MASTER GRADUATION PROJECT

Moral uncertainty for autonomous systems

The AiTech initiative together with the Interactive Intelligence group (Department of Intelligent Systems, EEMCS) and the Ethics/Philosophy of Technology group (Department of Values, Technology and Innovation, TPM), has an opening for an MSc-student looking for a master graduation project.

**Title:** Ethical decision making for autonomous systems considering moral uncertainty

**PROJECT DESCRIPTION**

As a result of rapid and impressive technological developments, possible applications for autonomous intelligent systems becomes broader. Such systems will have huge societal impact, for bad or for good, and will inherently make actions with ethical implications. If we want these systems to be beneficial, their actions must be aligned with our moral principles. Nevertheless, in many cases, there is no consensus on what is the “right” thing to do and according to which principles should the AI act.

Contributing to the field of beneficial AI, this master graduation project focuses on the understanding and modeling of decision-making mechanisms to incorporate normative uncertainty (uncertainty on which ethical theory is the most appropriate) for autonomous systems. Specific tasks to achieve the overall goal are:

- Studies on normative ethics and computational modeling of ethical theories: Understand the implications and limitations of normative ethics to AI, specifically focusing on the field of autonomous driving systems;
- Application of machine learning techniques on a large-scale dataset (e.g. the MIT’s moral machine dilemma dataset) to identify how human decisions relate to different ethical theories;
- Modeling of a decision-making mechanism rooted in ethical theories to incorporate moral uncertainty on the field of autonomous vehicles.

**PRACTICAL DETAILS**

- Starting date: As soon as possible
- Location: TU Delft, Interactive Intelligence Group
- Daily supervisor: Luciano Cavalcante Siebert
- Cosupervisors: Prof. Catholijn Jonker; Prof. Jeroen Van de Hoven

**HOW TO APPLY**

If you are interested in this position, please send your CV and a brief motivation to L.CavalcanteSiebert@tudelft.nl