

Perceptions of Compromise: Comparing Consequentialist and Contractualist Accounts

Jared Moore

Stanford University, Stanford, California, United States

Sydney Levine

Allen Institute for Artificial Intelligence, Seattle, Washington, United States

Yejin Choi

Allen Institute for Artificial Intelligence, Seattle, Washington, United States

Abstract

We are constantly faced with the question of how to aggregate preferences, views, perspectives and values. This is a problem for groups attempting to accommodate individuals with differing needs and interests, as will be our focus. The problem of “value aggregation” therefore crops up in myriads of places across the social sciences—in rational decision theory, social choice models, and proposals for systems of democratic voting, for instance. These sub-disciplines have formalized proposals for how to deal with value aggregation, though, remarkably, no research has yet directly compared people’s intuitions of two of the most obvious candidates for aggregation—taking the sum of all the values (the classic “Utilitarian” approach) and the product (a less well-known “contractualist” approach). In this paper, we systematically explore the proposals suggested by each algorithm, focusing on aggregating preferences across groups. We find that both humans and performant LLMs prefer a contractualist approach.