THE UNIVERSITY OF BRITISH COLUMBIA DEPARTMENT OF COMPUTER SCIENCE LABORATORY FOR COMPUTATIONAL INTELLIGENCE

Seeing the Forest Despite the Trees : Large Scale Spatial-Temporal Decision Making

Mark Crowley - Computer Science
http://www.cs.ubc.ca/~crowley

John Nelson - Forestry Resource Management
http://farpoint.forestry.ubc.ca/FP/search/Faculty_View.aspx?FAC_ID=3171

David Poole - Computer Science
http://www.cs.ubc.ca/~poole

Forestry Planning Problem

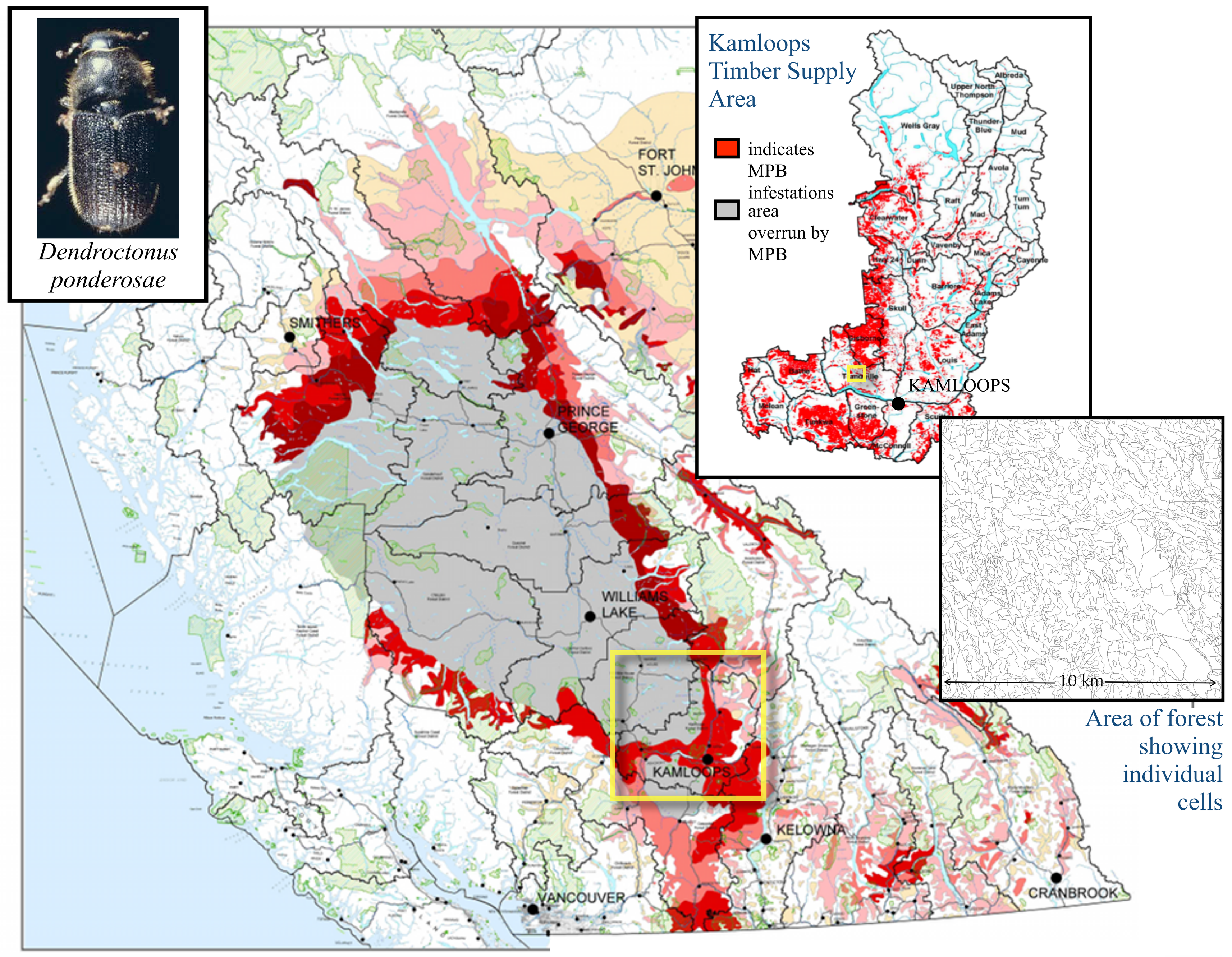
Forest planning area is divided spatially into thousands of **cells** called stands.
The planning problem is to :

- determine an **action at each cell for each year**
- satisfy complex **spatial constraints**
- maximize a complex **utility function** based on local cell state and landscape-wide patterns

Properties of the problem :

- very **long term** planning needed (300 years)
- **state uncertainty** exaggerated by **Mountain pine beetle infestation**
- complex dynamics defined in domain-specific **simulation models** which are expensive to sample from

We call problems with this structure **Large Scale Spatial-Temporal (LSST)** planning problems.
We provide a solution using **policy gradients.**



Area of forest showing individual cells

Current Practice in Forestry

Strategic Plan : entire province/region

- **Annual Allowable Cut (AAC)**
- manages social, political, ecological trade-offs

Tactical Plan : each forestry company in one region

- maximize return
- minimize fines incurred
- maintain sustainable flow of lumber

Methods currently used in forestry :

- mostly deterministic models
- do not take uncertainty into account while planning
- cannot deal with spatial relations

Examples :

- linear programming
- simulation modelling
- simulated annealing

Example Strategic Policy Report

Taken from : *Kamloops timber supply area : Rationale for allowable annual cut (AAC) Determination.*
http://www.for.gov.bc.ca/hts/tsa/tsa11/docs.htm

details of establishing the most fair, equitable and operationally effective means of administering these partitions.

In conclusion, I will reiterate that the primary purpose of the 'non-pine' partition is to maintain options to support the currently projected mid-term harvest level. In the short term, the implementation of this capped partition should allow for licensees to harvest in and sustain operations from Douglas-fir-, spruce- and balsam-leading stands during times when pine may not be economically viable, with the proviso that at the end of the five-year period the total amount of five times 1.7 million = 8.5 million cubic metres of harvested, Douglas-fir, spruce and balsam is not exceeded; this provides for continuation of the harvest at close to recent levels. To any extent that this were to be exceeded, the mid-term level would be compromised. Therefore, if ongoing monitoring shows a consistent trend toward over-reaching this limit, I may revisit this determination earlier than required by statute, to specify a new level of harvest in the 'non-pine' partition, consistent with achieving the mid-term level.

The partitions should also ensure that the AAC I have determined will not become a limiting factor on the priority for aggressive harvesting in the MPB-damaged pine stands.

Determination

Having considered and reasoned from all of the factors as documented above, including the risks and uncertainties in the information provided, it is my determination that:

A timber harvest level that accommodates as far as possible the objectives for all forest resources during the next five years, including the increased harvesting necessitated by the Mountain Pine Beetle infestation, that reflects current management practices as well as the socio- economic objectives of the Crown, and that includes the required adjustment in respect of the change in accounting for interior log grades, can be best achieved in the TSA by establishing an AAC of 4.0 million cubic metres, effectively a net reduction from the former AAC of 8.1 percent.

This new AAC includes partitions specifying allowable annual harvest volumes attributable to the following:

- for harvesting in stands predominated by Douglas-fir, spruce, or balsam, a maximum of 1 700 000 cubic metres (referred to in this document as the 'non-pine' partition;
- for harvesting of pine species, 1 994 000 cubic metres, with the possibility of an increase to the extent of any under-harvesting in the 'non-pine' partition;
- for harvesting in cedar- or hemlock- leading stands, 200 000 cubic metres;
- for harvesting in PA 16, 86 000 cubic metres; and
- for harvesting in deciduous-leading stands outside PA 16, within the Headwaters District, 20 000 cubic metres;

This determination, which excludes all woodlot licence volumes, becomes effective on June 1, 2008, and will remain in effect until a new AAC is determined, which must take place within five years of the effective date of this determination.

Implementation

In the period following this decision and leading to the subsequent determination, I encourage BCFS staff and licensees to undertake the tasks and studies noted below, the particular benefits of which are described in appropriate sections of this rationale document. I recognize that the ability of staff and licensees to undertake these projects is dependent on available resources including funding. These projects are, however, important to help reduce the risk and uncertainty associated with key factors that affect the timber supply in the Kamloops TSA.

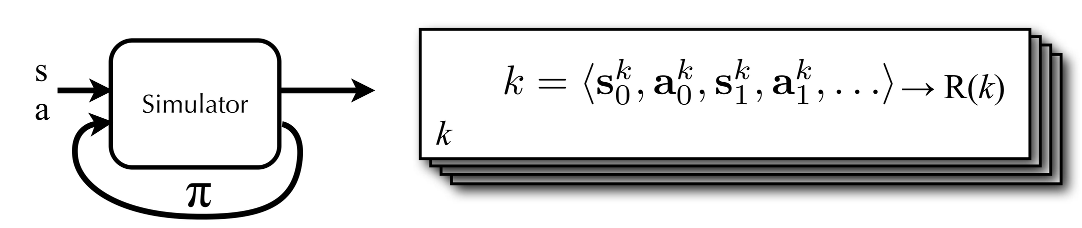
Page 52

Solution : LSST Policy Gradients

Motivation for using policy gradients:

- only simulation models available
- decomposable policy ⟶ decomposable gradient

1. Generate Trajectory



2. Update Policy

expected value of policy computed by combining weighted rewards from all trajectories

$$V^{\boldsymbol{\theta}} = \int_0^\infty p(k|\boldsymbol{\theta}) R(k) dk$$

approximate gradient using trajectories seen so far. Dynamics are not needed

$$\begin{aligned}\nabla_{\boldsymbol{\theta}} V^{\boldsymbol{\theta}} &\approx \frac{1}{|K|} \sum_k \nabla_{\boldsymbol{\theta}} \log p(k|\boldsymbol{\theta}) R(k) \\ &= \frac{1}{|K|} \sum_k \sum_t \nabla_{\boldsymbol{\theta}} \log \boldsymbol{\pi}(\mathbf{s}_t^k, \mathbf{a}_t^k, \boldsymbol{\theta}_t) \\ &= \frac{1}{|K|} \sum_k \sum_t \sum_c \nabla_{\theta} \log \pi(\mathbf{s}_t^k[c], \mathbf{a}_t^k[c], \boldsymbol{\theta}_t^C[c])\end{aligned}$$

gradient of cell policy

$$\begin{aligned}\nabla_{\alpha f} \log \pi(s,a,\theta) &= s[f](1-\pi(s,\alpha,\theta)) &&: \text{if } \alpha = a \\ &\quad - s[f]\pi(s,\alpha,\theta) &&: \text{if } \alpha \neq a\end{aligned}$$

$$\textit{where } \alpha \in A \textit{ and } f \in F \textit{ comprise } \theta$$

update the policy using the gradient scaled by a learning rate

$$\boldsymbol{\theta}' = \boldsymbol{\theta} + \mu \nabla_{\boldsymbol{\theta}} V^{\boldsymbol{\theta}}$$

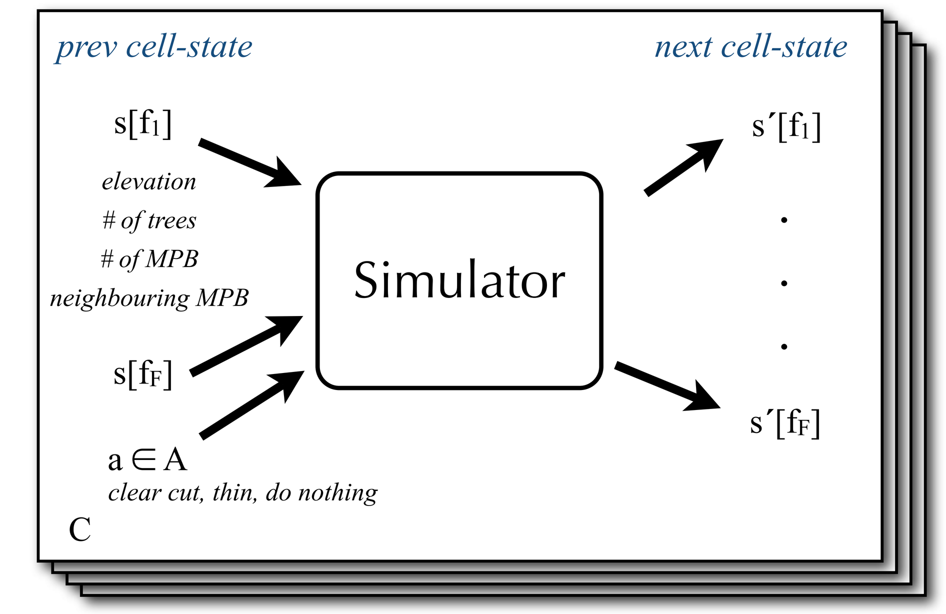
extensions to policy gradients

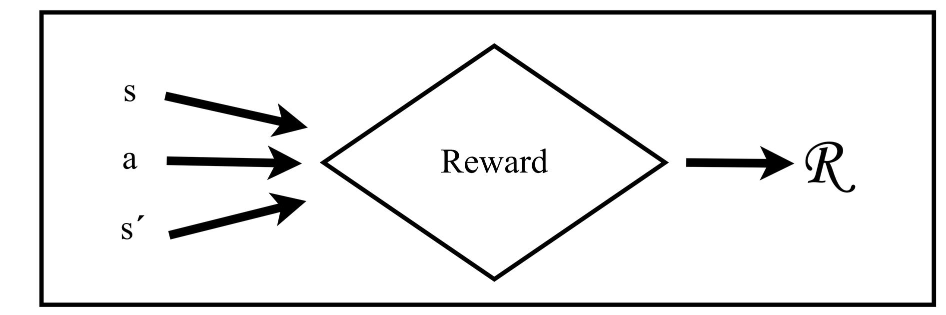
In our implementation we used some standard techniques to reduce the variance of the gradients via (Riedmiller, 2007):

- *optimal baselining* - subtract computed term from *R(k)*
- *Rprop* - follow direction of gradient only, not magnitude

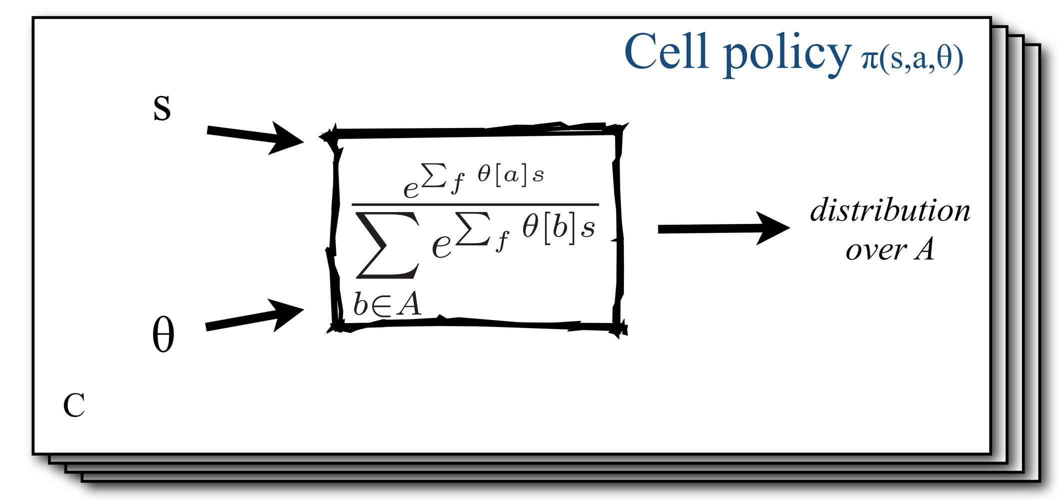
Components of our Algorithm

Input

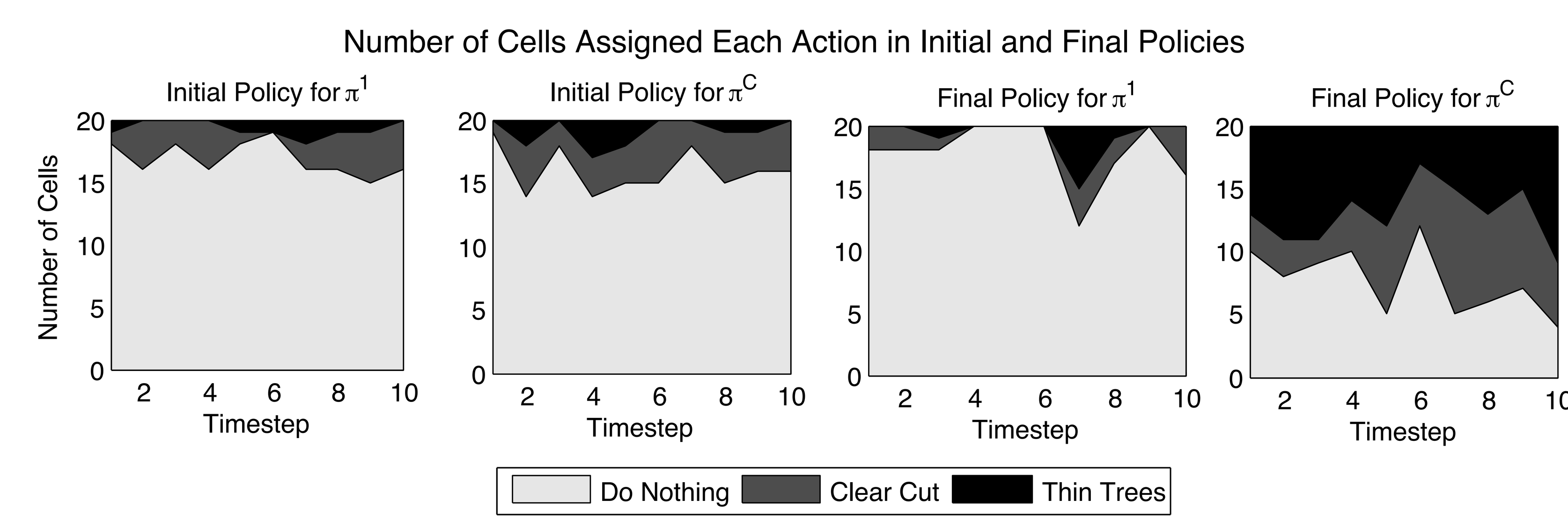




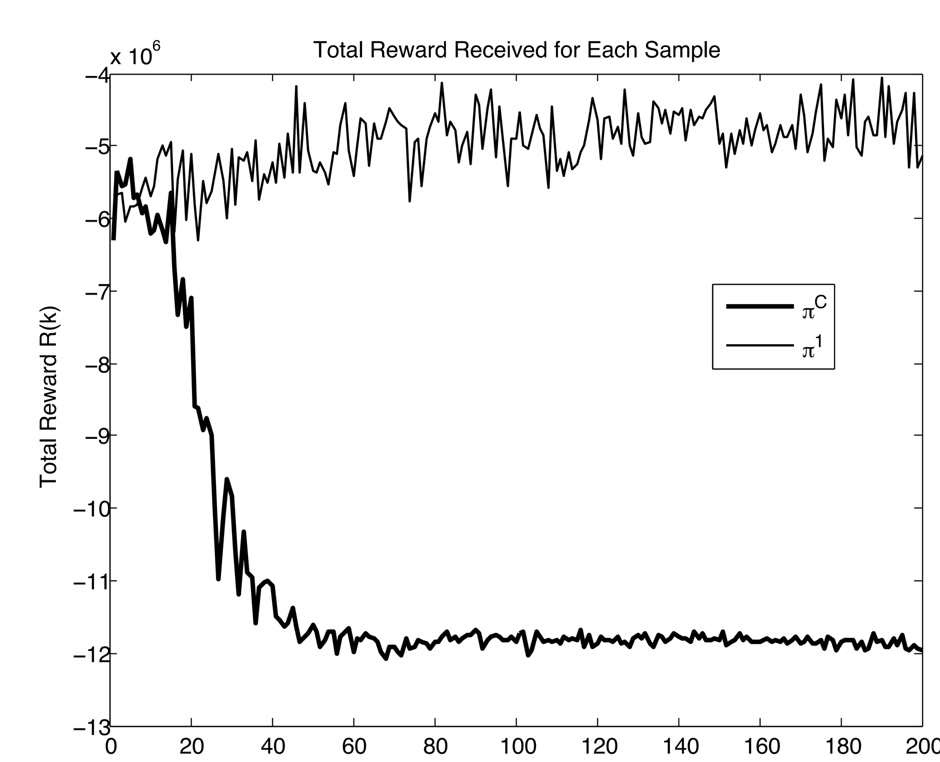
Output



Results



explicit policy: $\pi^C(s,a,\theta^C)$ = separate $\boldsymbol{\theta}$ for each cell
abstract policy : $\pi^1(s,a,\theta^1)$ = single, shared $\boldsymbol{\theta}$ for every cell
setup : 200 sampled trajectories, 10 timesteps, 20 cells.
initial policy : Do Nothing = 80% | Clear Cut = 15% | Thin = 5%



setup : 20 trials, 200 trajectories, 5 timesteps, 5 cells
initial policy :
Do Nothing = 100% | Clear Cut = 0% | Thin = 0%

Summary

Abstract policy algorithm achieves higher rewards than explicit policy algorithm and does not get stuck in local minima.

Explicit policy πC produces deterministic policies for each cell
(eg. "always clear cut in cell 5 at timestep 3")

Abstract policy π1 finds more robust by not fixating on individual cells.

Resulting policy is

- spatially stationary
- stochastic