

Chess Assignment

Introduction to Reinforcement Learning (Spring 2022)

van den Bergh Laurin
institute of informatics
University of Zurich
Zurich, Switzerland
laurin.vandenbergh@uzh.ch
16-744-401

Abstract—This document is a model and instructions for **LaTeX**. This and the **IEEEtran.cls** file define the components of your paper [title, text, heads, etc.]. ***CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

Index Terms—component, formatting, style, styling, insert

I. INTRODUCTION

II. METHODS

III. RESULTS

Question Answers

1. explain algos: Q-learning and SARSA are both temporal-difference (TD) methods for learning Q-values from interaction with an environment. Both can be used for any policy and both try to attribute future rewards to previous actions, these (not actually the rewards get discounted, but the q-value) get discounted with the hyper parameter $\alpha\gamma$. differences: (they behave differently: cliff example) Both algorithms choose their first action according to their chosen policy, e.g. ϵ -greedy. However, they differ in how they choose their second action which is then also used to update the Q-value of the first action (i.e. to compare the second action Q-value?). SARSA is an on-policy algorithm which means that it chooses the second action according to the same policy as it used for choosing the first action. Q-learning, on the other hand, chooses the second action greedily, i.e. it chooses the action for which it thinks is best at that point in time/training.

advantages/disadvantages

q learning taking optimal path, but sarsa taking riskier path. what does that mean for chess? shaping rewards changes the objective.

2. (group only)

3.

4. plot beta and gamma against performance metrics (maybe 3d)

5.

6. (group only)

7. (group only)

IV. CONCLUSION

V. APPENDIX

REFERENCES

- [1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove the template text from your paper may result in your paper not being published.