

Artificial Intelligence : Probabilistic & learning techniques

Chris PAL

Project (INF8225)

Farnoush FARHADI & Juliette TIBAYRENC

21 Apr. 2016



Table des matières

| | |
|---|----------|
| Introduction | 2 |
| 1. Reference article | 3 |
| 1.1 Literature review | 3 |
| 1.2 Article summary | 3 |
| 2. Theoretical elements | 4 |
| 2.1 Value iterating & policy iterating vs. Q-learning | 4 |
| 2.2 Reward matrix | 4 |
| 3. Comparing algorithms | 5 |
| 3.1 Classic Q-learning | 5 |
| 3.2 Speedy Q-learning | 5 |
| 3.3 Additional remarks | 5 |
| Further work and conclusion | 6 |
| Bibliographie | 7 |

Introduction

Our AI course has allowed us to discover Markov Decision Processes, which we decided to further explore in our project. Looking for an application for them, we stumbled upon Mastronarde & Van der Schaar's article, 'Fast Reinforcement Learning for Energy-Efficient Wireless Communications'[1], and we chose it as the basis for our work.

The article presents the two researchers' work in developing an algorithm that would enable systems to adopt the best policy to quickly send data packets in the most efficient way possible, an important goal to achieve since the majority of today's surfing is done on smartphones whose battery does not last longer than a day (at best).

In this report, we will first introduce the article, with a quick review of the state of the art at the time it was written (2011), then summarize it. We will continue with a presentation of some theoretical elements, highlighting the differences between what we saw in class and the processes on which the article and our implementations are based, and the way the reward matrix used is calculated (a process that does not appear in the article). Finally we will talk about our own implementation of classic Q-learning and one of its variants, speedy Q-learning, and compare the results we obtain with the results obtained with the authors' algorithm, plus the results we get by using value iteration and policy iteration.

1. Reference article

1.1 Literature review

1.2 Article summary

2. Theoretical elements

2.1 Value iterating & policy iterating vs. Q-learning

2.2 Reward matrix

3. Comparing algorithms

3.1 Classic Q-learning

3.2 Speedy Q-learning

3.3 Additional remarks

Further work & conclusion

Bibliographie

- [1] N. Mastronarde and M. Van der Schaar, “Fast reinforcement learning for energy-efficient wireless communication,” *IEEE Transactions on Signal Processing*, vol. 59, 12 2011.