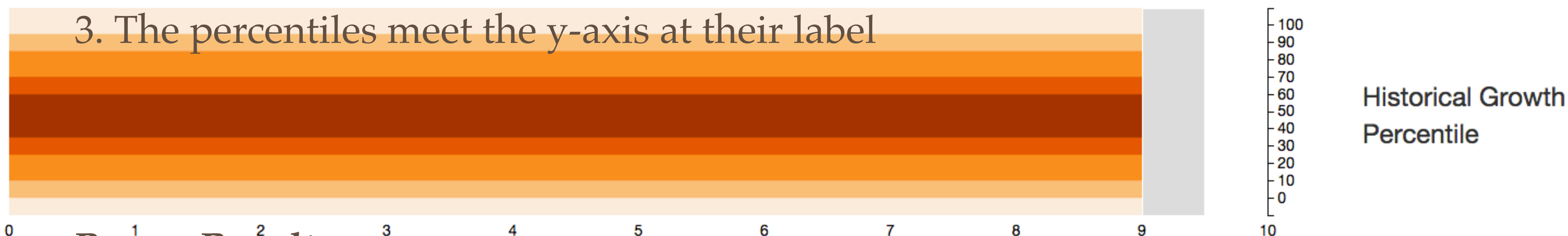
Interaction

2. Assign the policy to the historical policy (suppress all)



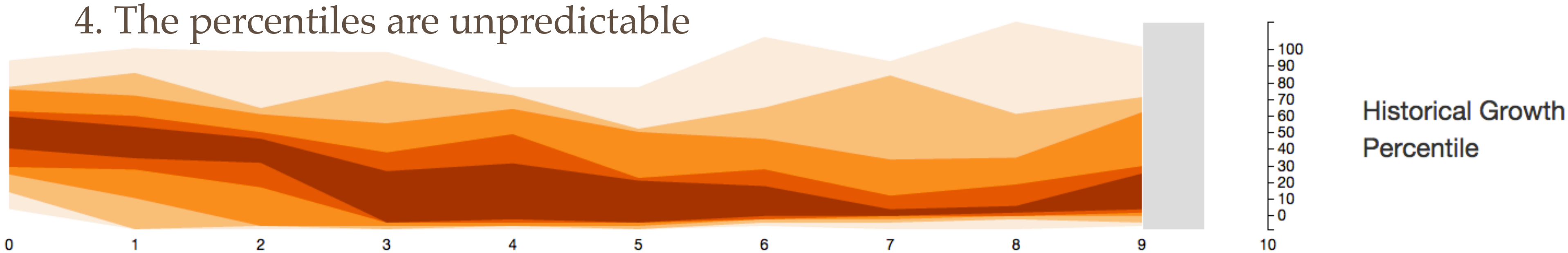
Expectation

3. The percentiles meet the y-axis at their label



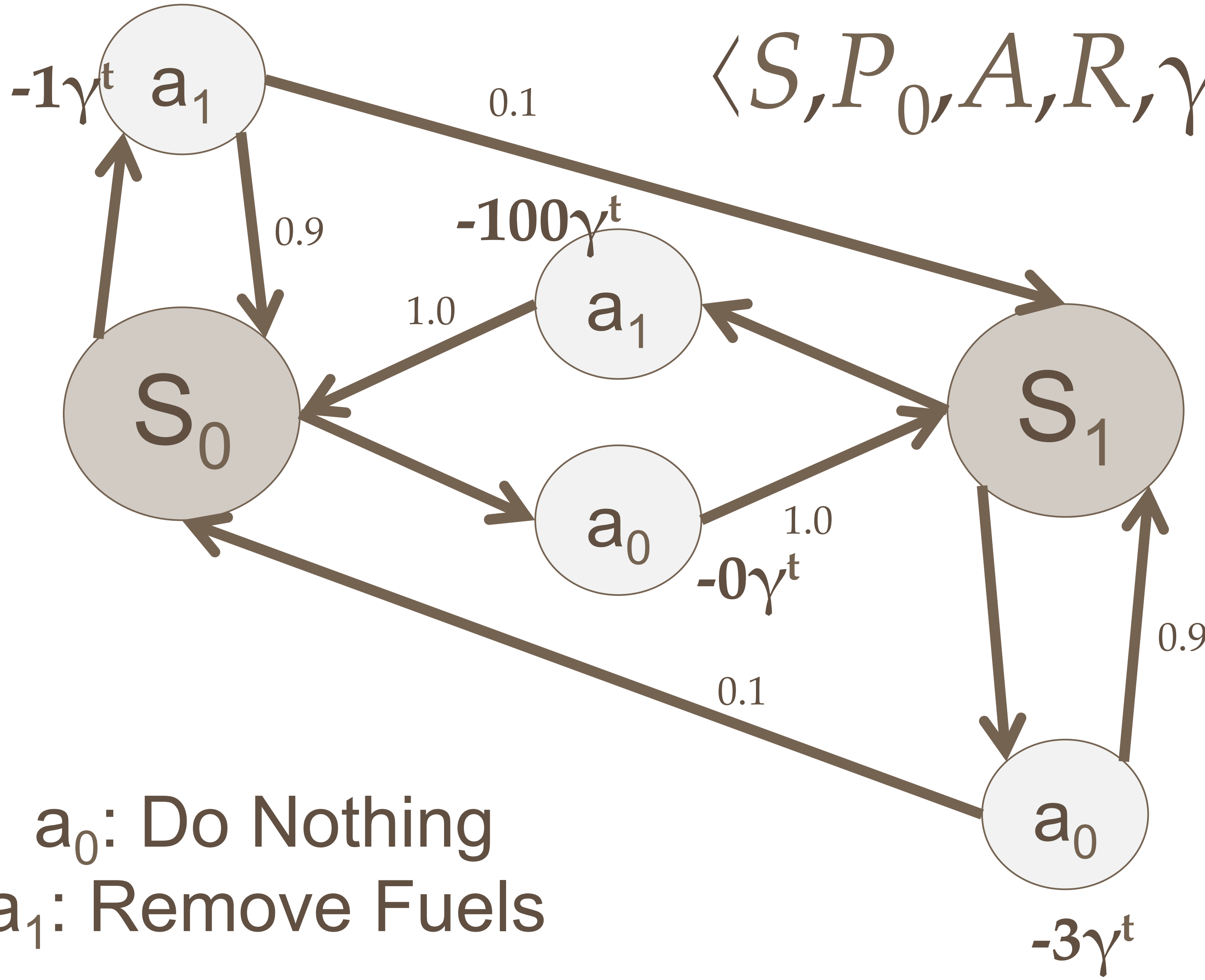
Buggy Result

4. The percentiles are unpredictable



Let's Construct a Simple MDP: "Pixel Forest"

$$\langle S, P_0, A, R, \gamma, P \rangle$$



a_0: Do Nothing
a_1: Remove Fuels