

Mechanism Design with Limited Information

Kate Larson

`klarson@cs.uwaterloo.ca`

David R. Cheriton School of Computer Science

University of Waterloo

200 University Avenue West

Waterloo, Ontario, N2L 3G1

CANADA

Abstract

The central mechanism design problem is to develop incentives for agents to truthfully reveal their preferences over different outcomes, so that the system-wide outcome chosen by the mechanism appropriately reflects these preferences. However, in many settings, agents' do not know their actual preferences a priori. Instead, an agent may need to compute or gather information to determine whether they prefer one possible outcome or another. Due to time constraints or the cost of acquiring information, agents must be deliberative in that they need to carefully decide how to allocate their computational resources when determining their preferences. In this paper we study the problem of designing mechanisms explicitly for deliberative agents. We propose a set of intuitive properties which we argue are desirable in deliberative-agent settings. We show that these properties are mutually incompatible, and that many classic approaches to mechanism design are not robust against undesirable behavior from deliberative agents.