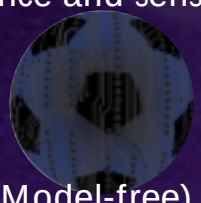
Simulation and Reinforcement Learning with Soccer Agents

Jinsong Leng and Lakhmi Jain KES School and Information Engineering University of South Australia
Colin Fyfe Applied Computational Intelligence Research Unit The University of Paisley Scotland

Research Objective

This research presents a novel reinforcement learning technique in stochastic and dynamic environments

- Investigate simulation-based optimisation and reinforcement learning techniques.
- Define the state space reduction methodology for large, stochastic and continuous spaces.
- Analyse and investigate the convergence and sensitivity of reinforcement learning.



Research Methodologies

- Adopt Temporal Difference learning (Model-free) combining dynamic programming and Monte Carlo techniques.
- Utilise eligibility traces for improving the learning speed and a linear approximation function technique known as Tile Coding to avoid the state space growing exponentially.
- Analyse the convergence and sensitivity using statistical techniques.

Simulation System

A soccer game called SoccerBots (Fig.1) is adopted as simulation environment. The learning will involve the following problems:

- Competitive skills: can the individual agents learn to intercept, shoot, dribble, clear, etc.
- Communication and cooperative skills: when and how to pass; when and where to move to receive a pass, etc.
- Learning and agent teaming: is there an interaction between the individual learning and that necessary for the team to perform optimally.

Empirical Results

- Successfully learn the desire skills.
- Convergence to optimality (Fig. 2).
- Improvement on learning speed by analysing the sensitivity of parameters (Fig. 3).

Conclusion and Future Work

- Proved learning abilities for decision-making in stochastic and unknown environments.
- Enhanced learning rate by analysing convergence and sensitivity.
- May combine with Bayesian learning techniques.

Figure 1 *SoccerBots Simulation Environment*

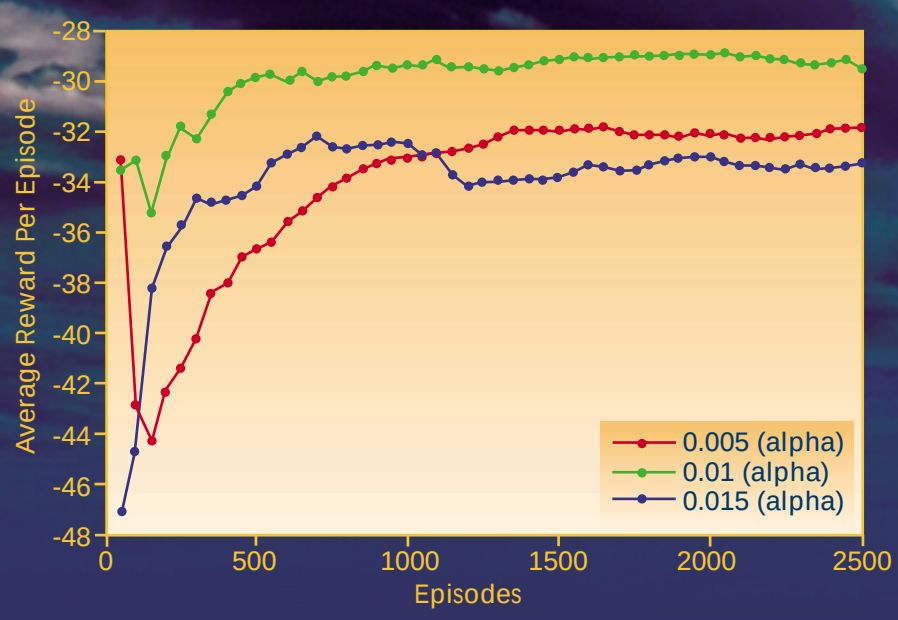


Figure 2 *Convergence with different learning rate*

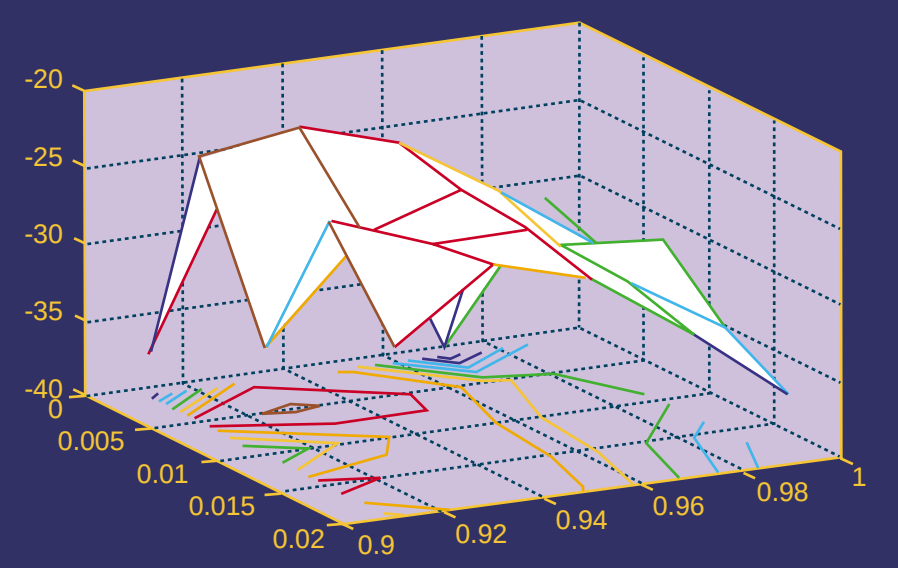


Figure 3 *Optimal values for learning rate and discount rate*