

# Προτεινόμενα Θέματα Διπλωματικής Εργασίας

Από

Καθηγητή Γεώργιο Βούρο

## A. Goal Conditioned Imitation Learning using Generative models

{excerpt from [2]} Goal-Conditioned Imitation Learning is a subdomain of Imitation Learning [1], where each demonstration is augmented with one or more goal-states that are indicative of the task that the demonstration was provided for. The goal-state contains information that a learning method can leverage to disambiguate demonstrations. Consequently, a goal-conditioned policy, i.e., a policy that includes the goal-state in its condition set, can use a given goal-state to adapt its behavior accordingly.

The goal of this MSc thesis is to study state of the art methods for Goal-Conditioned Imitation Learning using advanced generative models, such as diffusion models, propose advances, and benchmark these in advanced problems. The methods to be studied include BESO[2], Diffusion Policy [3] and Diffusion-BC [4].

### References

1. B.D.Argall et al., “A survey of robot learning from demonstration”,
2. M.Reuss et al., “Goal-Conditioned Imitation Learning using Score-based Diffusion Policies”, 2023
3. C. Cheng, et al., “Diffusion policy: Visuomotor policy learning via action diffusion”, 2023.
4. T. Pearce, et al., “Imitating human behaviour with diffusion models”, 2023.

## B. Modelling humans for human-AI collaboration.

This thesis aims to investigate answering the question how can we build reinforcement learning algorithms that optimize jointly own utility, reflecting their objective, as well the utility of humans? This need is emergent in safety-critical systems, where human need to trust the agent and the agent needs to learn human preferences and needs and align with them.

There are different ways to try to answer this question. The thesis will investigate alternatives based on reinforcement learning with human feedback (e.g. [1]), models calibration (e.g. [2]) and conformal prediction (e.g. [3]).

### References

- [1] D.Ziegler et al., “Fine-Tuning Language Models from Human Preferences”, 2019
- [2] K Vodrahalli et al., “Uncalibrated Models Can Improve Human-AI Collaboration”, 2022
- [3] A. Angelopoulos and S.Bates, A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification, 2021.

### C. Revisiting Behavioral Cloning.

This thesis aims to revisit the behavioural cloning technique taking advantage of latest achievement in machine learning and prediction, addressing distribution shift, compounding errors, causal confusion and need of abundance of data.

#### References

- [1] Ross et al 2011, A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning
- [2] De Haan et al, 2019, [Causal confusion in imitation learning](#)
- [3] Shah et al, 2023, GNM : A general Navigation model to drive any robot