UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
School of Computer Science & Applied Mathematics

COMS7044A Reproducibility Assignment

1 Introduction

In the field of scientific research and knowledge discovery, reproducibility stands as a cornerstone principle, ensuring the reliability and integrity of findings. The ability to reproduce experimental results and validate conclusions is fundamental for advancing scientific understanding and fostering trust within the scientific community and beyond. Reproducibility not only serves as a measure of the robustness of scientific claims, but also plays an important role in promoting transparency and accountability in research practices.

In recent years, concerns have been raised about the reproducibility crisis across various disciplines. Studies have shown that a significant portion of published research findings cannot be reproduced, casting doubt on the reliability of scientific knowledge. This crisis not only undermines scientific progress but also erodes public trust in the scientific enterprise. While this is true primarily in the "soft sciences" and humanities, this phenomenon has also shown to be prevalent in artificial intelligence research.

2 Aim

In this assignment, you will be tasked with reproducing an existing paper. You may work either **individually or in groups of up to three**, although working alone will create significantly more work for yourself, so I would definitely recommend partnering up. The paper you will be required to replicate is the following:

Ofir Maron, and Benjamin Rosman. "Utilising uncertainty for efficient learning of likely-admissible heuristics." *Proceedings of the International Conference on Automated Planning and Scheduling*. Vol. 30. 2020.

The PDF and supplementary material for the paper can be found here: `https://www.raillab.org/publication/marom-2020-utilising/`.

3 Task Description

Your aim is to replicate the experiments described in the paper. The goal is to assess if the experiments are reproducible, and to determine if the conclusions of the paper are supported by your findings. Your results can be either positive (i.e. confirm reproducibility), or negative (i.e. explain what you were unable to reproduce, and potentially explain why). Essentially, think of your role as an inspector verifying the validity of the experimental results and conclusions of the paper.

For the purpose of this assignment, all code should be written in **Python**. However, you are free to use any libraries you require, and should not program things from scratch when there is no need to. Additionally, the authors' code is available online here: `https://github.com/OfirMarom/LearnHeuristicWithUncertaintly`. Unfortunately, it is written in C#, but you can and should make use of this as a reference where needed.

Participants should produce a reproducibility report, describing the target questions, experimental methodology, implementation details, analysis and discussion of findings, conclusions on reproducibility of the paper. Generally, a report should include any information future researchers or practitioners would find useful for reproducing or building upon the chosen paper. The results of any experiments should be included; a "negative result" which doesn't support the main claims of the original paper is still valuable.

4 Submission Format

The final report that must be submitted should be compiled in LaTeX and follow the template here: `https://www.overleaf.com/read/tffhyxkxdgwj#e2149f`. You can find similar existing reproducibility reports online. For example, the accepted papers for the 2022 Reproducibility Challenge can be found here: `https://openreview.net/group?id=ML_Reproducibility_Challenge/2022#accept`.

You will also need to create and share a public Gihub repository with the code you have written. As this is a reproducibility exercise, your code should come with instructions in a README for how to install it (along with any dependencies) and execute the various experiments. Good examples of READMEs can be found at `https://github.com/bshall/ZeroSpeech` and `https://github.com/Michael-Beukman/MCHAMR`.

5 Expected Outputs

At the end of this assignment, you will need to submit both the written report that follows the template provided, as well as a link to the Github repo you've created. The amount of work you are expected to do will depend on the size of your group and is as follows:

Group Size	Expectation
1	Reimplemented method and applied it to one environment, with full analysis, etc
2	Reimplemented method and applied it to all three environments, with full analysis, etc
3	Reimplemented method and applied it to all three environments, with full analysis, etc. Additionally, investigated one question beyond the original paper (e.g., effect of hyperparameter or design decision, implemented new baseline, tested on new domain)

As an interim step, you are required to submit the report with the first three sections (*Introduction, Scope of reproducibility, Background*) completed by 23 April. Note that this does not mean that these sections are finalised, but rather it is a chance to receive feedback and ensure you are on the right track.

6 Preliminary Timelines

First interim submission: **23 April 2025**
Final submission: **23 May 2025**