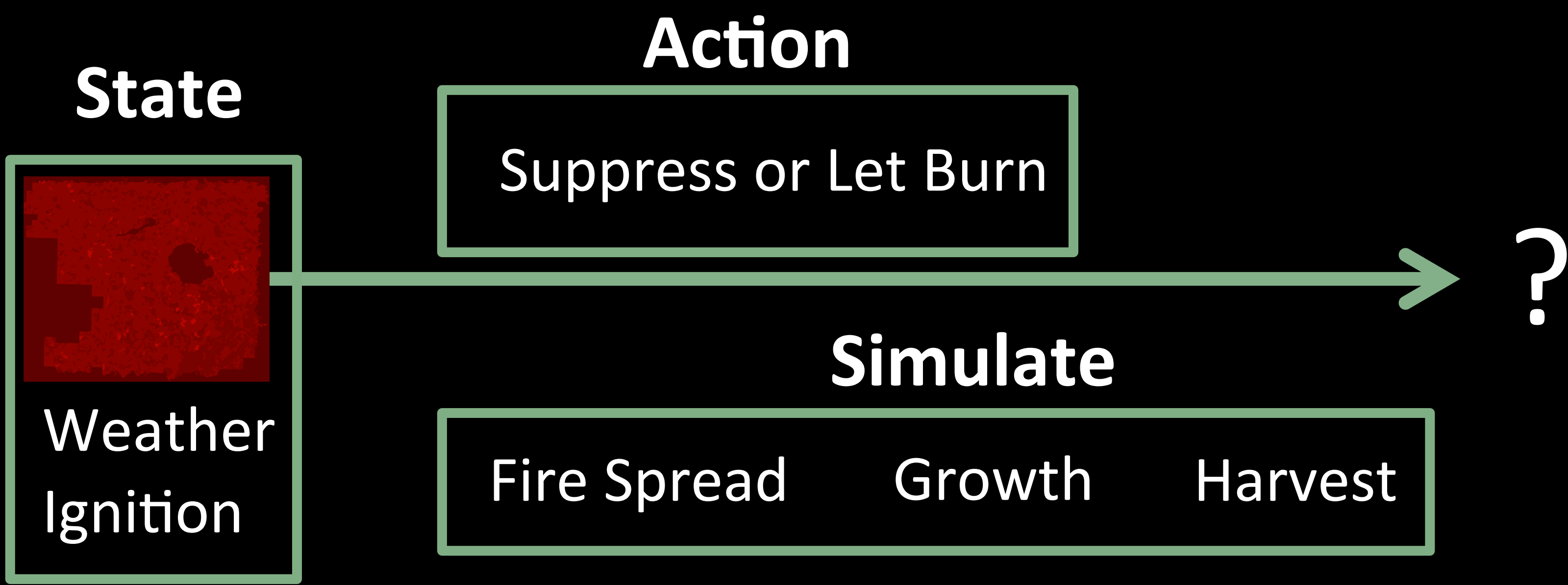
Fast Simulation for Computational Sustainability Sequential Decision Making Problems

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

Sequential Optimization in Wildfire Suppression Decisions



Simulate 100 Years of Decisions

Two Tasks

1. Optimizing policies
2. Validating policies

How do We Validate?

1. The optimization algorithm produced acceptable results
2. Simulation specification is correct

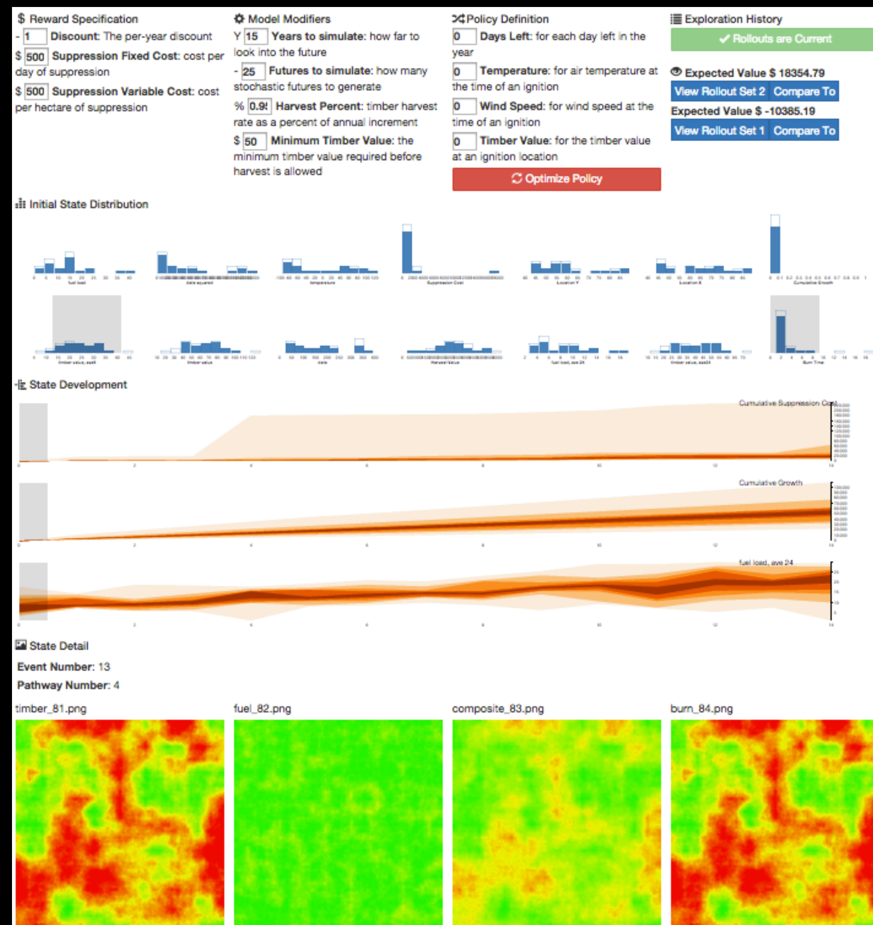
How do We Validate?

1. The optimization algorithm produced acceptable results
2. Simulation specification is correct

Visualization

Connect simulator to a visualization

1. Manually change parameters
2. Generate trajectories
3. Explore trajectories



How do We Validate?

1. The optimization algorithm produced acceptable results
2. Simulation specification is correct

Wildfire Reward Function

Timber Price * Timber Harvest

Reward = Timber Revenue + Ecology “Revenue” + Suppression Expenses

Ecological “Price” * Ecological State

Wildfire Reward Function

Suppress Everything!

Reward = Timber Revenue + Ecology “Revenue” + Suppression Expenses

Let Everything Burn!

Wildfire Reward Function

Selecting ecological reward selects policy

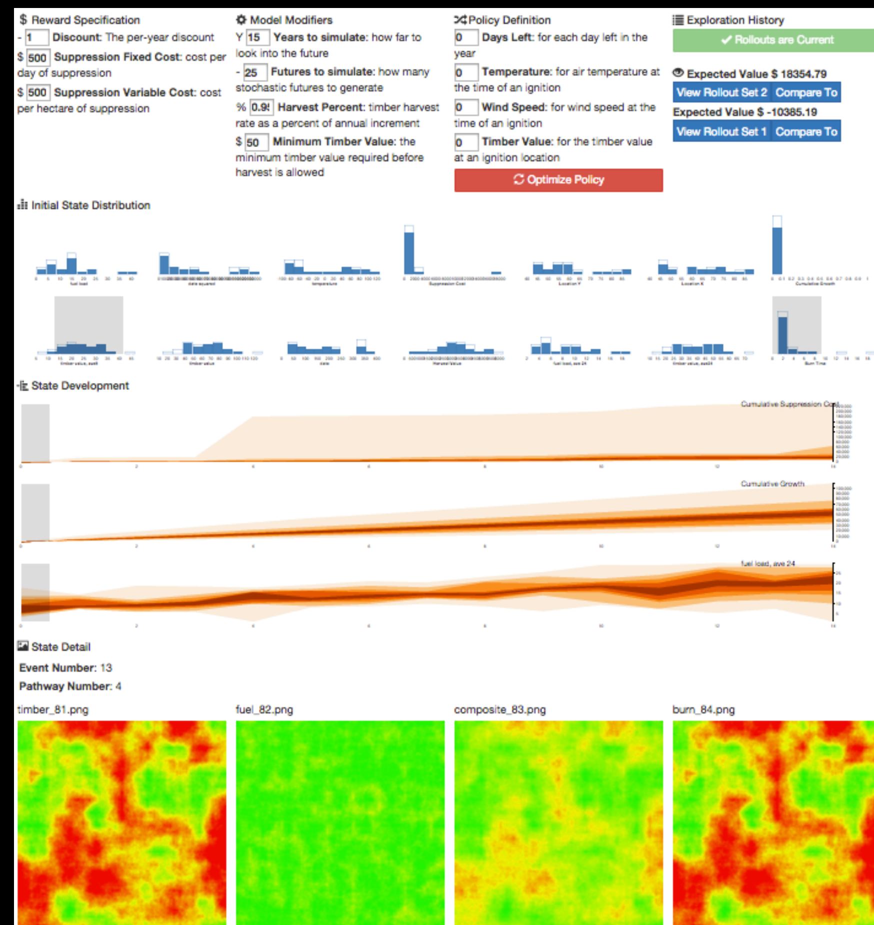
Visualization

Connect simulator to a visualization

1. Manually change parameters
2. Generate trajectories
3. Explore trajectories

Connect optimizer to visualization

1. Manually change parameters
2. Optimize policy
3. Generate trajectories
4. Explore trajectories



Problem

CompSust problems are (often) expensive to simulate

100 Years of Wildfire Simulation: several hours
Optimizing to 100 year horizon: *many days!*

Solution

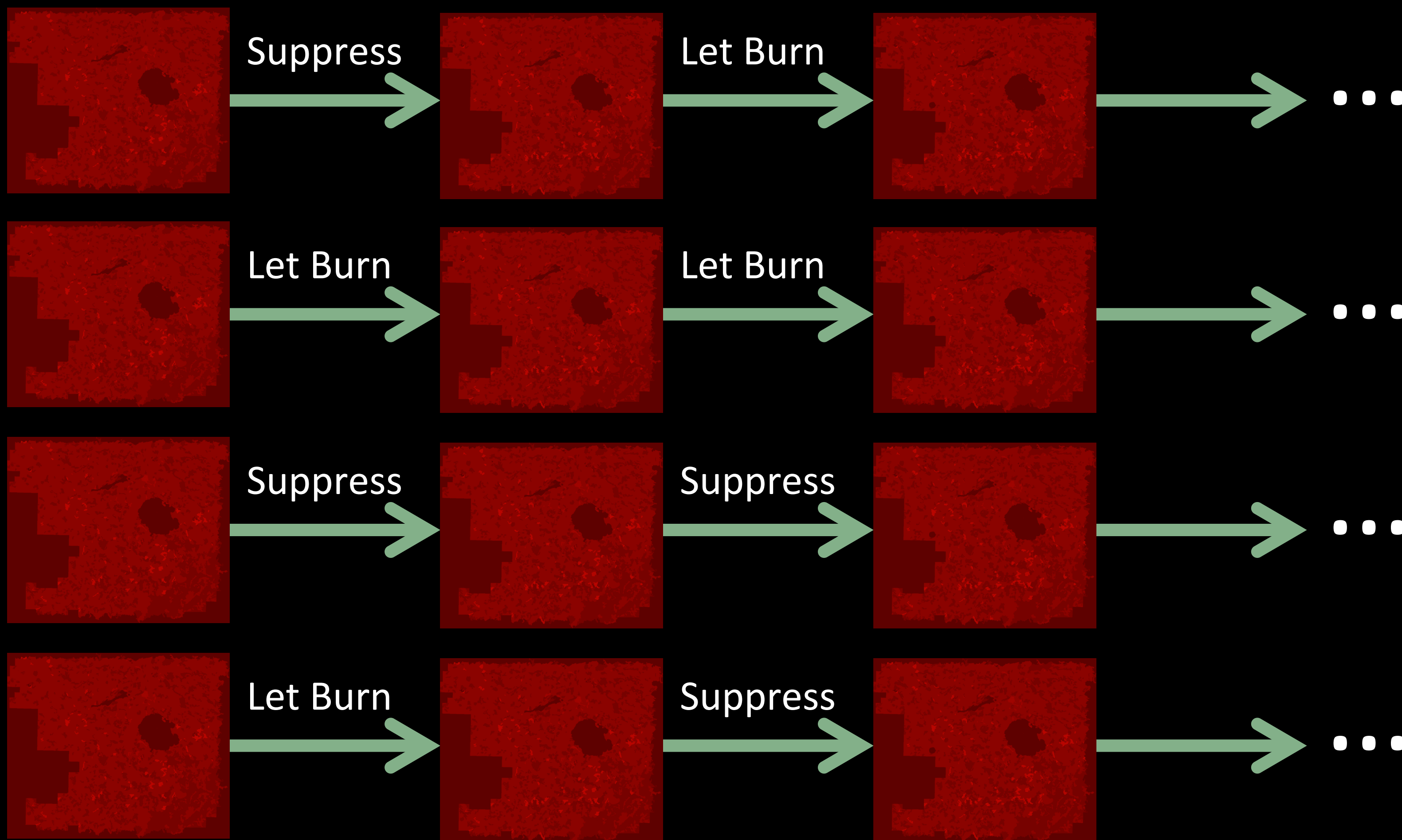
Trajectory Synthesis: Creating trajectories from a database of pre-computed state transitions

Benefits:

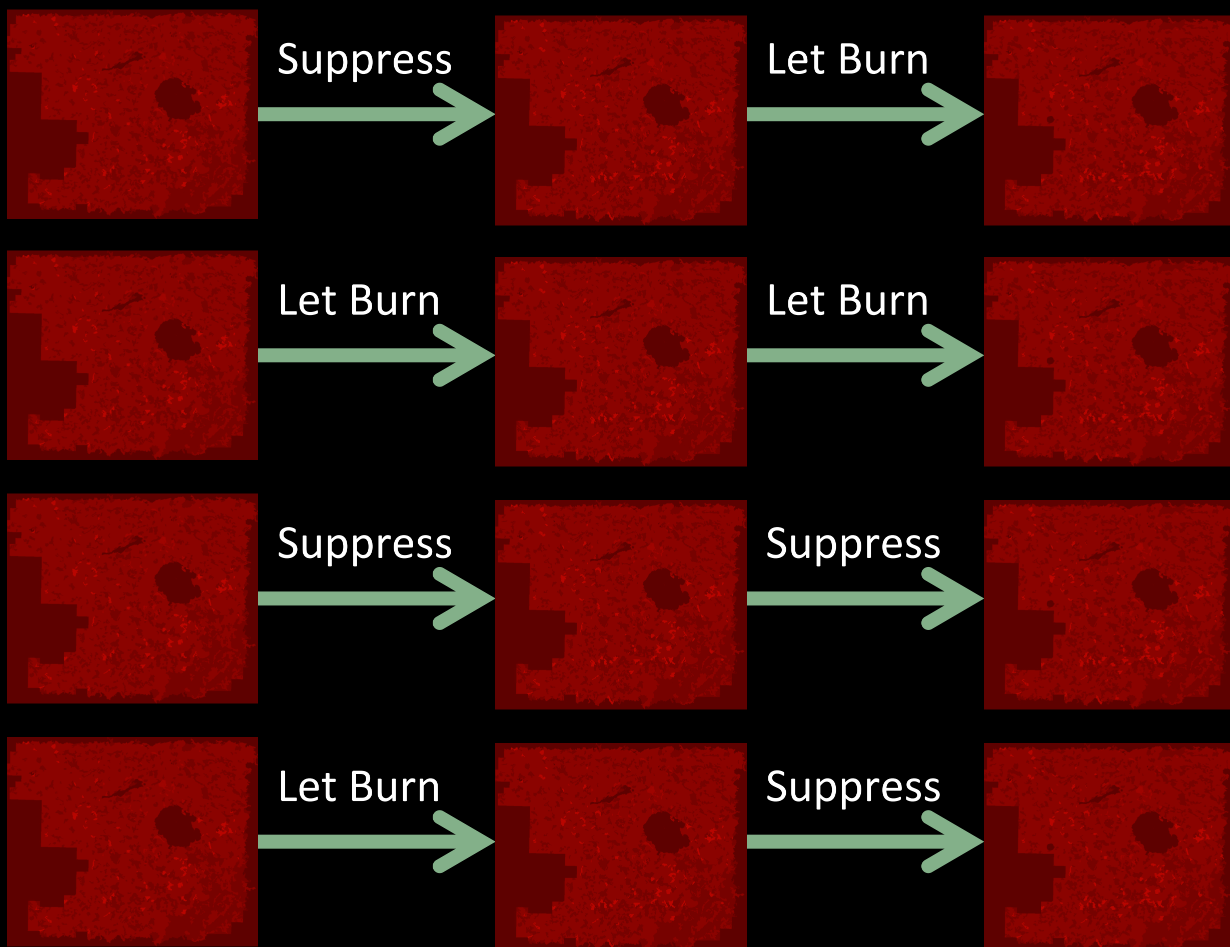
- Perform database queries at visualization time instead of expensive simulations
- Very sample efficient for exogenous variables

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.

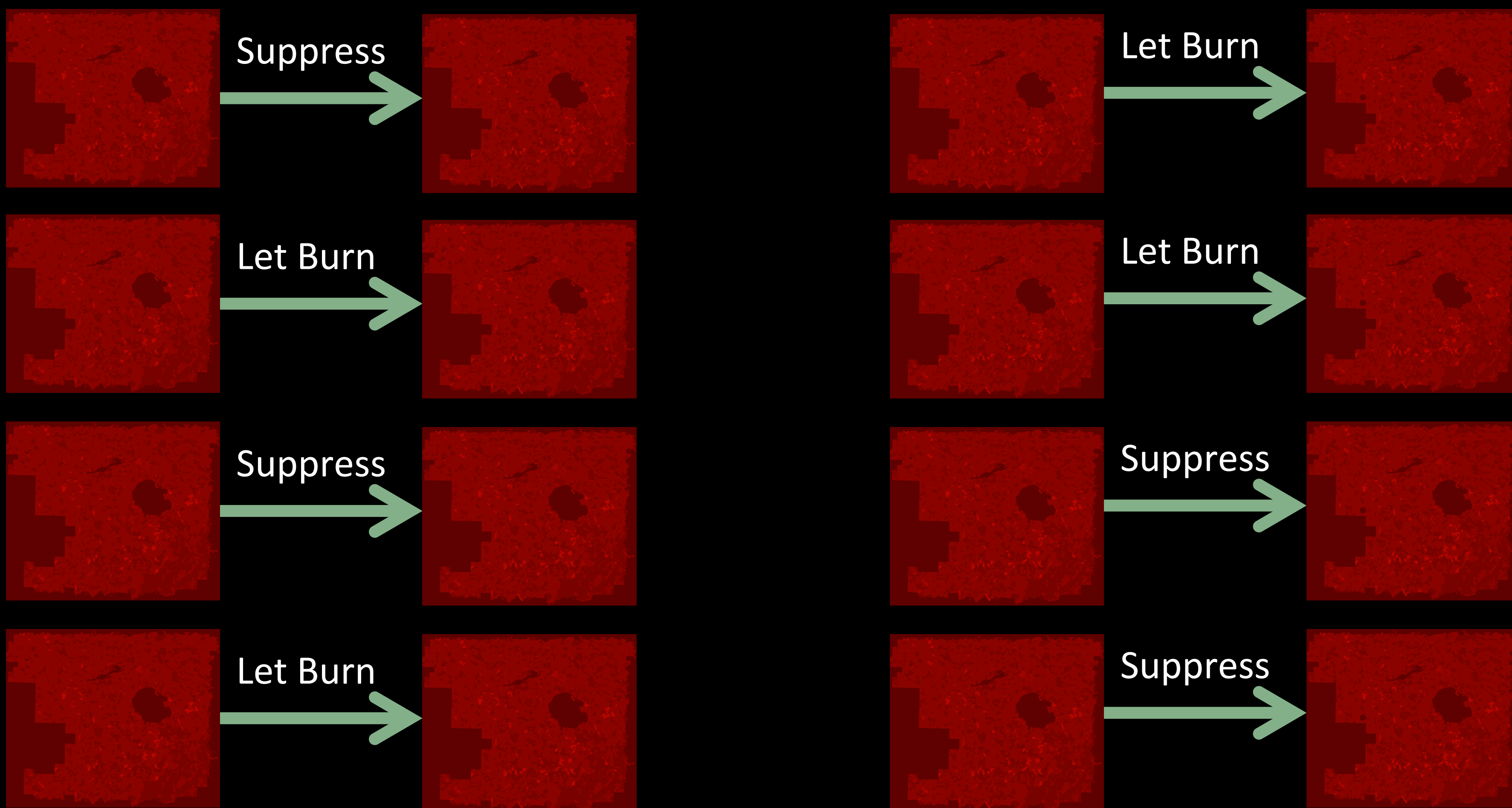
Constructing Database



Constructing Database

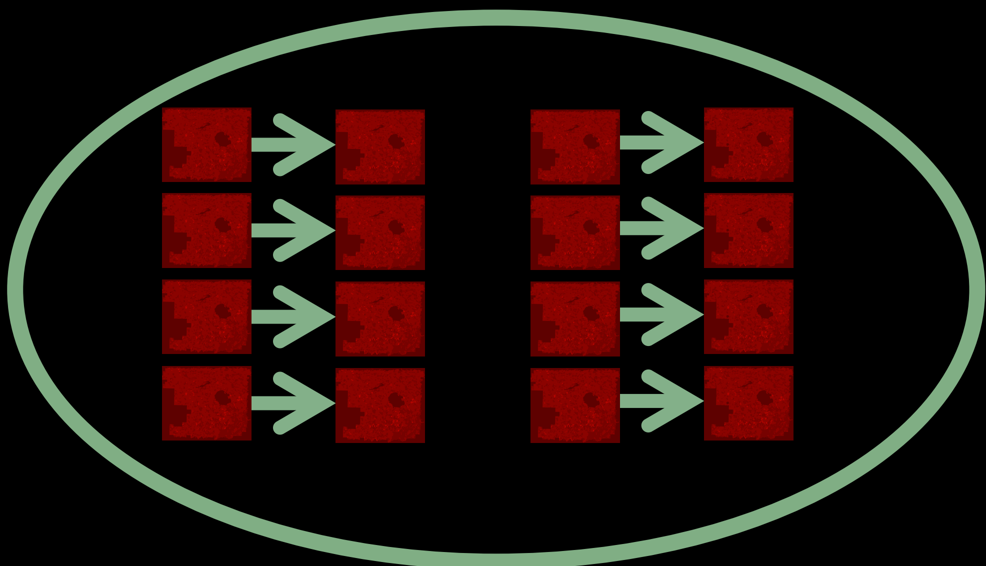


Constructing Database

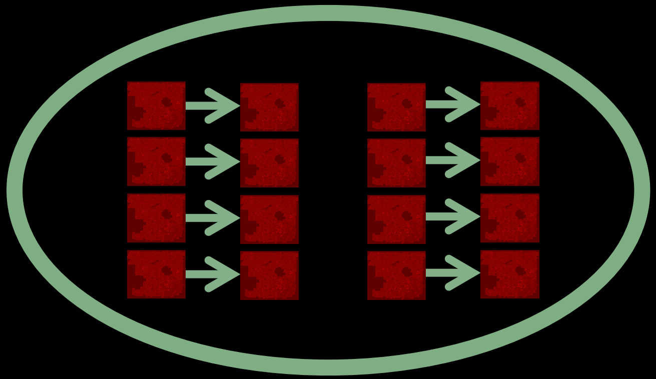


Constructing Database

Database

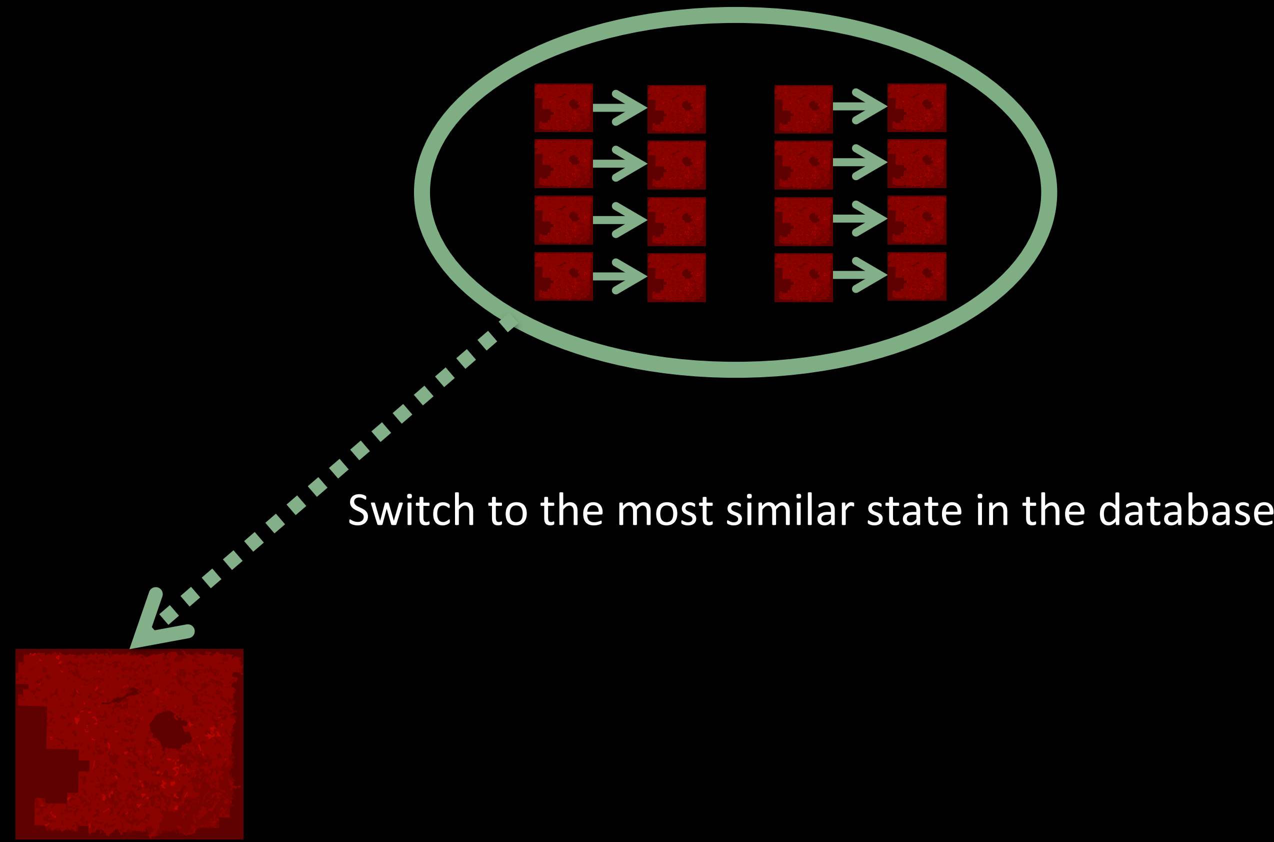


Fast Simulation

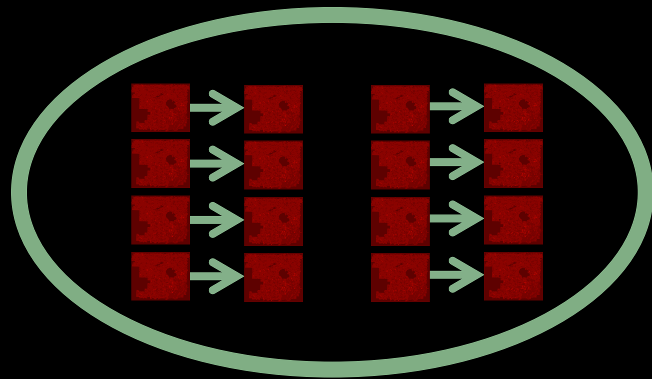


Draw Initial State

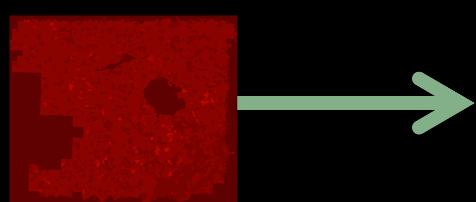
Fast Simulation



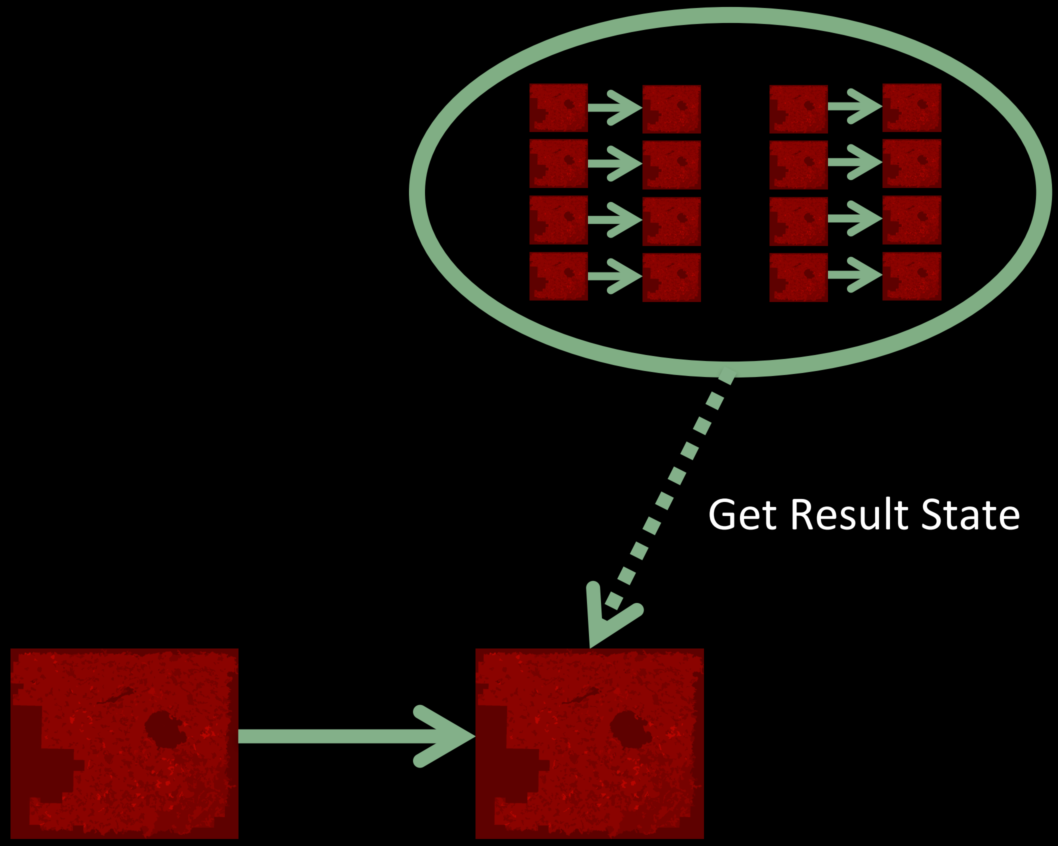
Fast Simulation



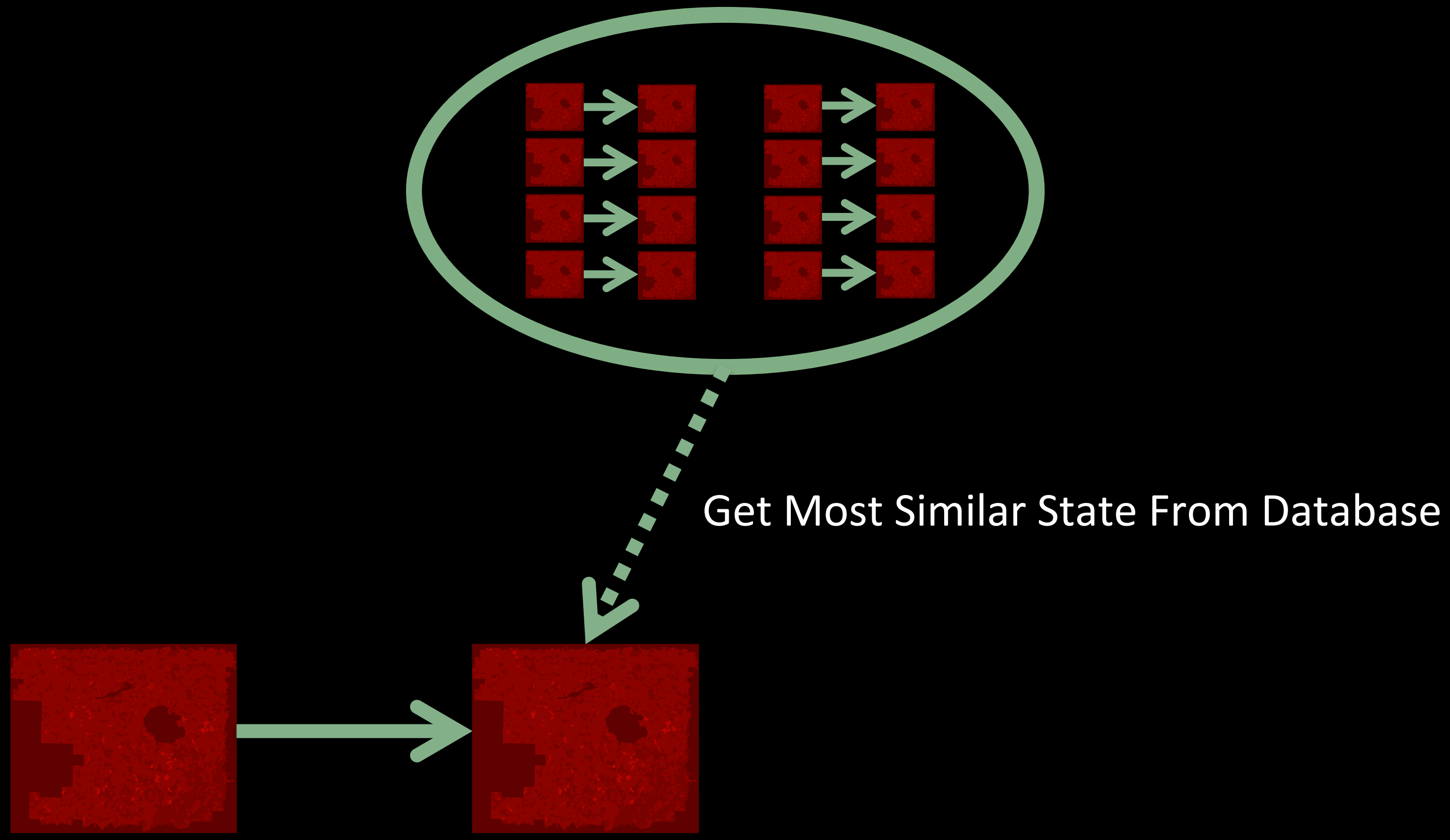
Apply Current Policy



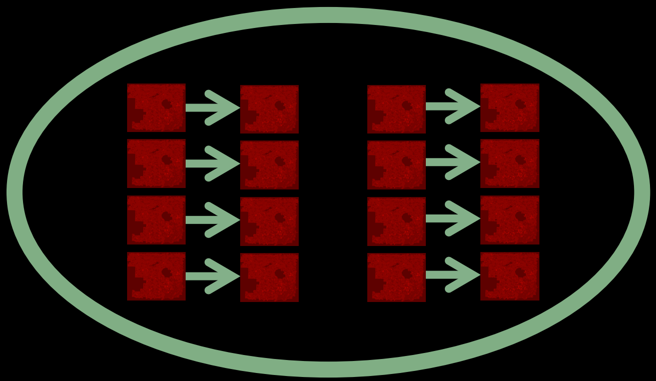
Fast Simulation

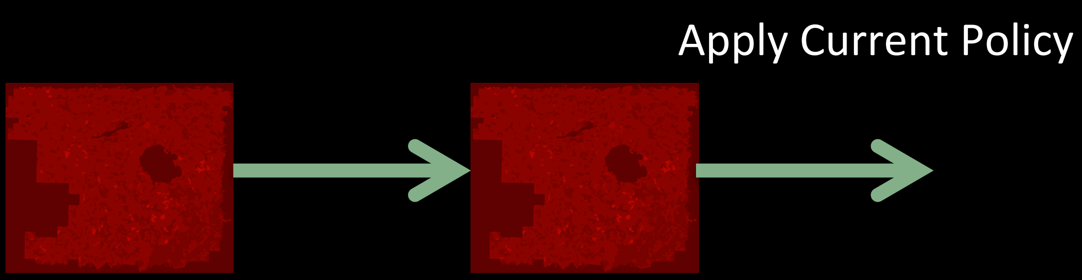


Fast Simulation

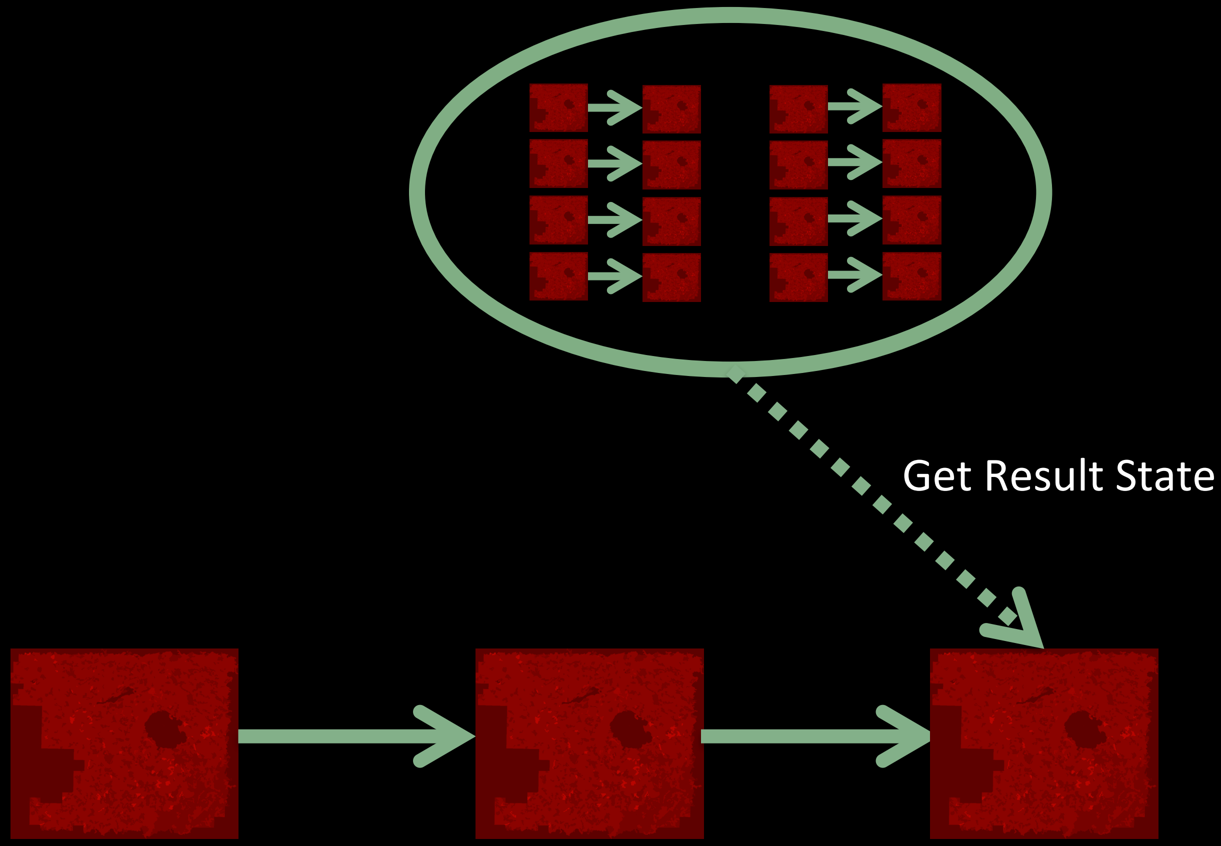


Fast Simulation

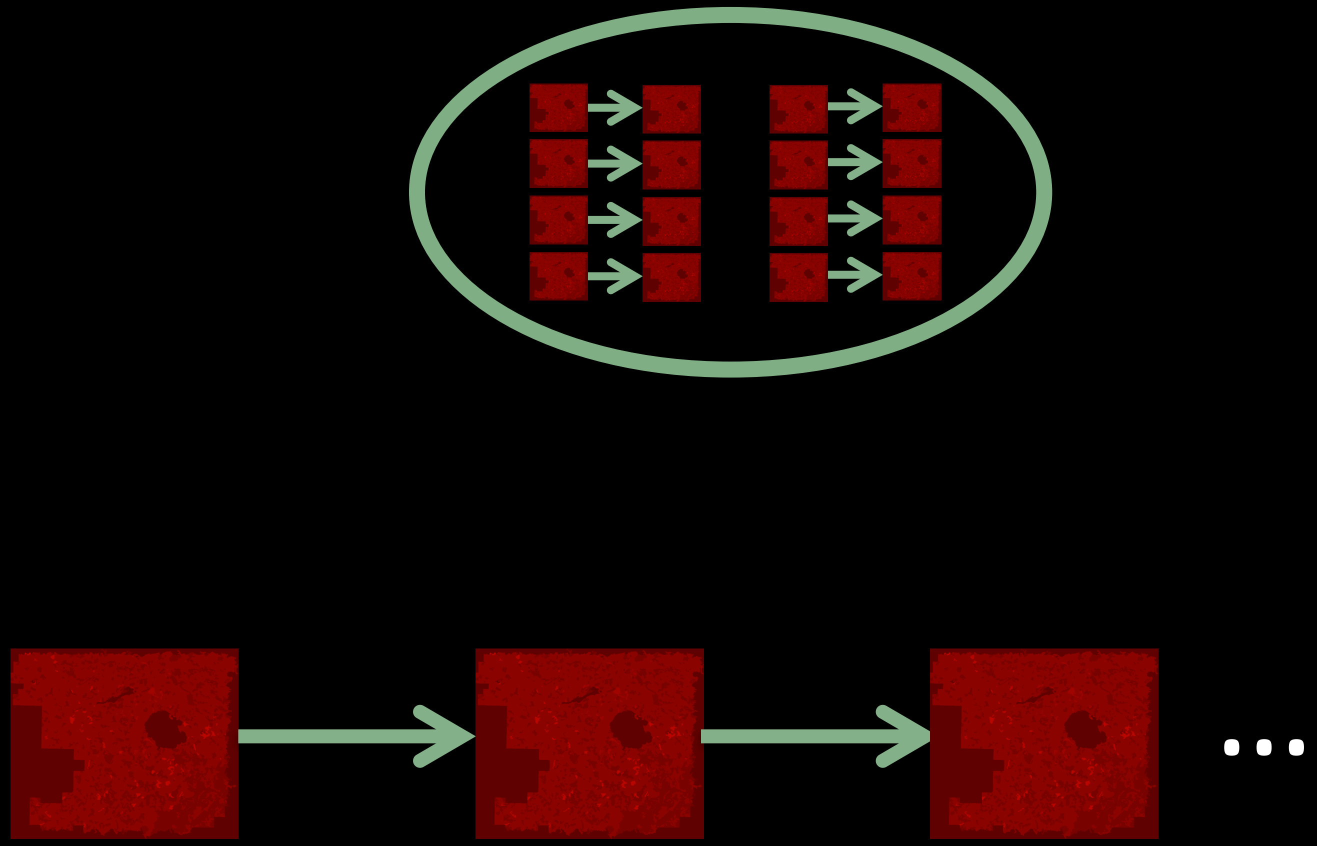




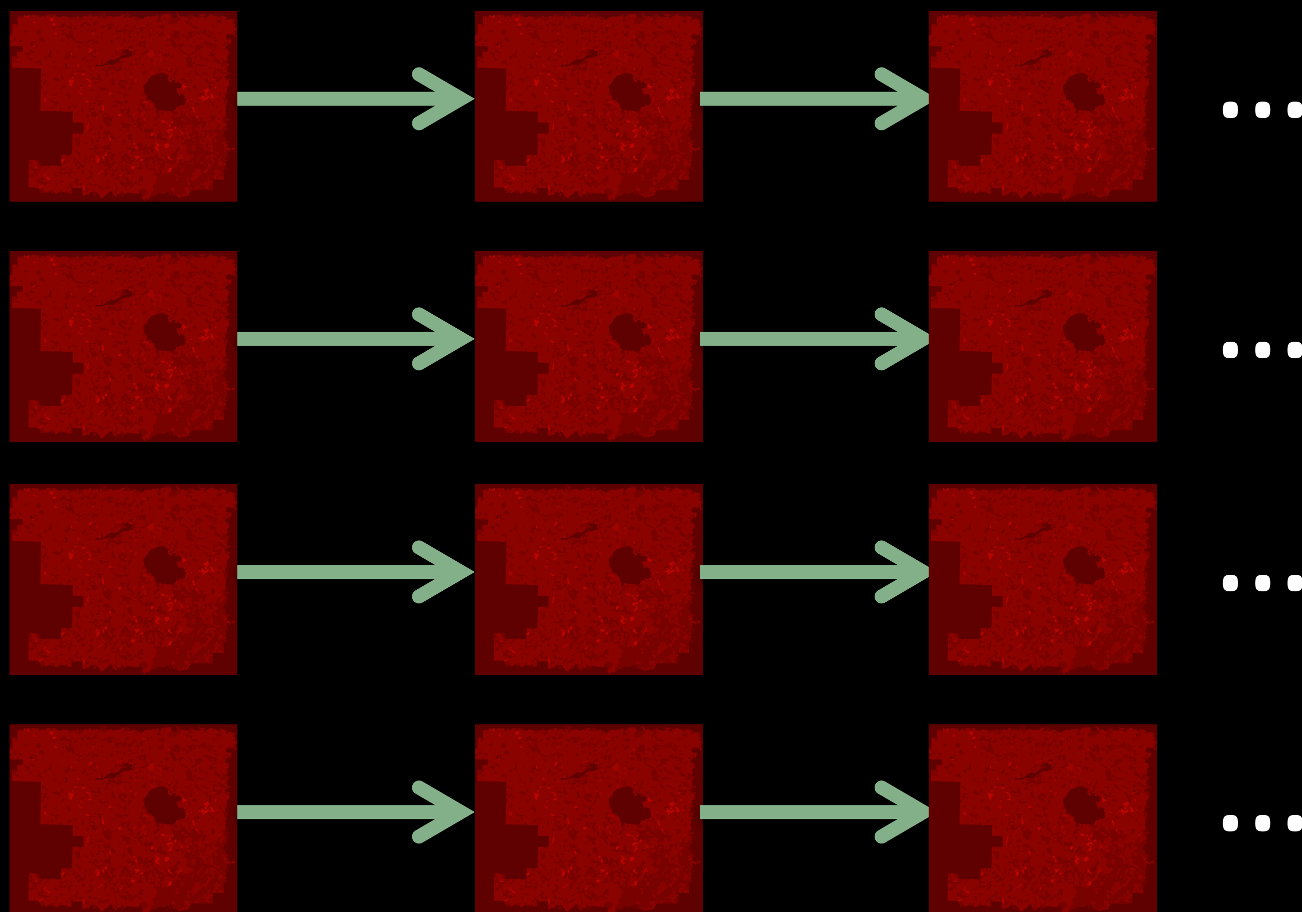
Fast Simulation



Fast Simulation

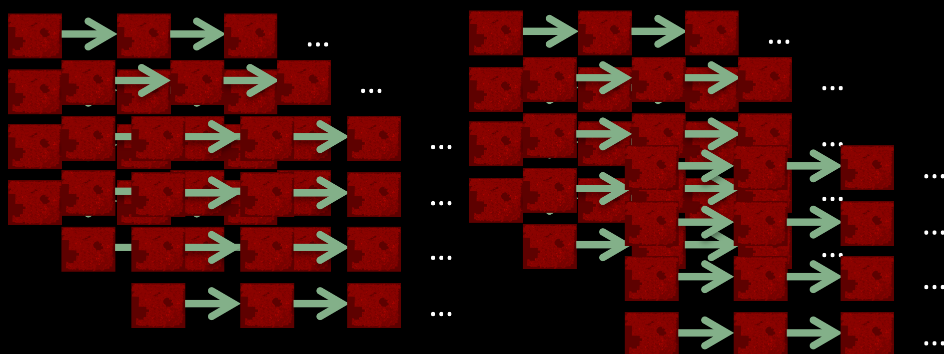


Fast Simulation

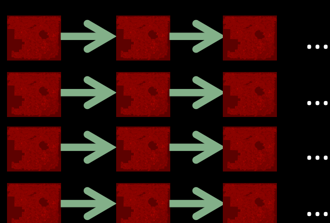


Fast Simulation

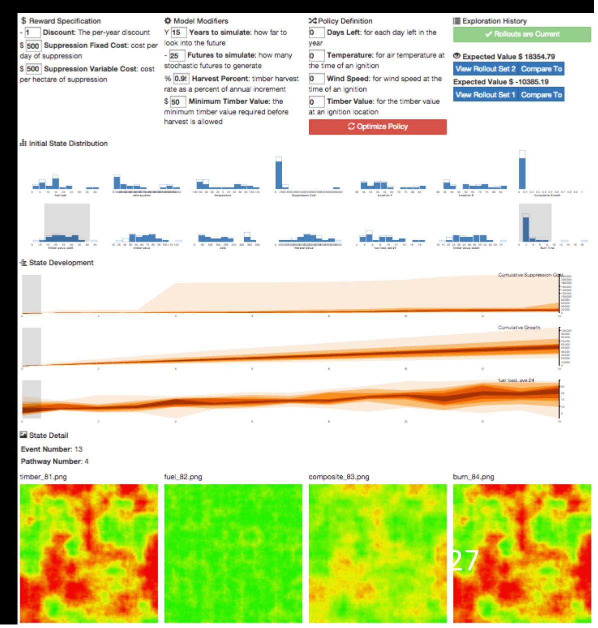
1. Optimize Over Many Sets of Trajectories



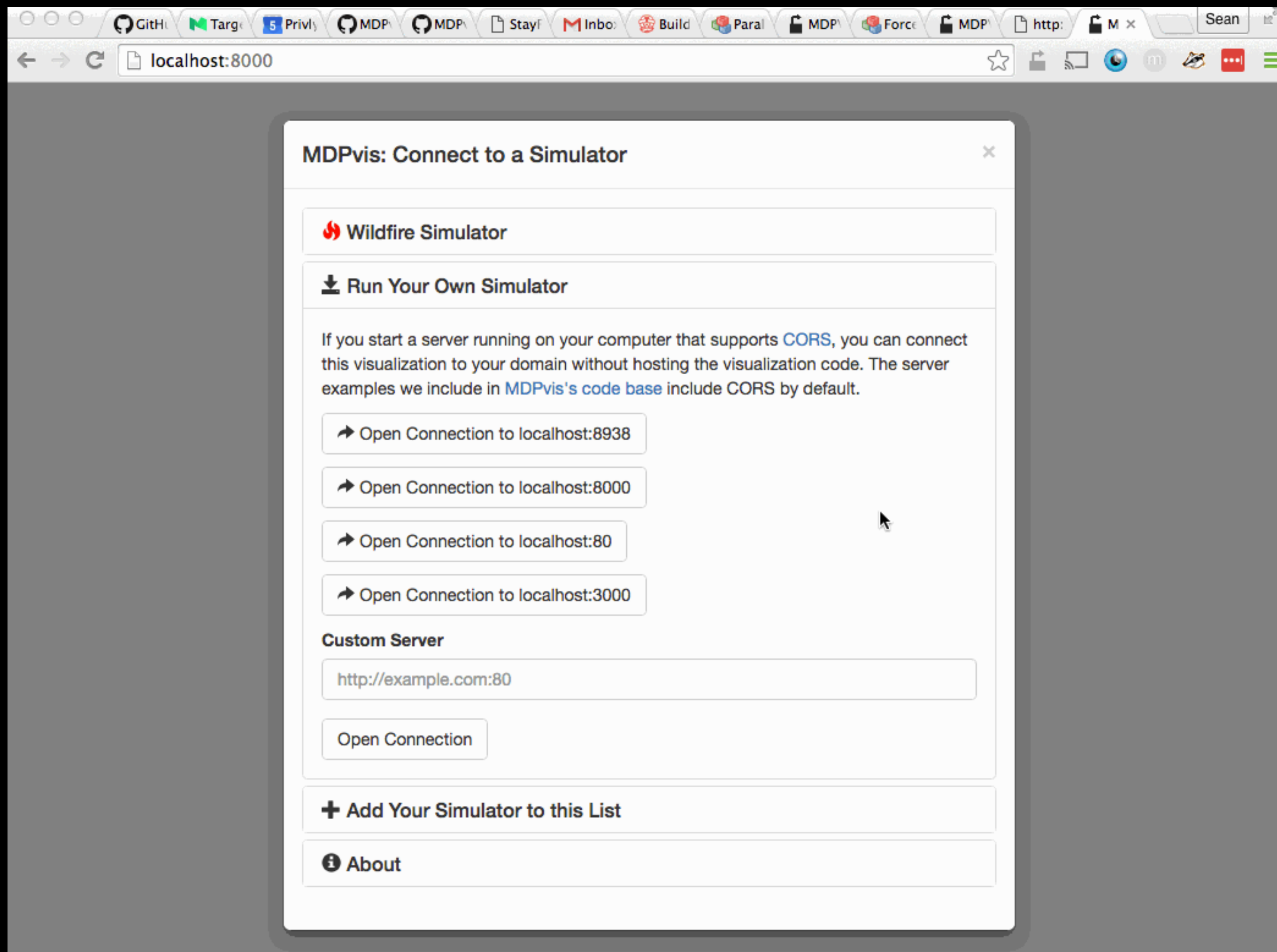
2. Generate a Final Set of Trajectories



3. Visualize!



Results


localhost:8000
Sean
MDPvis: Connect to a Simulator
Wildfire Simulator
Run Your Own Simulator
If you start a server running on your computer that supports CORS, you can connect this visualization to your domain without hosting the visualization code. The server examples we include in MDPvis's code base include CORS by default.
Open Connection to localhost:8938
Open Connection to localhost:8000
Open Connection to localhost:80
Open Connection to localhost:3000
Custom Server
http://example.com:80
Open Connection
Add Your Simulator to this List
About

Modeling Trajectory Synthesis

Much simpler than learning a full predictive model

Step 1: Sample a policy space to create a dataset
Step 2: Create a similarity metric for states

Similarity Metrics

State

~~Weather~~
~~Ignition~~

Action

Suppress or Let Burn

?

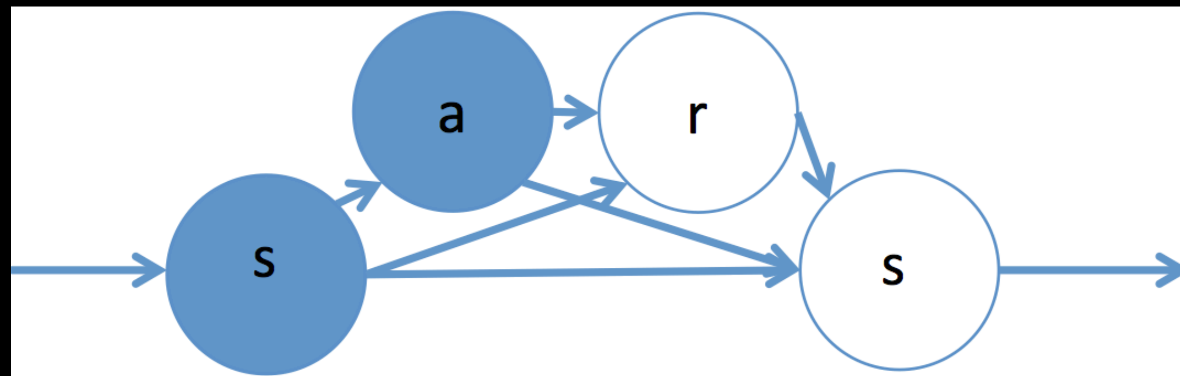
~~Simulate~~ **<u>Sample</u>**

Fire Spread Growth Harvest

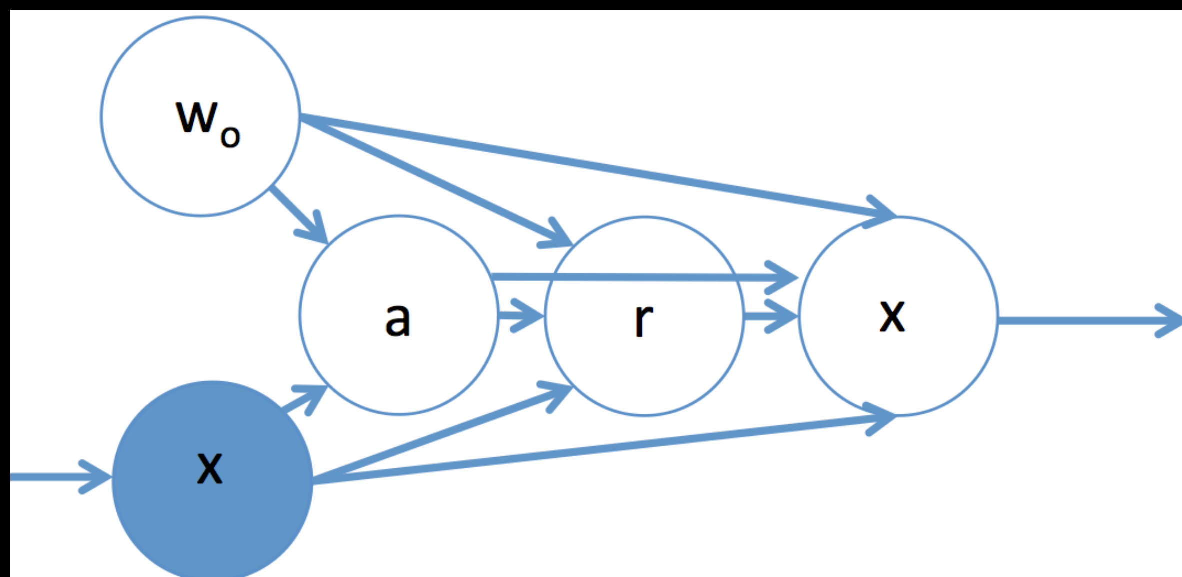
Eliminate Exogenous State!

S: Weather, Ignition Location, and Landscape

Standard State Transition

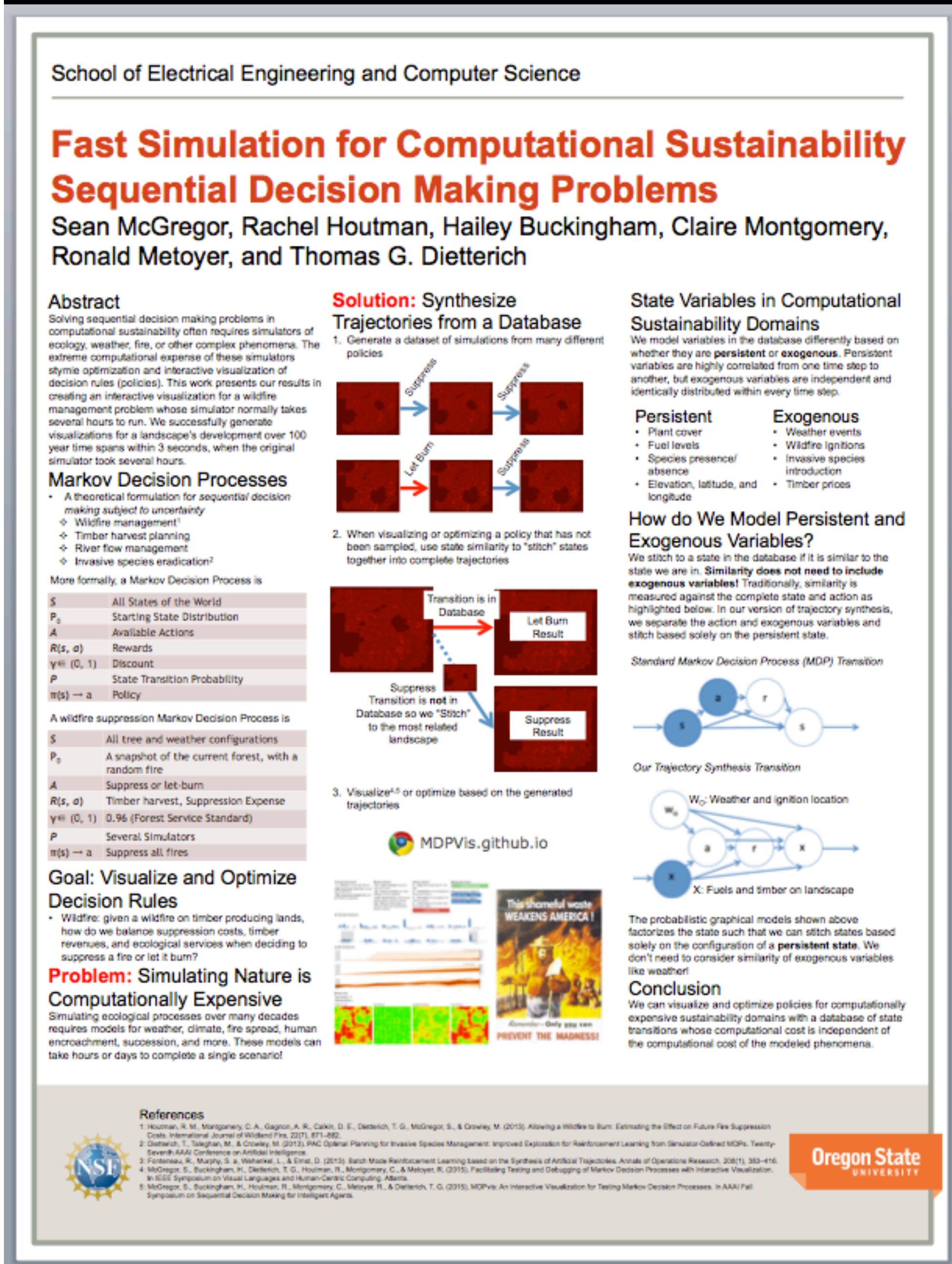


Factored State Transition



W_0: Weather and Ignition Location

X: Landscapes



Thanks!

Collaborators: Rachel Houtman, Hailey Buckingham, Claire Montgomery, Ronald Metoyer, Thomas Dietterich

Funder: NSF

Contact: CompSust@seanbmcgregor.com

Questions?