PhD projects (Robotics)

- World Models and Reinforcement Learning for Reliable Robot Autonomy
- Explaining robotic planning decision points along execution
- Civic and Responsible AI

World Models and Reinforcement Learning for Reliable Robot Autonomy

Supervisor: Matteo Leonetti

Project Description

Robots can perform individual skills such as grasping, placing and opening, but remain unreliable when a task requires many interdependent actions in a changing, partially observed environment. This project will investigate how learned predictive models of the physical world can support deliberate, long-horizon robot planning.

The central question is how a robot can use experience to represent the objects, relations and action consequences that matter for a task, then use this representation to select promising actions, test whether they are feasible, and revise its behaviour when execution fails. The project may address object- or relation-centred world models; learning action effects, affordances or causal structure; integrating task-level reasoning with motion planning; language or vision-language models as sources of semantic guidance; uncertainty estimation, verification and recovery; or transfer from simulation to physical robots.

The successful candidate will define a focused research question within this area and develop a rigorous methodology combining machine learning and robotics. Proposals should identify a current limitation in reliable long-horizon autonomy, explain the relevant literature, formulate testable hypotheses, and specify evaluation in simulation, on physical hardware, or both.

References

Ferraro S, Mazzaglia P, Verbelen T and Dhoedt B (2025) FOCUS: object-centric world models for robotic manipulation. Front. Neurorobot.
Z. Yang, C. Garrett, D. Fox, T. Lozano-Perez and L. P. Kaelbling, "Guiding Long-Horizon Task and Motion Planning with Vision Language Models," 2025 IEEE International Conference on Robotics and Automation (ICRA), 2025.
Q. Weiming, W. Yaoyao and L. Dingsheng, "State Representation Learning for Task and Motion Planning in Robot Manipulation," 2023 IEEE International Conference on Development and Learning (ICDL), 2023.
Mendez-Mendez, J., Kaelbling, L.P. & Lozano-Perez, T.. (2023). Embodied Lifelong Learning for Task and Motion Planning. Proceedings of The 7th Conference on Robot Learning.
Ichter, B., et al. (2023). Do As I Can, Not As I Say: Grounding Language in Robotic Affordances. Proceedings of The 6th Conference on Robot Learning.

Explaining robotic planning decision points along execution

Supervisor: Gerard Canal

Project Description

Explanation of robotic behaviours has been proved to be very important to improve the understanding of the users of such robots, which improves their trust in the robotic system.

However, explanations in robotics are tricky as they need to be given at the correct moment and based on what happened in the execution. In robotic-based planning, an interesting explanation is that of decision points, where the robot could have taken a different action with a different outcome.

This project focuses on the explanation of such decision points at execution time, integrating information om the current events and past events that may help explain the decision to a user. For this, we will look into explainability in the space of plans where, knowing the committed plan and what has happened in the execution, we compare with the other alternatives that the robot had at a certain decision point. This will evolve towards generating explanations along the execution of plans, as well as determining when some decisions may not be obvious to the user, thus warranting explanations.

References

[1] Canal, G., Torras, C., & Alenya, G. (2023). Generating predicate suggestions based on the space of plans: an example of planning with preferences. User modeling and user-adapted interaction, 33(2), 333-357.
[2] Eifler, R., Cashmore, M., Hoffmann, J., Magazzeni, D., & Steinmetz, M. (2020, April). A new approach to plan-space explanation: Analyzing plan-property dependencies in oversubscription planning. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 06, pp. 9818-9826).
[3] Wachowiak, L., Celiktutan, O., Coles, A., & Canal, G. (2023, June). A Survey of Evaluation Methods and Metrics for Explanations in Human—Robot Interaction (HRI). In ICRA2023 Workshop on Explainable Robotics.
[4] Wachowiak, L., Tisnikar, P., Canal, G., Coles, A., Leonetti, M., & Celiktutan, O. (2022, August). Analysing eye gaze patterns during confusion and errors in human—agent collaborations. In 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) (pp. 224-229). IEEE.

Civic and Responsible AI

Supervisor: Martim Brandao

Project Description

This project is open to a variety of subjects falling under the themes of "Civic AI" and "Responsible Robotics", except those falling under Education, Security or Military applications.

The topic should be one the student is passionate about and has already had some exposure to, but is quite free as long as aligned with the interests of the CRAIL lab (www.civicandresponsibleai.com). Interdisciplinary backgrounds are welcome.

Examples of Civic AI projects:
- New system designs and techniques that apply AI/robotics to solve problems of corporate accountability and human rights organisations
- Leveraging public data and collaborative citizen infrastructure for oversight of AI and robot deployments

Examples of Responsible Robotics projects:
- Algorithms for fair, safe, explainable or accountable robots
- Ergonomics of work with robots and worker-centred robotics and HRI
- Algorithms for sustainable algorithms
- Algorithms for robot accident investigation