

Reinforcement Learning Project

Anonymous Authors

Abstract—This document is a rough guide to producing the project report. You should enter the title, but do *not* enter any author names or anything that identifies any of the authors (in any part of the document). The structure (i.e., sections) outlined here is offered as a suggestion, but feel free to change if convenient. And, of course, replace these hints/instructions/examples with your own text. But you must use this IEEE template.

Hint: shared tools like <http://overleaf.com/> are great tools for collaborating on a multi-author report in L^AT_EX. There are also Word templates (<https://www.ieee.org/conferences/publishing/templates.html>) if you wish; but you must convert to pdf format for submission. Recall the page limit of 5 pages.

I. INTRODUCTION

In this section

- 1) An overview of what you did, and why
- 2) Provide access to your anonymous code¹.

II. BACKGROUND AND RELATED WORK

It is absolutely essential to provide sufficient background to your work. Elaborate (*in your own words*) the material required to understand your work (for someone who has attended the course but needs a pedagogical reminder about the parts relevant to your project).

References are essential. You may cite lectures, e.g., [2], book chapters, e.g., Chapter x from [1], or articles from the literature, e.g., [3], [4], even blog posts and code repositories. In all cases you *must* properly cite any work that is not your own.

Don't hesitate to use and reference equations, but rigorously check that each part of your notation is introduced clearly. For example, Eq. (1) is a multi-label prediction, conditioned on input $\mathbf{x} \in \mathbb{R}^D$ (where D is the input dimension), with regard to outputs $\mathbf{y} \in \{0, 1\}^L$ for L labels.

$$\hat{\mathbf{y}} = \operatorname{argmax}_{\mathbf{y} \in \{0, 1\}^L} P(\mathbf{y}|\mathbf{x}) \quad (1)$$

III. THE ENVIRONMENT

Describe your environment, addressing the following points:

- State space; What observations does an agent have of the environment
- Action space: what actions can be taken in this environment
- How is the reward function defined
- Is the environment deterministic or stochastic, fully observed, partially observed, not observed, etc.

¹You can make use of a footnote like this one. Code found here: <http://anonymouslinktoyourcode.zip>. Note that results should be reproducible using the technologies from the labs (i.e., Python, and relevant libraries)

- What are the main challenges this environment poses (for an agent)
- Are there potential real-world applications, or is it for theoretical/educational interest?

Don't hesitate to use diagrams, figures, and screenshots wherever they are useful; as exemplified in Fig. 1.

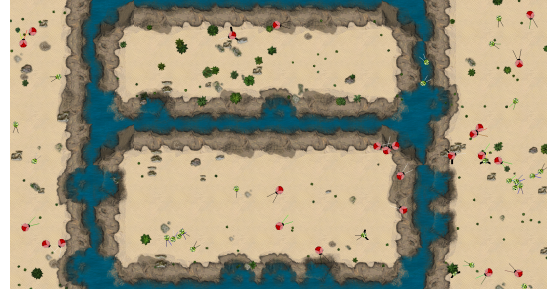


Fig. 1. Figure captions should be descriptive (not like this one).

IV. THE AGENT

Describe your agent, answering the following points:

- What type(s) of agent(s) did you select/design for this environment,
- Why this selection/design?
- What are the main advantages/disadvantages.
- How did you implement/configure/parametrize it.

V. RESULTS AND DISCUSSION

This is one of the most important sections.

- 1) Test your agent(s) in the environment(s),
- 2) show the results, – and most importantly –
- 3) discuss and *interpret* the results (don't just explain the results, but say what they mean); this includes highlighting strong points and also weaknesses.

Make use of plots, e.g., Fig. 2, tables (e.g., Table I), etc; anything that illustrates the performance of your agent in the environment under different configurations. Make sure to clearly indicate which parameters (γ , etc.) you have set.

TABLE I
TABLE CAPTIONS SHOULD ADEQUATELY DESCRIBE THE CONTENTS OF TABLES (UNLIKE THIS ONE).

Environment config.	SARSA	Q-Learning
Simulation 1	10	15
Simulation 2	12	11

VI. CONCLUSION AND FUTURE WORK

Summarize the project: Main outcome, discoveries, lessons learned, possible future work if you had had more time, etc.

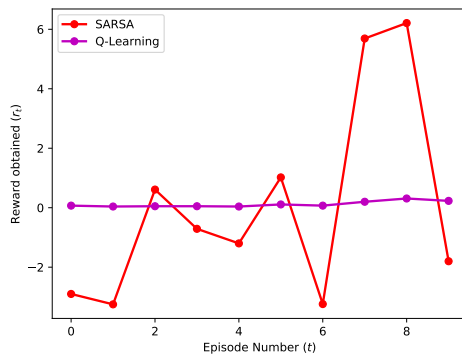


Fig. 2. An example plot. Your plots should be understandable from labels and the caption.

APPENDIX

Some other interesting things you tried that aren't essential to the main outcome? (For example, additional results and tables, lengthy derivations, ...). You can include them here (does not count towards page limit).

REFERENCES

- [1] D. Barber. Bayesian Reasoning and Machine Learning, *Cambridge University Press*, 2012.
- [2] As mentioned in Lecture III - Multi-Output Learning. *INF581 Advanced Topics in Artificial Intelligence*, 2020.
- [3] D. Mena et al. A family of admissible heuristics for A* to perform inference in probabilistic classifier chains. *Machine Learning*, vol. 106, no. 1, pp 143-169, 2017.
- [4] O. Vinyals et al. StarCraft II: A New Challenge for Reinforcement Learning. <https://arxiv.org/abs/1708.04782>, 2017.