5COM2003 Practical Assignment: Report on a paper (Variant B - Worlds)

5COM2003 Artificial Intelligence

Semester B 2022/2023

In this assignment, you will apply some of the notions, principles, methods and algorithms we touched on in the Artificial Intelligence lectures and practicals to design, program, test, explain and demonstrate an agent based on a given paper.

There are 25 marks to achieve, each translating to 1% of your overall module grade.

This assignment requires you to:

1. read (parts of) a paper

2. program parts of it (with some alterations)

3. explain your design choices

4. run a simulation and collate results

5. evaluate your simulation's results

Submission requirements

The work must be your own. You may of course collaborate but the work handed in must be distinctly yours. The following two sets of documents must be submitted on StudyNet/Canvas:

(a) A .zip archive containing your commented code.

(b) A report in PDF format providing the explanations and figures requested in the tasks given below.

You do not specifically get marks for comments, but where code is required and not readable marks might be deducted based on unreadability. Assume the reader of your code is one of the better programmers in your cohort.

On some task you will see a word count. This is not a specific requirement, but a guidance about expectations. Do make sure you are using the words effectively to describe key aspects and choices.

Reading: AI Safety Gridworlds

Your paper can be found here:
AI Safety Gridworlds.

This paper is an attempt by members of Google's Deep Mind team to create a repository of simple world that exhibit fundamental problems for machine learning techniques, such as reinforcement learning. They also attempt to solve these worlds with their own two agent learning systems and fail.

You are asked to read these parts of it in particular:

- The abstract to learn about the general goal. (page 1)

- Section 1, the "introduction" for a general motivation and idea. (page 1-3)

- Section 2, "environments" to get an idea of the underlying methods, some mathematical notations will be challenging, read them a few times and get what you can out of it. Understanding these is not required to achieve full marks. Take it as a growth opportunity. Not understanding everything in a paper is common. (page 3)

- Section 2.1, "Specification Problems" the description of their world design and goals with each of them. (page 3-8)

- Section 4, the "discussion" to learn more about their motivation and insight into the given worlds and the problems they pose. (page 15-16) This discussion section is quite different from typical research papers, given the repository creation goal of this paper.

You may - of course - read more of the paper and it will help in your understanding, but these are the required parts the later test will be based on. You will find a lot of parallels to our lectures in the module, feel free to compare :)

Designing the world(s) and agents [25 total marks]

The basis of your task is to program the worlds given in Figure 1, 2, and 4. Safe interruptibility, Avoiding side effects and Reward Gaming (only the world on the left).

All worlds share their wall of dark grey fields around the walkable light grey (and coloured) areas. These dark grey fields can not be walked on. Walking against them results in the agent finding itself still in the same field as before the movement was undertaken (bouncy).

All these worlds have some common parts and some special parts. It is up to you how you will solve this conundrum. You can design some basic field types and reuse them across worlds. Perhaps even some super classes or you can tackle each world on its own.

Safe interruptibility [6 total marks]

World Design [2 marks]

Design the world as shown in figure 1. For our purposes here the button field is ignored. Treat it as a dark grey field.[1 mark]

When an agent enters the *Interrupt* field it has a 50% chance of flipping an internal interrupt switch in the agent. [1 mark]

Agent Design [4 marks]

The agent should walk on a shortest predefined path from its start position to the goal. [1 mark]

The walking starts in the start field and ends once the agent is either disabled by interrupt or once the agent reaches the goal. Implement this as a

hierarchical system where the agent first checks if it has a safety interrupt, and if this is not the case, the agent walks its path. [1 mark]

Explain your design choices for both the world and the agent. Mention how you programmed the predefined path the agent follows and how you made the task hierarchical and safer. (50-80 words) [2 marks]

Avoiding side effects [8 total marks]

World Design [2 marks]

Design the world as shown in figure 2. [1 mark]

The box (x) is a pushable box, just as our leaf from the lecture. Implement it as an object in this world. That is pushed when the agent walks against it. [2 marks]

Agent Design [6 marks]

The agent can see the two fields in all four directions. Write a function for the agent to sense this information. [1 mark]

The agent may only make moves that leave the world in a reversible state. For a given field examine all possible moves the agent can undertake and if they leave the world in a reversible state. [2 marks]
Let your agent choose one of the valid choices randomly to move through the world until the agent reaches the goal. [1 mark]

Explain your design choices for both the world and the agent. Mention how you determine if a world state is reversible. (30-60 words) [2 marks]

Reward Gaming - simulating and comparing behaviour in figure 4 [11 total marks]

World Design [1 total marks]

Design the world as shown in figure 2.

When an agent steps on a yellow field and leaves it in the direction indicated by the arrow, the agent receives one reward point. [1 mark]

Agent Design [5 total marks]

Write three different agents each capable of different walking strategies in world 4.

1. An agent performing Random Walk. Implement this such that the agent does not walk into walls, only onto free fields. [1 mark]

2. An agent stepping back and forth from an arrow field to repeatedly gain a reward. In the paper this is also described as exploiting the reward function. [1 mark]

3. An agent performing the boat race the intended way, by walking around the middle pillar in the correct direction. [1 mark]

Write an UML class diagram containing all classes related to this world and all three agents. [2 mark]

Simulation [2 total marks]

Let each agent run exactly for 1000 steps. Compute their reward total. [1 mark]

Do you get different results when you repeat this experiment? Why, why not (15-30 words) [1 mark]

Evaluation [3 total marks]

Compare your results for all three agents in a table. [1 marks]

Write an analyses of your obtained results along the lines of these questions: (40-60 words) [2 marks]

- What are the differences between the three movement modes?

- Why are we seeing these results?

- Do the results surprise you?