

CSE 497 Engineering Project Presentation

USING HIERARCHIES IN REINFORCEMENT LEARNING FRAMEWORK WITH NON-STATIONARY ENVIRONMENTS

Group Members

Kutalmış Coşkun

Ezdin Aslancı

Advisor

Assoc. Professor Borahan Tümer

Outline

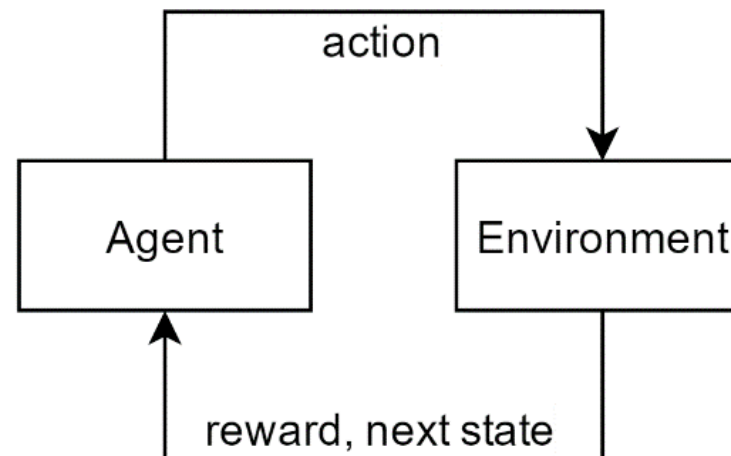
1. Project Description
2. Project Goals
3. Related Work
4. Scope
5. Methodology and Technical Approach
6. Tasks Accomplished
7. Difficulties Encountered
8. Tasks to be Completed in the Second Semester

Project Description

- “Using Hierarchies in Reinforcement Learning Framework with Non-Stationary Environments”
- Few questions pop up:
 - ? What is Reinforcement Learning?
 - ? What is Hierarchical Reinforcement Learning?
 - ? What is a non-stationary environment?
 - ? How to detect changes in the environment?

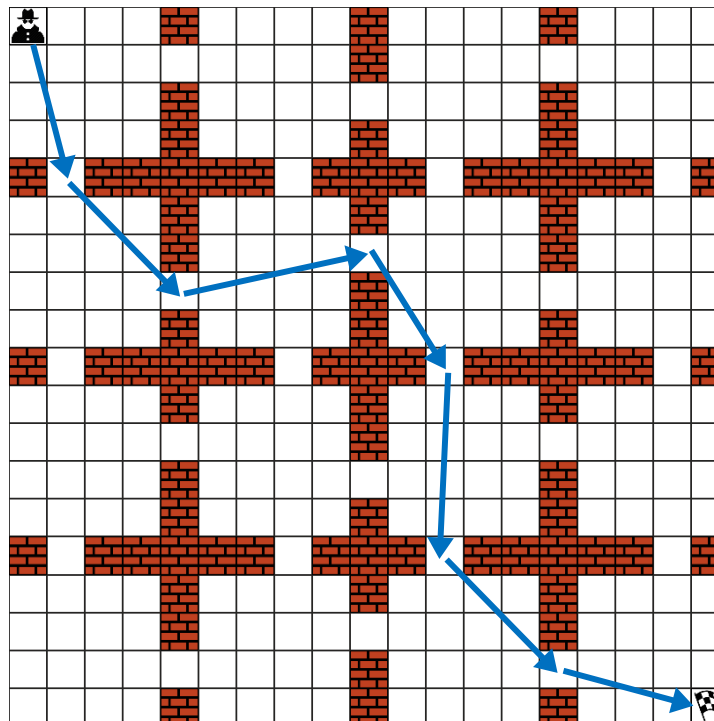
Reinforcement Learning (RL)

- RL is a behavioral learning approach to solve sequential decision making problems.



Hierarchical Reinforcement Learning (HRL)

- RL becomes infeasible when the state space is large or continuous.
- Idea: Divide the problem into solvable independent tasks
- How?: Define and use *options* together with actions.

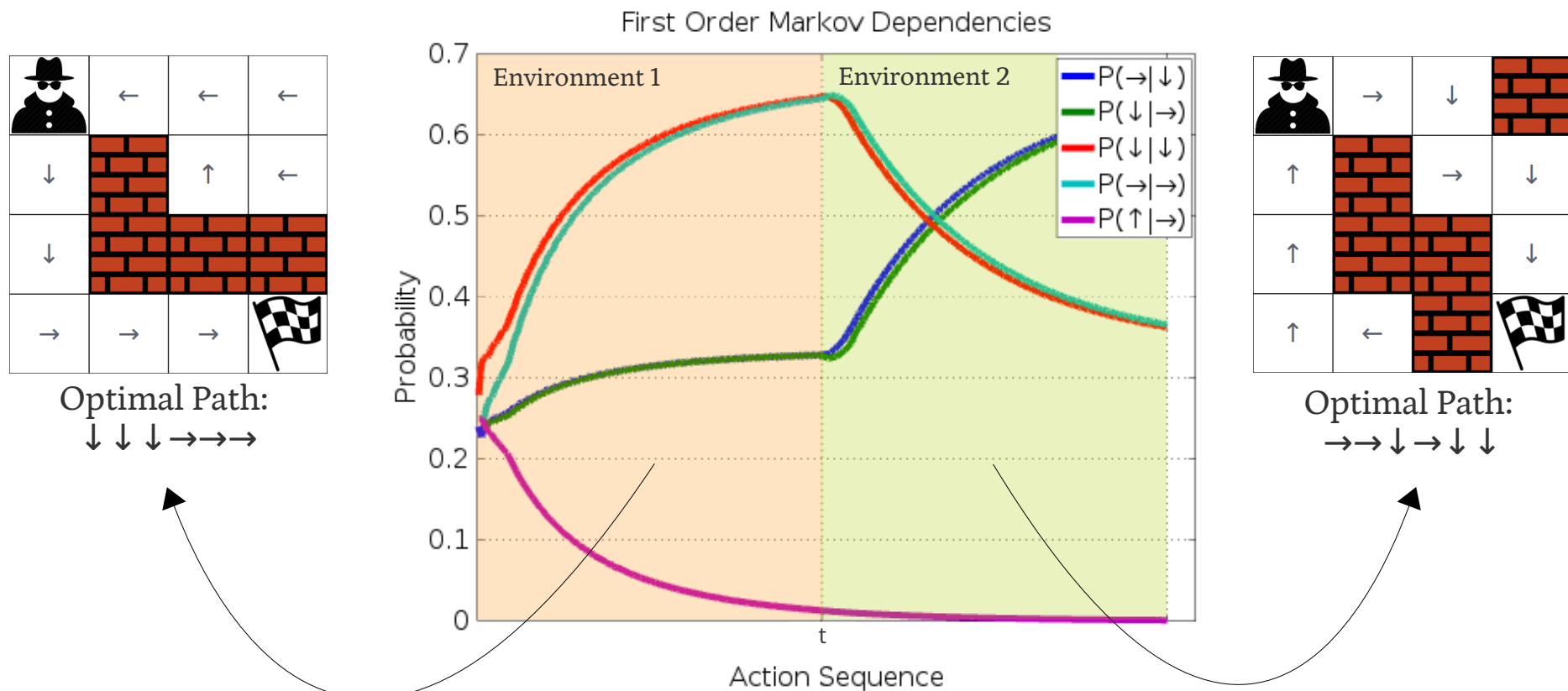


RL in Non-Stationary Environments

- Real world problems are often non-stationary.
- Following properties hold for non-stationary environments:
 - ♦ Multiple stationary regimes with distinct dynamics
 - ♦ Regimes,
 - ✓ change independently of agent's actions
 - ✓ change relatively infrequently
 - ✓ are not directly observable

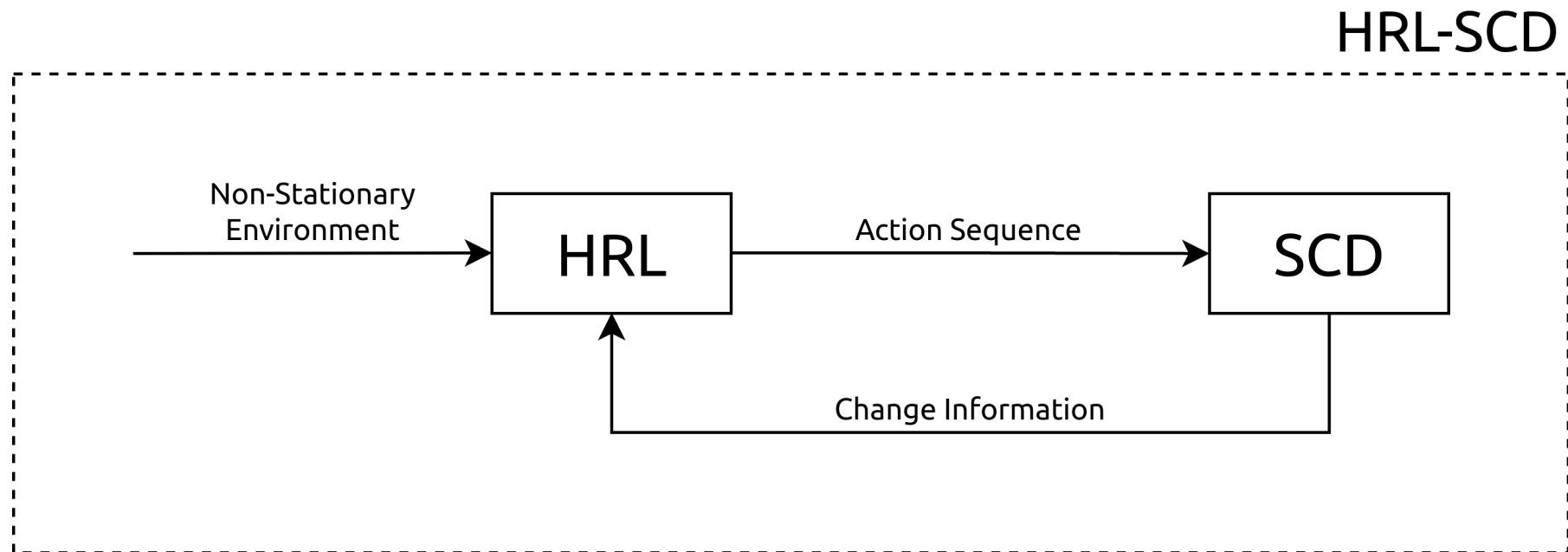
Detecting Changes in Non-Stationary Environments

- Idea: Stationary behaviors (i.e. regimes) of the environment can be distinguished by tracking the action sequence.



Project Goals

- Hierarchical Reinforcement Learning (HRL)
- Change Detection Algorithm (SCD)

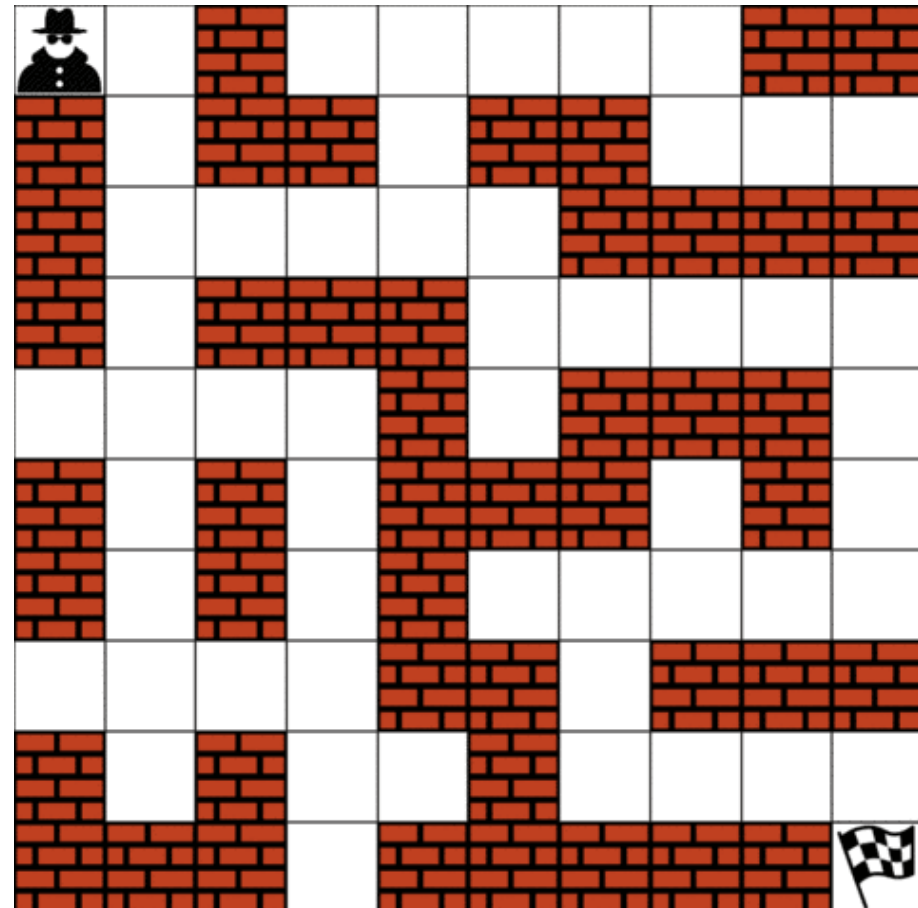


Related Work

- Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning
Sutton, R.S., Precup, D. & Singh, S.
- Dealing with Non-Stationary Environments using Context Detection
Silva, B.D. da, Basso, E.W., Bazzan, A.L.C. & Engel, P.M.
- Hierarchical Reinforcement Learning with Context Detection (HRL-CD)
Yücesoy, Y.E. & Tümer, B.

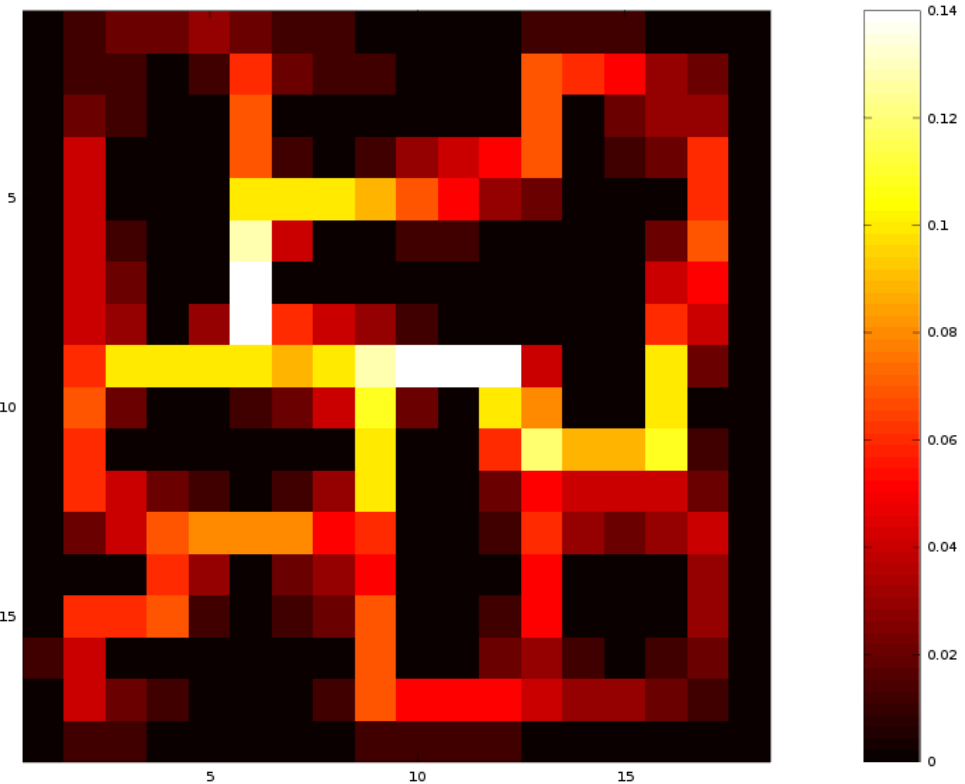
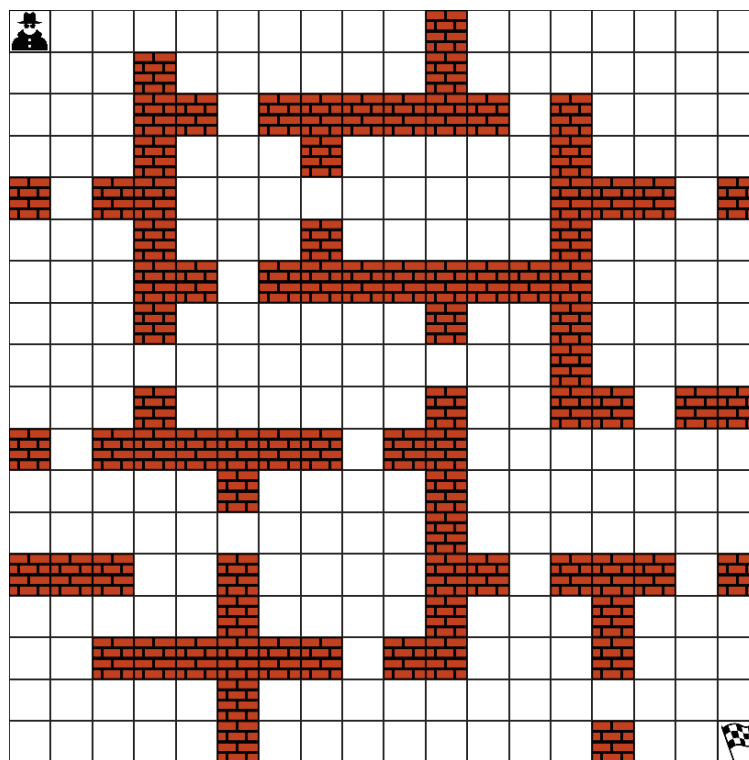
Scope

- Deterministic
- Discrete
- Non-Stationary



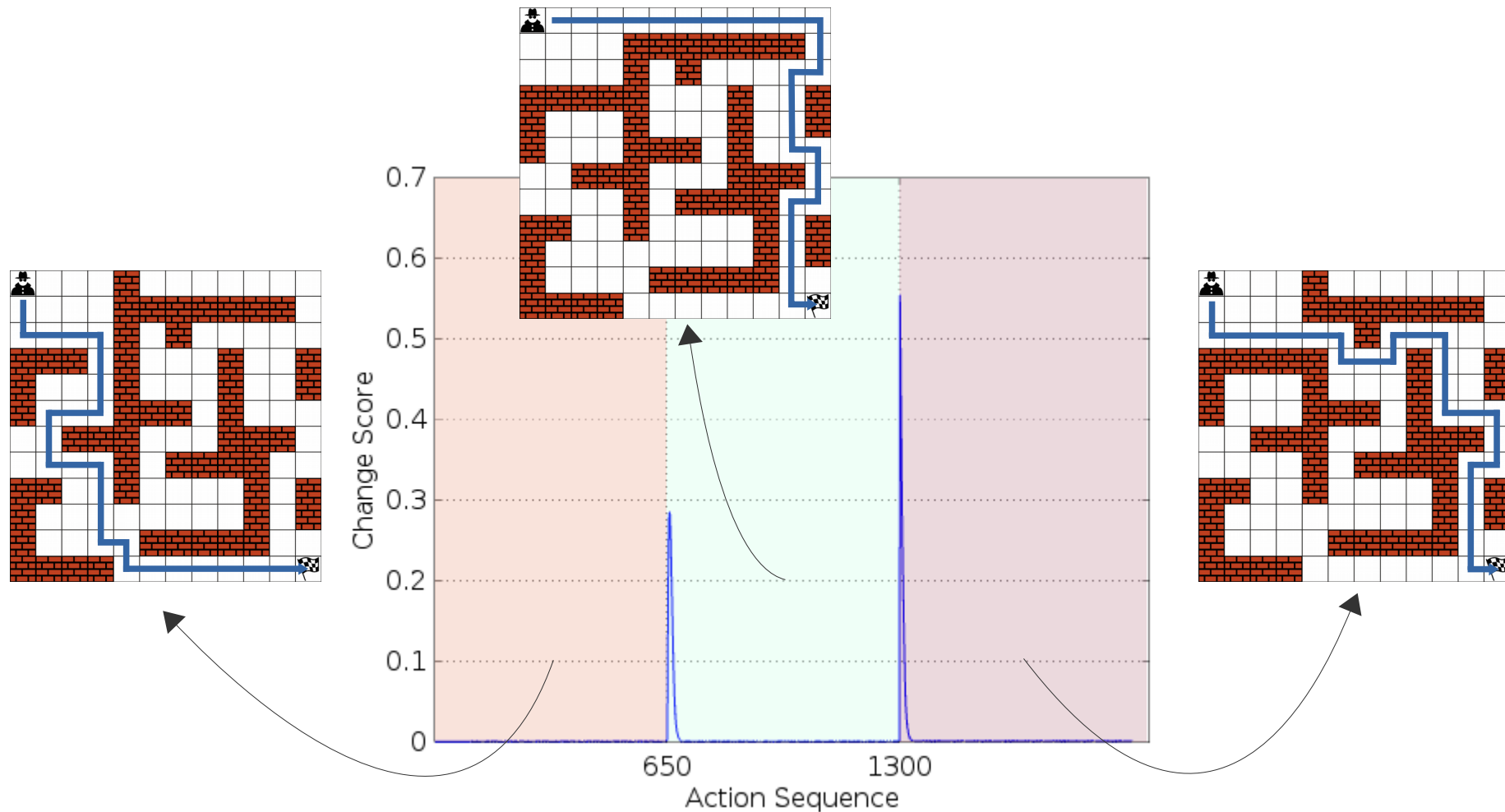
Methodology and Technical Approach

- Subgoal Detection with **Betweenness Centrality**



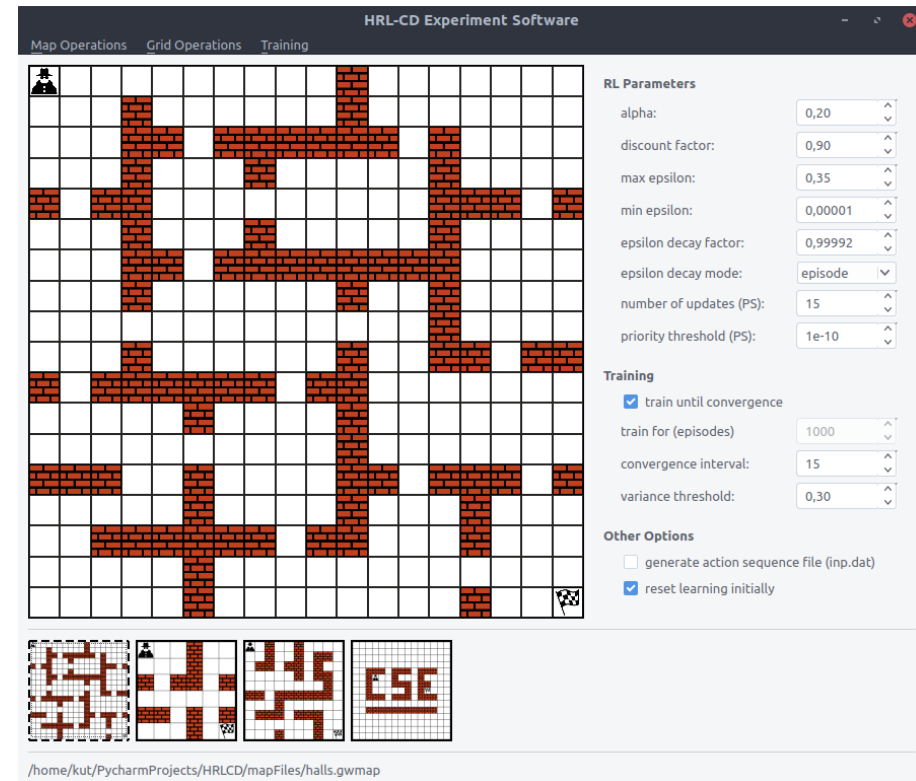
Methodology and Technical Approach

- An experiment on a non-stationary environment with 3 regimes:



What we achieved so far?

- We have found a way to detect changes.
(even slight ones)
- We have implemented:
 - ✓ Change Detection Module
 - ✓ Experiment software with multiple environment support
 - ✓ Q-Learning and Prioritized Sweeping Algorithms
 - ✓ Betweenness centrality analysis module



Difficulties Encountered

- How to detect changes?
- Billions of actions

Plans for Second Semester

- HRL implementation
- Generating a significant structure for each regime
- Using previous experiences on new environments

References

- [1] Silva, B.D. da, Basso, E.W., Bazzan, A.L.C. & Engel, P.M., Dealing with Non-Stationary Environments using Context Detection. 23rd International Conference on Machine Learning (ICML) , 2006.
- [2] Sutton, R.S., Precup, D. & Singh, S., Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. Artificial Intelligence , pages 181-211, 1999.
- [3] Sutton, R.S. & Barto, A.G., Reinforcement learning. Learning 3, 322, 2012.
- [4] Yücesoy, Y.E. & Tümer, M.B., Hierarchical Reinforcement Learning with Context Detection (HRL-CD). International Journal of Machine Learning and Computing , 7763, 2015.
- [5] Oommen, B.J. & Rueda, L. Stochastic learning-based weak estimation of multinomial random variables and its applications to pattern recognition in non-stationary environments. Pattern Recognition 39, 328-341, 2006.