

Accepted Manuscript

Computational Models of Ethical Decision-Making: A Coherence-Driven Reflective Equilibrium Model

Levent Yilmaz, Ana Franco-Watkins, Timothy S. Kroecker

PII: S1389-0417(16)30121-8

DOI: <http://dx.doi.org/10.1016/j.cogsys.2017.02.005>

Reference: COGSYS 544

To appear in: *Cognitive Systems Research*

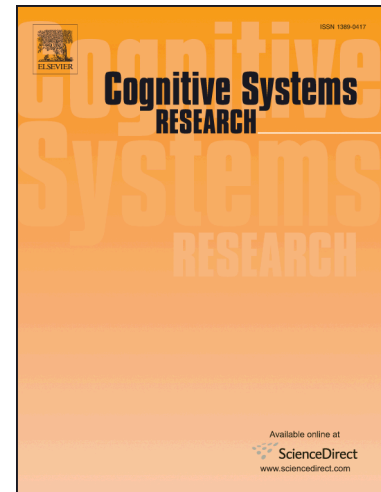
Received Date: 6 July 2016

Revised Date: 8 December 2016

Accepted Date: 9 February 2017

Please cite this article as: Yilmaz, L., Franco-Watkins, A., Kroecker, T.S., Computational Models of Ethical Decision-Making: A Coherence-Driven Reflective Equilibrium Model, *Cognitive Systems Research* (2017), doi: <http://dx.doi.org/10.1016/j.cogsys.2017.02.005>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Computational Models of Ethical Decision-Making: A Coherence-Driven Reflective Equilibrium Model

Levent Yilmaz

*Department of Computer Science and Software Engineering
Auburn University*

Ana Franco-Watkins

*Department of Psychology
Auburn University*

Timothy S. Kroecker

Air Force Research Lab

Abstract

There are scientific and technical challenges that must be addressed in developing systems that interact with humans and work along with other agents in complex, dynamic, and uncertain environments where ethical concerns may arise. In such systems relationships between users and autonomous components will be driven as much by issues such as trust, responsibility, and acceptability, as technical ones such as planning and coordination. This paper provides a comprehensive review and classification of existing methods in machine ethics, resulting in delineation of specific challenges and issues. To address the identified challenges, we introduce a method that leverages the method of reflective equilibrium and the multi-coherence theory as a unifying constraint satisfaction framework to simultaneously assess multiple ethical principles and manage ethical conflicts in a context-sensitive manner.

Keywords: decision-making, machine ethics, cognitive coherence, reflective equilibrium, cognitive agent

Download English Version:

<https://daneshyari.com/en/article/4942377>

Download Persian Version:

<https://daneshyari.com/article/4942377>

[Daneshyari.com](https://daneshyari.com)