Value-Oriented Ethico-Legal Reasoning in LogiKEy

Christoph Benzmüller
Freie Universität Berlin

David Fuenmayor
University of Luxembourg

September 6, 2021

There exists critical application areas in which a naive deployment of modern AI technology could cause significant damage or harm. Simple pleas towards "transparency" and "explanation", due to their "posthoc nature", are not offering convincing solutions in such contexts, since they are hardly suited to prevent disasters in the first place; they could even be seen as invitations for adversial attacks. It is, however, of societal interest to invest in preventive measures.

We argue for the development of ethico-legal governors to assess, justify and legitimate options for critical actions of an intelligent artificial agent before action execution is granted. Such governor technology calls for the provision of deliberative legal and moral reasoning competencies in intelligent artificial agents.

Our research therefore focuses on the provision of flexible and expressive symbolic means to represent and reason with normative theories. To address this challenge we have developed the LogiKEy formal framework, methodology, and associated tool support [5, 1]. LogiKEy supports the design and engineering of ethical reasoners, normative theories and deontic logics in a most flexible way, and it also provides a fruitful link between different research communities, including knowledge representation and reasoning in AI, the deduction systems community and formal ethics. In particular, LogiKEy enables the application of off-the-shelf interactive [6] and automated [7] theorem proving technology for classical higher-order logic in ethico-legal reasoning.

In more general terms our goal is to work towards some form of legally and ethically reasonable machines [4], capable of exchanging rational justifications for the actions they take. Building up a capacity in such reasonable machines to engage in *value-oriented ethico-legal argumentation* is thus a relevant challenge to address. As a possible solution we currently explore suitable adaptions of the multilayered LogiKEy framework to not only enable flexible and expressing non-classical reasoning, but to also take different value systems and preferences into account [2, 3].

At the PhDs in Logic event we will offer a demo of our recent achievements. In particular, we will demonstrate how LogiKEy supports the exploration and modelling of value ontologies and value preference systems in combination with

the exploration and application of combinations of expressive non-classical logics. As a prominent, exemplary use case we will demonstrate the formal reconstruction of prominent legal cases in property law (wild animal court cases).

References

[1] Christoph Benzmüller, Ali Farjami, David Fuenmayor, Paul Meder, Xavier Parent, Alexander Steen, Leendert van der Torre, and Valeria Zahoransky. LogiKEy workbench: Deontic logics, logic combinations and expressive ethical and legal reasoning (Isabelle/HOL dataset). *Data in Brief*, 33(106409):1–10, 2020.

[2] Christoph Benzmüller and David Fuenmayor. Value-oriented legal argumentation in Isabelle/HOL, 2021. Preprint: `https://dx.doi.org/10.13140/RG.2.2.21514.80320`.

[3] Christoph Benzmüller, David Fuenmayor, and Bertram Lomfeld. Encoding legal balancing: Automating an abstract ethico-legal value ontology in preference logic. In *Workshop on Models of Legal Reasoning, hosted by 17th Conference on Principles of Knowledge Representation and Reasoning (KR 2020)*, 2020. Preprint: `https://arxiv.org/abs/2006.12789`.

[4] Christoph Benzmüller and Bertram Lomfeld. Reasonable machines: A research manifesto. In Ute Schmid, Franziska Klügl, and Diedrich Wolter, editors, *KI 2020: Advances in Artificial Intelligence – 43rd German Conference on Artificial Intelligence, Bamberg, Germany, September 21–25, 2020, Proceedings*, volume 12352 of *Lecture Notes in Artificial Intelligence*, pages 251–258. Springer, Cham, 2020. Preprint: `https://dx.doi.org/10.13140/RG.2.2.28918.63045`.

[5] Christoph Benzmüller, Xavier Parent, and Leendert van der Torre. Designing normative theories for ethical and legal reasoning: LogiKEy framework, methodology, and tool support. *Artificial Intelligence*, 287:103348, 2020. Preprint: `https://www.researchgate.net/publication/342146653`.

[6] Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. *Isabelle/HOL - A Proof Assistant for Higher-Order Logic*, volume 2283 of *Lecture Notes in Computer Science*. Springer, 2002.

[7] Alexander Steen and Christoph Benzmüller. Extensional higher-order paramodulation in Leo-III. *Journal of Automated Reasoning*, 2021. Accepted for publication, preprint: `https://arxiv.org/abs/1907.11501`.