

Department of Computer Science

Towards Computational Morality



Ari Saptawijaya

PhD Student in Computer Science
Webpage: <http://goo.gl/kJ5mV>

- M.Sc. in Computational Logic (TU Dresden, Germany)
- B.Sc. in Computer Science (Universitas Indonesia, Indonesia)



Centro de Inteligência Artificial (CENTRIA)
Advisor: Prof. Luís Moniz Pereira
Starting: October 2011
Foreseen Conclusion: August 2014

CENTRIA

Objectives

- To develop a Logic Programming based system with features needed in modeling moral settings, to represent agents' knowledge in morality settings, and to allow moral reasoning under some aspects studied in moral philosophy, moral psychology, and other related fields.
- To employ the developed system for modeling several moral situations and reasoning thereof, under some aspects studied in morality-related fields. Whenever empirical results from those fields are available, we shall verify, whether the computational results are corroborated by the empirical ones.

Methodology

1. **Design the System:** identify morality aspects from literature; study logic programming based formalisms allowing knowledge updates and abduction; adapt the chosen formalisms and extend it with the features for moral reasoning.
2. **Implement the System:** implement each feature needed in the system using XSB Prolog; test the implementation and assess the implementation by recapitulating its strengths and weaknesses.
3. **Modeling Morality Aspects:** model each aspect that is identified from the literature; assess the results and verify with those from the moral philosophy and psychology fields.

Expected Results

1. The first step of the methodology results in an architecture of the system for moral reasoning under the considered morality aspects.
2. The second step of the methodology results in a working and tested implementation, following the architecture and based on the specification, ready for modeling moral situations and moral reasoning under the considered morality aspects.
3. The third step of the methodology results in several models of moral situations, with respect to the aspects considered, and analysis of their results (comparison with empirical and theoretically prescribed examples and results gleaned from the literature).

Funding: Bolsa de Doutoramento – FCT/MEC (SFRH/BD/72795/2010)



Principles of Double & Triple Effect

