Ethical and proactive decision-making in autonomous robots

Lead supervisor: Dr Samuele Vinanzi

Co-supervisor(s): Prof Alessandro Di Nuovo

Research Centre or Department: Department of Computing

Contact for applicant queries: Dr Samuele Vinanzi s.vinanzi@shu.ac.uk

Project summary

This summer internship aims to investigate ethical decision-making in autonomous robots. We are seeking a talented and motivated individual looking to build research experience by collaborating with the Smart Interactive Technology (SIT) group in developing the next generation of cognitive robots. We are particularly interested in developing agents capable of incorporating ethical reasoning into their decision-making processes.

The prevailing assumption regarding artificial intelligence (AI) is that an agent will always adhere to the commands it receives. However, in the context of developing intelligent agents for human-robot collaboration, this compliance-based strategy is suboptimal: while robots may experience failures, the same holds true for humans. In such scenarios, a robot needs to possess cognitive mechanisms to identify and anticipate mistakes or deceptions in its human partner's strategy and make appropriate decisions to ensure the successful completion of the shared task. For instance, an autonomous vehicle may disobey its human operator's command to drive through a red traffic light.

Your role will involve building a system that possesses a set of ethical values and can reason about the consequences of its actions to select the most appropriate response to a given situation. You will be engaged in the development of a Belief-Desire-Intention (BDI) agent, utilizing existing frameworks such as BDIPython or GWENDOLEN. The expected deliverable will be a foundational system that can be used as a building block for further research on the topic.

This internship offers a unique opportunity to contribute to the growing field of robotics and AI ethics while gaining practical experience in agent-based systems and ethical decision-making frameworks. You will have the opportunity to develop skills in Python programming, robotics and ethical artificial intelligence.

Specific skills and experience required for this project

Please also refer to the advert on our jobs pages for the person specification for these internships

The ideal candidate for this internship should possess a combination of technical skills, research aptitude, and a strong interest in the topic under consideration.

Essential skills:

- Strong proficiency in Python programming.
- Experience or coursework in machine learning and/or artificial intelligence.

Desirable skills:

- Familiarity with concepts and frameworks related to robotics and autonomous systems.
- Strong problem-solving abilities and analytical skills.
- Excellent English communication skills, both written and verbal.
- Ability to work independently as well as in a team environment.
- A proactive and motivated approach to learning and tackling new challenges.

Project location

City Campus

Home working may be available

Project delivery

The project is suitable both for full-time and part-time delivery.