A Comparison of Genetic Algorithms and Reinforcement Learning for Optimising Sustainable Forest Management

Verena Rieser, Derek T. Robinson, Dave Murray-Rust, Mark Rounsevell

School of GeoSciences, University of Edinburgh, UK
Telephone: (+44) (0)131 650 2270
Fax: (+44) (0)131 650 2524
Email: verena.rieser@ed.ac.uk

1. Introduction

Sustainable forest management is defined as "*the stewardship and use of forests and forest lands in a way, and at a rate, that maintains their biodiversity, productivity, regeneration capacity, vitality and their potential to fulfil, now and in the future, relevant ecological, economic and social functions* […]." (MCPFE, 1994). As such, forest management has to satisfy multiple and often conflicting goals. Furthermore, forest planning is characterised by the long-term horizon of its outcomes. Since long-term plans are made in the face of uncertain futures, long-term sustainable forest management should incorporate some measure of risk. Uncertainty emerges from a variety of sources, including irregular or unknown fluctuations in the demand for timber, or the occurrence of extreme events. In addition, forest management is dynamic in time and space, for example, different stands have different properties, and the likelihood of stochastic events may change over time. Forest planning may be suboptimal if it ignores these sources of uncertainty and risk.

Previous work on multi-objective optimisation in forest management has mainly used heuristic search methods. For example, Bettinger et al. (2002), Pukkala and Kurttila (2005) compare various heuristic optimisation techniques and conclude that Genetic Algorithms (GAs) perform well for more complex spatial problems. However, the studies did not investigate the algorithms' performance under uncertainty.

Reinforcement Learning (RL) is an alternative approach for optimal policy selection. RL is a Machine Learning approach frequently used with agent-based systems (Sutton and Barto, 1998). Contemporary research using RL in the context of forest management has shown that it can find robust optimal solutions to multi-objective forest management problems, e.g. (Bone and Dragicevic, 2009). To further explore the potentials that RL provides over heuristic optimisation approaches, we perform a systematic comparison between RL and GA for sustainable forest management for tasks with increasing uncertainty.

2. Problem Descriptions for Sustainable Forest Management

We present several different hypothetical task environments that are used to test the performance of GA and RL. The task descriptions are meant to provide a proof-ofconcept and are not striving to incorporate the multitude of complex factors in a real-world task environment. In particular, we investigate three aspects of the forest management problem with increasing levels of uncertainty: (1) multi-objective planning, (2) temporal planning with increasing uncertainty over time, (3) planning in environments, which are dynamic in time and space.

The overall task is to decide on a management option for a forest management unit (a "cell"), where the two management options available are to *preserve* or to *harvest* a cell. For task types (1) and (3) the optimisation task is to decide *how many* cells to harvest according some trade-off, reflected in the multi-objective goal. The forest is composed of 10 cells, where the decision for each of the cells is made sequentially. Task type (2) deals with temporal decision making, where the optimisation task is to decide *when* to harvest an individual cell over 10 time intervals.

2.1 Task 1: Multi-objective goal

The multi-objective goal implements the trade-off between economic return versus forest conservation: to satisfy the existing demand for timber while cutting as few forest cells as possible. Equation (1) formulates the objective as a weighted sum:

$$objective = w_f \times forestCells + (-w_d) \times unsatisfiedDemand; \quad (1)$$

We assume that the environment is static and behaves in a deterministic way, e.g. the demand can always be satisfied by harvesting five cells, and each cell has the same potential to satisfy demand.

2.2 Task 2: Increasing uncertainty over time

In Task (2) we explore uncertainty, which is introduced by the temporal nature of forest management. Within our modeling framework uncertainty increases over time, which is operationalised as an increasing probability of disturbance affecting a forest cell.

2.3 Task 3: Spatial Dynamics

In Task (3) we model the likelihood of forest disturbance as a function of tree age, similar to Bone and Dragicevic (2009). However, we extend the model to also include the spatial proximity to neighbouring cells and their average age. This implements the notion that forest disturbances tend to spread. The likelihood of forest disturbance is now a linear function of the cell's own age and the average age of its neighbouring cells, where we use a Moore neighbourhood. The cell's age is also positively related to the amount of demand it can satisfy: the older the cell, the more demand it can satisfy.

3. Problem Implementation

3.1 Problem Implementation in RL

RL addresses the problem of how a forest manager should take actions in an uncertain environment so as to maximise some notion of cumulative, long-term utility or "reward". RL uses Markov Decision Processes (MDPs) as its underlying representation for decision making and learning. At each time step t the process is in some state s_t and the forest manager may choose any action $a(s)$, that is available in state s. The process responds at the next time step by moving into a new state s' according to the probability $P(s'|s,a)$, which is defined by the transition function $T_{ss'}$, and giving the decision maker a corresponding reward $R_{ss'}$ (see (Sutton and Barto, 1998) for further details). In our case, the reward corresponds to the multi-objective goal as formulated by Equation (1).

We use an implementation of the well-known SARSA algorithm. The state-action space of the MDP is defined as in fig.1.

$$\left[\text{ACTION:} \begin{bmatrix} \texttt{harvest} \\ \texttt{perserve} \end{bmatrix} \text{STATE:} \begin{bmatrix} \texttt{forestCells}\{1\text{–}10\} \\ \texttt{demandSatisfied:}\{0/1\} \\ \texttt{forestCycle:}\{0\text{–}10\} \\ \texttt{ageCell:}\{\texttt{young/med/old}\} \\ \texttt{ageNeighbours:}\{1\text{–}300\} \end{bmatrix} \right]$$

Figure 1. RL State-action space for the forest management problem

The state space keeps track of the number of preserved *forestCells* and whether the *demand* is satisfied or not. The feature *forestCycle* is only used for Task 2 to keep track of the temporal progression. The feature *ageCell* and *ageNeighbours* are only used for Task 3.

3.2 Problem Implementation in GA

Genetic algorithms (GAs) use mechanisms inspired by biological evolution: reproduction, mutation, recombination, and selection (see (Holland, 1975) for further details). We implement GA using binary encoding, as widely used in the forest modelling community, e.g. (Falcao and Borges, 2001; Pukkala, 2006). A gene represents a cell and an allele a binary forest management option. For Task (1) and Task (3), a chromosome represents the whole forest of 10 cells and the binary options represent preserve or harvest. For the temporal problem type in Task (3), a chromosome represents same cell over time. The fitness function corresponds to the multi-objective goal in Equation (1).

4. Results

RL outperforms GA with increasing significance the more uncertainty is introduced into the planning environment. We explain RL's superior performance by its ability to explicitly represent uncertainty in its transition function and to monitor dynamic changes in the environment in its state-space. Table 1 summarises the results and reports the average performance of RL and GA in terms of their average objective value (see Equation 1). We compare them for significant differences using a 2-tailed paired Student's T-test (n=300). Note that, subtasks (denoted by x.x) use different weights in their objective function. We will discuss and interpret the results in more detail in the full version of the paper.

Task	GA	RL
Task 1.1	95.00 (±0.00)	95.00 (±0.00)
Task 1.2	5.00 (±0.00)	5.00 (±0.00)
Task 2.1	-6.47 (±9.03)	-4.63 (±8.23) *
Task 2.2	12.27 (±8.25)	14.13 (±6.41) **
Task 3	-10.80 (±31.85)	15.00 (±13.09) ***

Table 1. Comparing mean performance of RL and GA for task types with increasing uncertainty, where *denotes $p<0.01$, ** denotes $p<0.005$, and *** $p<0.001$.

5. Discussion

Our implementation of GA follows a binary encoding as widely used in the forest modelling community. Unlike RL, this implementation of GA doesn't have an internal representation of the decision process, e.g. feature states, transition probabilities, or the expected return of taking an action in a state, as used by MDPs. In future work, we will investigate the performance of advanced evolutionary algorithms, such as Linear Classifier Systems (Holland, 1975). We will also test the algorithms with real data.

7. References

P. Bettinger, D. Graetz, J. Sessions and W. Chung. Eight heuristic techniques applied to three increasingly difficult wildlife planning problems. Silva Fennica 36(2):561-584.

C. Bone and S. Dragicevic. GIS and intelligent agents for multiobjective natural resource allocation: A Reinforcement Learning approach. Transactions in GIS, 13:253–272, 2009.

A.O. Falcao and J.G. Borges. Designing an evolution program for solving integer forest management scheduling models: an application in Portugal. Forest Science, 47(2):158—168, 2001.

J. H. Holland. Adaptation in Natural and Artificial Systems. University of Michigan Press, 1975.

MCPFE. Ministerial conference on the protection of forests in Europe. Documents, 16-17 June 1994.

T.Pukkala and M. Kurttila. Examining the performance of six heuristic optimisation techniques in different forest planning problems. Silva Fennica, 39(1), 2005.

T. Pukkala. The use of multi-criteria decision analysis and multi-objective optimisation in forest management. In Sustainable Forest Management: Growth Models of Europe. Springer Berlin / Heidelberg, 2006.

R. Sutton and A. Barto. Reinforcement Learning: An Introduction. MIT Press, 1998.