**TCSS 600-A Winter 2014**

**Independent Study Proposal**

**Title: Automated Planning and Reinforcement Learning**

**Student Name: Abdelrahman Elogeel**

**Study Advisor’s Name: Dr. Matthew Alden**

**Justification:**

**This independent study will provide me with a background in automated planning and reinforcement learning. Ideally I would have liked to take a course on automated planning and reinforcement learning techniques, but have instead elected to study the topic as an independent study since no such course was offered last year. This independent study will also help me shape my Master’s thesis background which I plan to work on during the Summer 14 and Autumn 14 quarters.**

**Overview:**

Automated Planning is the reasoning side of acting [1]. It is an abstract, explicit deliberation process that chooses and organizes actions by anticipating their expected outcomes.

Reinforcement learning [2]is an area of machine learning inspired by behaviorist psychology, concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward.

Blending automated planning and reinforcement learning together has proven to result in reactive and adaptive agents capable of solving real-world problems that are in interest for many industries, like robotics, gaming industry, military and deliberative agents. A reactive and adaptive agent can react to environment changes and adapt to the new state. We aim to study automated planning and reinforcement learning and their use to make adaptive and reactive agents. This 10 week study includes: studying topics from the chapters of the selected textbooks “Automated Planning Theory and Practice”, “Reinforcement Learning: An Introduction”, and relevant papers in the areas of Monte Carlo methods, Temporal-Difference learning, planning under uncertainty, decision trees and feature selection.

**Description**:

Automated planning and scheduling is a branch of artificial intelligence that concerns the realization of strategies or action sequences, typically for execution by intelligent agents, autonomous robots and unmanned vehicles. Unlike classical control and classification problems, the solutions are complex and must be discovered and optimized in multidimensional space. Planning is also related to decision theory. In known environments with available models, planning can be done offline. Solutions can be found and evaluated prior to execution. In dynamically unknown environments, the strategy often needs to be revised online. Models and policies must be adapted. Solutions usually resort to iterative trial and error processes commonly seen in artificial intelligence. These include dynamic programming, reinforcement learning and combinatorial optimization. Languages used to describe planning and scheduling are often called action languages.

Reinforcement learning is an area of machine learning inspired by behaviorist psychology, concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward. The problem, due to its generality, is studied in many other disciplines, such as game theory, control theory, operations research, information theory, simulation-based optimization, statistics, and genetic algorithms. In the operations research and control literature, the field where reinforcement learning methods are studied is called approximate dynamic programming. The problem has been studied in the theory of optimal control, though most studies there are concerned with existence of optimal solutions and their characterization, and not with the learning or approximation aspects. In economics and game theory, reinforcement learning may be used to explain how equilibrium may arise under bounded rationality.

The study entails completing the listed readings from Automated Planning: Theory & Practice by Malik Ghallab, Dana Nau, Paolo Traverso, reinforcement learning: An Introduction by Richard S. Sutton, Andrew G. Barto, writing annotated bibliographies for the technical papers identified and doing a comparison between the planning & learning blended techniques and when they can be useful.

**List of Topics**:

The main topics covered in this Independent Study are as follows:

1. Control Rules in Planning, HTN, Control Strategies in Deductive Planning (Chapters 10, 11 and 12)
2. Planning with Time & Resources, Temporal Planning (Chapters 13 and 14)
3. Planning & Resources Scheduling and Planning based on MDP (Chapter 16)
4. Introduction and Evaluation Feedback (Chapters 1 and 2)
5. Monte Carlo Methods, Temporal Difference Learning (Chapters 5 and 6)
6. Eligibility Traces and Function Approximation (Chapters 7 and 8)
7. Learning & Planning (Chapter 9)

**List of Readings**:

1. Text Book: Automated Planning: Theory and Practices by Malik Ghallab, Dana Nau, Paolo Traverso
2. Reinforcement Learning: An Introduction by Richard S. Sutton, Andrew G. Barto
3. Jasmina NOVAKOVIĆ, Perica STRBAC, Dusan BULATOVIĆ “Toward Optimal Feature Selection Using Ranking Methods and Classification Algorithms”, 2011
4. Jimenez, Fernandez and Borrajo, “Integrating Planning, Execution, and Learning to Improve Plan Execution”, 2013

**Deliverables**:

1. Review paper on comparison between possible combinations of planning and learning techniques and when they can be used
2. Summaries of the weekly reading assignments
3. 30-45 minutes presentation on the comparison review paper at the end of the quarter colloquium

**Grading**:

|  |  |
| --- | --- |
| Description | Final Grade |
| Review paper | 40% |
| Presentation | 20% |
| Weekly Summaries | 40% |

**Tentative Schedule**

|  |  |  |  |
| --- | --- | --- | --- |
| **Week #** | **Topic** | **Readings** | **Deliverable** |
| 1 | Control Rules in Planning, HTN, Control Strategies | Chapters 10 and 11 from Automated Planning | Summary |
| 2 | Deductive Planning and Planning with Time | Chapters 12 and 13 from Automated Planning | Summary |
| 3 | Resources, Temporal Planning and Planning & Resources Scheduling | Chapters 14 and 15 from Automated Planning | Summary |
| 4 | Planning based on MDP & Feature Selection | Chapter 16 from Automated Planning and “Toward Optimal Feature Selection Using Ranking Methods and Classification Algorithms” | Summary |
| 5 | Introduction and Evaluation Feedback | Chapters 1 and 2 from Reinforcement Learning: An Introduction | Summary |
| 6 | Monte Carlo Methods, Temporal Difference Learning | Chapters 5 and 6 from Reinforcement Learning: An Introduction | Summary |
| 7 | Eligibility Traces and Function Approximation | Chapters 7 and 8 from Reinforcement Learning: An Introduction | Summary |
| 8 | Learning & Planning | Chapter 9 from Reinforcement Learning: An Introduction and Integrating Planning, Execution, and Learning to Improve Plan Execution | Summary |
| 9 | Authoring the paper |  | Review paper |
| 10 | Final Presentation |  | Final review paper and presentation |

# References

|  |  |
| --- | --- |
| [1] | Wikipedia, "Automated planning and scheduling," Wikimedia Foundation, 19 November 2013. [Online]. Available: http://en.wikipedia.org/wiki/Automated\_planning\_and\_scheduling. [Accessed 2 December 2013]. |
| [2] | Wikipedia, "Reinforcement learning," Wikimedia Foundation, 25 November 2013. [Online]. Available: http://en.wikipedia.org/wiki/Reinforcement\_learning. [Accessed 2 December 2013]. |