Linköping University | Department of Computer and Information Science
Master's thesis, 30 ECTS | Datateknik
2017 | LIU-IDA/LITH-EX-A--2017/001--SE

A very very long title
– with a subtitle

En himla bra svensk titel

Oskar Lundin

Supervisor : Sourabh Balgi
Examiner : Jose M. Peña

External supervisor : Fredrik Bissmarck

Linköpings universitet
SE–581 83 Linköping
+46 13 28 10 00 , www.liu.se

Abstract

The abstract resides in file `Abstract.tex`. Here you should write a short summary of your work.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Pellentesque in massa suscipit, congue massa in, pharetra lacus. Donec nec felis tempor, suscipit metus molestie, consectetur orci. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Curabitur fermentum, augue non ullamcorper tempus, ex urna suscipit lorem, eu consectetur ligula orci quis ex. Phasellus imperdiet dolor at luctus tempor. Curabitur nisi enim, porta ut gravida nec, feugiat fermentum purus. Donec hendrerit justo metus. In ultrices malesuada erat id scelerisque. Sed sapien nisi, feugiat in ligula vitae, condimentum accumsan nisi. Nunc sit amet est leo. Quisque hendrerit, libero ut viverra aliquet, neque mi vestibulum mauris, a tincidunt nulla lacus vitae nunc. Cras eros ex, tincidunt ac porta et, vulputate ut lectus. Curabitur ultricies faucibus turpis, ac placerat sem sollicitudin at. Ut libero odio, eleifend in urna non, varius imperdiet diam. Aenean lacinia dapibus mauris. Sed posuere imperdiet ipsum a fermentum.

Nulla lobortis enim ac magna rhoncus, nec condimentum erat aliquam. Nullam laoreet interdum lacus, ac rutrum eros dictum vel. Cras lobortis egestas lectus, id varius turpis rhoncus et. Nam vitae auctor ligula, et fermentum turpis. Morbi neque tellus, dignissim a cursus sed, tempus eu sapien. Morbi volutpat convallis mauris, a euismod dui egestas sit amet. Nullam a volutpat mauris. Fusce sed ipsum lectus. In feugiat, velit eu fermentum efficitur, mi ex eleifend ante, eget scelerisque sem turpis nec augue.

Vestibulum posuere nibh ut iaculis semper. Ut diam justo, interdum quis felis ac, posuere fermentum ex. Fusce tincidunt vel nunc non semper. Sed ultrices suscipit dui, vel lacinia lorem euismod quis. Etiam pellentesque vitae sem eu bibendum. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Pellentesque scelerisque congue ullamcorper. Sed vehicula sodales velit a scelerisque. Pellentesque dignissim lectus ipsum, quis consectetur tellus rhoncus a.

Nunc placerat ut lectus vel ornare. Sed nec dictum enim. Donec imperdiet, ipsum ut facilisis blandit, lacus nisi maximus ex, sed semper nisl metus eget leo. Nunc efficitur risus ac risus placerat, vel ullamcorper felis interdum. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Duis vitae felis vel nibh sodales fringilla. Donec semper eleifend sem quis ornare. Proin et leo ut dolor consectetur vehicula. Lorem ipsum dolor sit amet, consectetur adipiscing elit.

Nunc dignissim interdum orci, sit amet pretium nibh consectetur sagittis. Aenean a eros id risus aliquam placerat nec ut lectus. Curabitur at quam in nisi sodales imperdiet in at erat. Praesent euismod pulvinar imperdiet. Nam auctor mattis nisi in efficitur. Quisque non cursus ipsum, consequat vehicula justo. Fusce varius metus et nulla rutrum scelerisque. Praesent molestie elementum nulla a consequat. In at facilisis nisi, convallis molestie sapien. Cras id ullamcorper purus. Sed at lectus sit amet dolor finibus suscipit vel et purus. Sed odio ipsum, dictum vel justo sit amet, interdum dictum justo. Quisque euismod quam magna, at dignissim eros varius in. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas.

Acknowledgments

`Acknowledgments.tex`

Contents

List of Figures

1 Introduction

In this thesis project, the problem of searching for targets in unknown but familiar environments is addressed. This chapter presents the motivation behind the project, the research questions that are addressed, and the delimitations.

1.1 Motivation

Searching for targets (objects of interest, goals) in unknown environments is a well-studied problem that appears in many areas ranging from robotics [] to computer vision [] and neuroscience []. Applications include search and rescue,

In many cases, the environment is familiar in that it has characteristics that are similar to previously seen environments.

This is similar to the exploration problem, where a robot is tasked with maximizing the knowledge of a certain area.

If the environment is known, this problem becomes drastically easier.

In this project, an instance of the target search problem is considered where the environment is searched by a pan-tilt camera fixed in place. The camera has a limited view of the environent. Automating this task is of interest for multiple reasons. Manually controlling a camera may be costly, and the performance of a human operator may be suboptimal. Crucial to the problem is generalization.

1.2 Aim

The aim of this thesis is to implement and evaluate an autonomous agent that intelligently searches its environment for targets. The agent should learn common characteristics of environments and utilize this knowledge to search for targets in new environments more effectively.

1.3 Research questions

This thesis will address the following questions:

1. How can the visual search problem be solved by a learning agent?

2. How can a simulator that tests the ability of an agent to solve the presented problem be implemented?
3. How does the learning agent compare to common non-learning methods?

1.4 Delimitations

This thesis will be focused on the behavioural aspects of the presented problem. To train and test agents, a simplified environment will be used. This will test the desired characteristics of the agent as presented above, but will not simulate realistic environments.

2 Theory

This chapter introduces relevant theory and related work

2.1 Active Vision and Visual Search

An *active vision* system is a system that can manipulate the viewpoint of the camera in order to investigate the environment and get better information from it.

Studies in humans have shown that

- We do not use memory during visual search
- We have easier to differentiate unknown distractors from targets (or vice versa?)

2.2 Reinforcement Learning

Reinforcement learning (RL) is a subfield of machine learning concerned with learning from interaction how to achieve a goal. An *agent* and its *environment* interact continually over discrete time steps. At each time step the agent selects some *action* that updates the state of the environment, and gives it a *reward*. The agent selects actions using a stochastic *policy* with the goal of maximizing the *return* which is usually defined as the discounted sum of future rewards.

2.2.1 Markov Decision Process

The RL setup is usually formalized as a (finite) Markov decision process (MDP).

The problem of learning from interaction to achieve a goal is usually framed as a (finite) Markov Decision Process (MDP). For regular MDPs it is assumed that the learning agent has access to some representation of the underlying *state* of the environment which it uses to select *actions*. For many problems this is not true. A partially observable Markov decision process (POMDP) is a generalization of an MDP in which it is assumed that the environment has some well defined underlying latent state, but the agent only perceives a partial *observation* of it from the environment.

A POMDP is formally defined as a 7-tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, \mathcal{R}, \mathcal{T}, \Omega, \gamma \rangle$, where

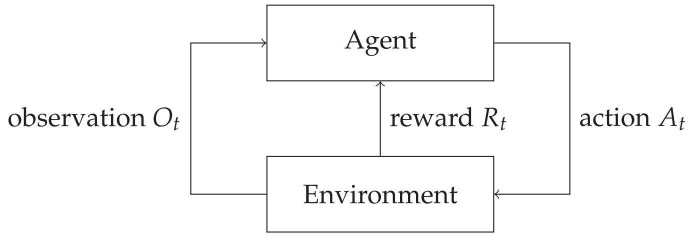

Figure 2.1: Partially observable Markov decision process.

- $\mathcal{S}$ is a finite set of states,
- $\mathcal{A}$ is a finite set of actions,
- $\mathcal{T}$ is a set of conditional state transition probabilities,
- $\mathcal{R} : S \times A \to \mathbb{R}$ is a reward function,
- Ω is a finite set of observations,
- $\mathcal{O}$ is a set of conditional observation probabilities, and
- $\gamma \in [0, 1]$ is a discount factor.

The agent interacts with the environment at discrete time steps $t = 0, 1, 2, \ldots T$. At each time step t, the agent receives an observation of the environment's state $O_t \in \Omega$ and selects some action $A_t \in \mathcal{A}$. In the next time step the agent receives a reward

action $a \in \mathcal{A}$ which causes the environment to transition to state s' with probability $\mathcal{T}(s'|s, a)$. It receives an observation $o \in \Omega$ with probability $\mathcal{O}(o|s', a)$, as well as a reward r given by $\mathcal{R}(s, a)$.

This interaction is repeated until the end of the episode at time step T. The goal of the agent is to maximize the *discounted return*, defined as the discounted sum of future rewards $G_t \doteq \sum_{k=t+1}^{T} \gamma^{k-t-1} R_k$ where γ reflects the uncertainty of the environment.

2.2.2 Policies and Value Functions

Most RL algoerithms estimate both a *value function* that tells the agent how good it is to be in a given state, and a

2.2.3 Policy Optimization

This work will focus on policy optimization algorithms.

2.2.4 Taxonomy of Algorithms

- Model-free vs. model-based
-

2.2.5 Generalization

3 Method

In this chapter, the method is described in a way which shows how the work was actually carried out. The description must be precise and well thought through. Consider the scientific term replicability. Replicability means that someone reading a scientific report should be able to follow the method description and then carry out the same study and check whether the results obtained are similar. Achieving replicability is not always relevant, but precision and clarity is.

Sometimes the work is separated into different parts, e.g. pre-study, implementation and evaluation. In such cases it is recommended that the method chapter is structured accordingly with suitable named sub-headings.

4 Results

This chapter presents the results. Note that the results are presented factually, striving for objectivity as far as possible. The results shall not be analyzed, discussed or evaluated. This is left for the discussion chapter.

In case the method chapter has been divided into subheadings such as pre-study, implementation and evaluation, the result chapter should have the same sub-headings. This gives a clear structure and makes the chapter easier to write.

In case results are presented from a process (e.g. an implementation process), the main decisions made during the process must be clearly presented and justified. Normally, alternative attempts, etc, have already been described in the theory chapter, making it possible to refer to it as part of the justification.

5 Discussion

This chapter contains the following sub-headings.

5.1 Results

Are there anything in the results that stand out and need be analyzed and commented on? How do the results relate to the material covered in the theory chapter? What does the theory imply about the meaning of the results? For example, what does it mean that a certain system got a certain numeric value in a usability evaluation; how good or bad is it? Is there something in the results that is unexpected based on the literature review, or is everything as one would theoretically expect?

5.2 Method

This is where the applied method is discussed and criticized. Taking a self-critical stance to the method used is an important part of the scientific approach.

A study is rarely perfect. There are almost always things one could have done differently if the study could be repeated or with extra resources. Go through the most important limitations with your method and discuss potential consequences for the results. Connect back to the method theory presented in the theory chapter. Refer explicitly to relevant sources.

The discussion shall also demonstrate an awareness of methodological concepts such as replicability, reliability, and validity. The concept of replicability has already been discussed in the Method chapter (3). Reliability is a term for whether one can expect to get the same results if a study is repeated with the same method. A study with a high degree of reliability has a large probability of leading to similar results if repeated. The concept of validity is, somewhat simplified, concerned with whether a performed measurement actually measures what one thinks is being measured. A study with a high degree of validity thus has a high level of credibility. A discussion of these concepts must be transferred to the actual context of the study.

The method discussion shall also contain a paragraph of source criticism. This is where the authors' point of view on the use and selection of sources is described.

In certain contexts it may be the case that the most relevant information for the study is not to be found in scientific literature but rather with individual software developers and open source projects. It must then be clearly stated that efforts have been made to gain access to this information, e.g. by direct communication with developers and/or through discussion forums, etc. Efforts must also be made to indicate the lack of relevant research literature. The precise manner of such investigations must be clearly specified in a method section. The paragraph on source criticism must critically discuss these approaches.

Usually however, there are always relevant related research. If not about the actual research questions, there is certainly important information about the domain under study.

5.3 The work in a wider context

There must be a section discussing ethical and societal aspects related to the work. This is important for the authors to demonstrate a professional maturity and also for achieving the education goals. If the work, for some reason, completely lacks a connection to ethical or societal aspects this must be explicitly stated and justified in the section Delimitations in the introduction chapter.

In the discussion chapter, one must explicitly refer to sources relevant to the discussion.

6 Conclusion

This chapter contains a summarization of the purpose and the research questions. To what extent has the aim been achieved, and what are the answers to the research questions?

The consequences for the target audience (and possibly for researchers and practitioners) must also be described. There should be a section on future work where ideas for continued work are described. If the conclusion chapter contains such a section, the ideas described therein must be concrete and well thought through.