

Engineering Kindness: Building a Machine with Compassionate Intelligence

Cindy Mason, University of California, Berkeley, Berkeley, CA, USA

ABSTRACT

The author provides first steps toward building a software agent/robot with compassionate intelligence. She approaches this goal with an example software agent, EM-2. She also gives a generalized software requirements guide for anyone wishing to pursue other means of building compassionate intelligence into an AI system. The purpose of EM-2 is not to build an agent with a state of mind that mimics empathy or consciousness, but rather to create practical applications of AI systems with knowledge and reasoning methods that positively take into account the feelings and state of self and others during decision making, action, or problem solving. To program EM-2 the author re-purposes code and architectural ideas from collaborative multi-agent systems and affective common sense reasoning with new concepts and philosophies from the human arts and sciences relating to compassion. EM-2 has predicates and an agent architecture based on a meta-cognition mental process that was used on India's worst prisoners to cultivate compassion for others, Vipassana or mindfulness. She describes and presents code snippets for common sense based affective inference and the I-TMS, an Irrational Truth Maintenance System, that maintains consistency in agent memory as feelings change over time, and provides a machine theoretic description of the consistency issues of combining affect and logic. The author summarizes the growing body of new biological, cognitive and immune discoveries about compassion and the consequences of these discoveries for programmers working with human-level AI and hybrid human-robot systems.

Keywords: Affective Inference, AI, Biology, Cognition, Compassion, Compassionate Intelligence, Deep Learning, EQ, Emotion, Heidigger, Human-Level AI, Human-Robot Interaction, Human-Robot Interface, Machines, Meditation, Meta-Cognition, Neuroplasticity, Robots, Semantic Network, Semantic Web, Software Agents, Software Requirements, Time, Truth Maintenance Systems, User Experience

INTRODUCTION

How do you tell the difference between a human and an emotionally intelligent robot? No matter how many times you tell your truly sad story, the robot will cry with you every time. Dr. Ed Weiss

Repeated interactions with the artifacts we create have a rub-off effect. They are changing us. Recent co-discoveries across a number of different fields support the idea that positive emotion has many good side effects so at this juncture in AI, it is time to consider the idea of design-

DOI: 10.4018/IJSE.2015010101

ing compassionate intelligence (CI). The paper outlines a generalized software requirements description and describes the EM-2, a software agent with CI. The work represents first steps towards building machines (robots and software agents) that have a stake in us, by programming with compassionate intelligence based on both state of mind and state of heart. We offer the generalized software requirements description for anyone who wants to pursue this direction of programming in AI/Robotics. This direction is significant not just for AI but for user interfaces, healthcare, education, and design in other fields. We illustrate the CI ideas with code snippets and architectures of software agent, EM-2.

What does it mean to give an AI system compassionate intelligence or have a stake in us? Simply, we mean the ability to program an AI system that makes decisions and actions that take into account positive regard for others and self. We describe software agent EM-2 as a first step to an AI program that represents and reasons not only with logical concepts of mental state but state of the heart. EM-2 components leverage the sensory analysis and multi-agent functions of prior agent systems, however in the paper we focus on elements of EM-2 related to CI. At the programming level this includes multi-level representations and predicates of feelings of self and others as well as logical concepts and the logical concepts about feelings. The work presented here covers several AI concepts and software agent systems: cooperative multi-agent systems technology, emotion oriented programming, common sense knowledge, affective inference, default reasoning and belief revision. We do not presume the agent to “have” feelings, nor do we address this issue here. Rather we address the computational aspects of creating a reasoning apparatus that uses representation of these concepts to accomplish compassionate decision making. Essential to the system is a pro-social agential stance – this includes but is not limited to a) agents do not lie about concepts or feelings and there is common sense knowledge of positive emotion, society and culture.

The motivation for this work is based on a growing body of recent discoveries from social and cognitive neuroscience, psychoneuroimmunology, and genetics indicating humans benefit greatly from compassionate experiences. User experience studies, neuroplasticity and genetic plasticity studies indicate we are literally changed by repeated interactions with objects and relations in our environment. With our growing symbiotic relationship with gadgets (phones, cars, robotic assistants, browsers, appliances, IoT, etc.) there is significant biological imperative to intentionally design for humane and pro-social AI systems and interfaces. It's not to say all AI needs such features, but that we have a choice when it makes sense to do so. The technical components of the paper are divided into three parts. In Part I we present a generalized software requirements guide for building AI systems with compassionate intelligence and then detail the EM-2 implementation of the first three design requirements, namely: a) agent philosophy of mind and architecture that supports it b) a representation and reasoning calculus that supports notions of combining emotion, standard logic, self and other and c) an inferencing and learning component that supports a) and uses b) to make decisions and take action based on consideration of self and other. In Part II we increase our exploration of the issues of combining affect and logical fact components in multiple agents by deepening the descriptions of algorithms and code snippets given in Part I with a machine theoretic architecture and description of the introspective machines for beliefs and feelings of EM-2, including the predicates for introspection and meta-cognition, conflict resolution and naming. Part III of the paper discusses social and technical issues surrounding this topic such as human-level AI, compassionate intelligence, etc.

Our programming approach to building agents with compassionate intelligence includes 1) multi-agent programming methods for the representing, communicating and reasoning about multiple agents cognitive state(s) during decision making and for being responsive to changes of agents cognitive state over time *with a pro-social stance* 2) organizing models of memory and

21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:

www.igi-global.com/article/engineering-kindness/138576?camid=4v1

This title is available in InfoSci-Journals, InfoSci-Journal Disciplines Computer Science, Security, and Information Technology. Recommend this product to your librarian:

www.igi-global.com/e-resources/library-recommendation/?id=2

Related Content

Beyond the ILS: A New Generation of Library Services Platforms

Marshall Breeding (2013). *Robots in Academic Libraries: Advancements in Library Automation* (pp. 13-36).

www.igi-global.com/chapter/beyond-ils-new-generation-library/76457?camid=4v1a

Stiffness Modeling and Analysis of Passive Four-Bar Parallelogram in Fully Compliant Parallel Positioning Stage

X. Jia, Y. Tian, D. Zhang and J. Liu (2013). *Advanced Engineering and Computational Methodologies for Intelligent Mechatronics and Robotics* (pp. 59-75).

www.igi-global.com/chapter/stiffness-modeling-analysis-passive-four/76439?camid=4v1a

Empathy and Human-Machine Interaction

Florence Gouvrit (2013). *International Journal of Synthetic Emotions* (pp. 8-21).

www.igi-global.com/article/empathy-and-human-machine-interaction/97675?camid=4v1a

Bond Graph Modeling and Computational Control Analysis of a Rigid-Flexible Space Robot in Work Space

Amit Kumar, Pushparaj Mani Pathak and N. Sukavanam (2011). *International Journal of Intelligent Mechatronics and Robotics* (pp. 18-30).

www.igi-global.com/article/bond-graph-modeling-computational-control/58320?camid=4v1a