Moral Permissibility of Actions in Smart Home Systems

In the near future, we will see the realization of smart homes, homes where appliances etc. are wholly or partially controlled via artificial intelligence. In such homes, many everyday decisions will have to be made by artificial agents, and these decisions and plans must be ethically acceptable. With this poster, we present ongoing work of how to operate a smart home via a Hybrid Ethical Reasoning Agent (HERA), see (Lindner, Bentzen, and Nebel 2017). This work is part of the broader scientific effort to implement ethics on computer systems known as machine ethics, see also (Dennis, Fisher, Slavkovik, and Webster, 2016; Lindner and Bentzen, 2018). We showcase an everyday example involving a mother and a child living in the smart home. Our formal theory and implementation allows us to evaluate actions proposed by the smart home from different ethical points of view, i.e. utilitarianism, Kantian ethics and the principle of double effect. When points of view differ, ethical uncertainty ensues, and this is the case in the showcased example. We suggest various ways of coping with the ensuing ethical uncertainty, e.g. human in the loop, one overriding ethics. We discuss how formal verification, in the form of model-checking, can be used to check that the modeling of a problem for reasoning by HERA conforms to our intuitions about ethical action.