Conniving with Continuations: Representing Goals in a Domain-Specific Language of Thought

Kartik Chandra
MIT CSAIL

Rebecca Saxe
MIT BCS

Jonathan Ragan-Kelley
MIT CSAIL

Joshua B. Tenenbaum
MIT BCS

Abstract

Wanting composes flexibly with knowing: we can want to know, want to know what someone wants, and so on. This paper offers a goal representation that allows for this type of rich integration with theory-of-mind concepts. We extend the memo programming language (https://github.com/kach/memo), which is specialized for theory-of-mind reasoning, with a new syntactic construct called "wants" that represents goals. To implement "wants," we borrow an idea from programming language theory: the notion of a "continuation," which allows reasoning about possible future states of a program. We show that this construct can be used to represent a variety of interesting goals, and discuss what it can teach us about the concept of desire more broadly.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*