

Computational Philosophy: On Fairness in Automated Decision Making

Suresh Venkatasubramanian

School of Computing, University of Utah, Salt Lake City, USA
suresh@cs.utah.edu

Abstract

As more and more of our lives are taken over by automated decision making systems (whether it be for hiring, college admissions, criminal justice or loans), we have begun to ask whether these systems are making decisions that humans would consider fair, or non-discriminatory. The problem is that notions of fairness, discrimination, transparency and accountability are concepts in society and the law that have no obvious formal analog.

But our algorithms speak the language of mathematics. And so if we want to encode our beliefs into automated decision systems, we must formalize them precisely, while still capturing the natural imprecision and ambiguity in these ideas.

In this talk, I'll survey the new field of fairness, accountability and transparency in computer science. I'll focus on how we formalize these notions, how they connect to traditional notions in theoretical computer science, and even describe some impossibility results that arise from this formalization. I'll conclude with some open questions.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems, J.4 Social and Behavioral Sciences

Keywords and phrases fairness, transparency, accountability, impossibility results

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2017.2

Category Invited Talk



© Suresh Venkatasubramanian;

licensed under Creative Commons License CC-BY

28th International Symposium on Algorithms and Computation (ISAAC 2017).

Editors: Yoshio Okamoto and Takeshi Tokuyama; Article No. 2; pp. 2:1–2:1

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany