Fairness and Equity in Resource Allocation and Decision-Making: An Annotated Reading List

FAIDRA MONACHOU
Stanford University
and
ANA-ANDREEA STOICA
Columbia University

Fairness and equity considerations in the allocation of social goods and the development of algorithmic systems pose new challenges for decision-makers and interesting questions for the EC community. We overview a list of papers that point towards emerging directions in this research area.

Categories and Subject Descriptors: F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity; J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*

General Terms: Algorithms, Machine Learning, Economics, Applied Modeling

Additional Key Words and Phrases: Fairness, Equity, Resource Allocation, Decision-making

Improving fairness and equity of decision-making systems used in the allocation of social goods is an important priority in a wide range of domains. For example, how can a government fairly and equitably allocate scarce medical resources to citizens? What does it mean for a machine learning algorithm used for loan allocation to be fair? How can an online matching platform for freelance workers ensure equal access to employment opportunity for all?

While many different disciplines have studied these problems, we argue that they lie at the very center of the economics and computation field, as algorithms constitute a central mechanism in many traditional decision-making systems. Whether algorithms are in the role of assisting in resource allocation at scale, diagnosing differences in access, or governing new environments in which resources are distributed (e.g. online platforms), fairness questions lie at the core of their design and implementation. Thus, these challenging questions pose a unique opportunity, not only to uncover untapped research insights utilizing tools from mechanism design, algorithms, machine learning, and optimization, but most importantly to collectively contribute potential research-grounded solutions to emerging societal problems.

In an attempt to highlight some of the most promising directions for future research, in this article we offer a list of papers that could serve as a useful starting point for the interested readers in the EC community. We acknowledge that the list below is far from exhaustive and fully representative, since due to space constraints, we omitted many important and closely related papers. Nevertheless, we hope it serves as a useful starting point for the interested readers in the EC community, highlighting some of the most promising directions for future research.

Authors' addresses: monachou@stanford.edu, astoica@cs.columbia.edu

(1) Phelps, E. S. 1972. The statistical theory of racism and sexism. *The American Economic Review 62*, 4, 659–661.

In his seminal paper, Phelps, often credited together with K. Arrow, introduces *statistical discrimination*: the theory that, even in the absence of prejudice, discrimination can arise due to uncertainty about individuals' true characteristics.

(2) Coate, S. and Loury, G. C. 1993. Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review 83*, 5, 1220–1240.

This classic paper introduces discrimination as *coordination failure*, i.e., when groups of ex ante identical agents choose different equilibrium strategies. Its simple equilibrium model is the basis of multiple subsequent works until today.

(3) Kleinberg, J., Mullainathan, S., and Raghavan, M. 2017. Inherent trade-offs in the fair determination of risk scores. In *8th Innovations in Theoretical Computer Science Conference*. Vol. 67. 43:1–43:23

A staple of the fairness research literature, this paper presents a theoretical analysis exploring the trade-offs between three main statistical definitions of group fairness, showing that not all can co-exist when the underlying data is not completely unbiased.

(4) Hardt, M., Price, E., and Srebro, N. 2016. Equality of opportunity in supervised learning. *Advances in Neural Information Processing Systems 29*.

This paper proposes an alternative definition of fairness to demographic parity that shows a better alignment between objectives and diversity considerations. Formalized as *equality of opportunity*, this definition bridges notions of equality and fairness and opens up avenues of research in designing mechanisms that equalize chances of obtaining resources across different groups.

(5) Hu, L. and Chen, Y. 2020. Fair classification and social welfare. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 535–545

This paper explores the connections between fair classification and social welfare maximization, pointing towards how "more fair" classifiers can worsen welfare outcomes for all social groups.

(6) Kilbertus, N., Carulla, M. R., Parascandolo, G., Hardt, M., Janzing, D., and Schölkopf, B. 2017. Avoiding discrimination through causal reasoning. *Advances in Neural Information Processing Systems 30*.

This paper proposes a generalized conceptual method for defining fairness through a *causality* criterion, generalizing from observational methods that define fairness. Through its practical distinction between protected attributes and their proxies, this paper opens up avenues of interdisciplinary research for establishing interventions that remove discrimination through causal pathways.

(7) Bertsimas, D., Farias, V. F., and Trichakis, N. 2012. On the efficiency-fairness trade-off. *Management Science 58*, 12, 2234–2250.

Following upon the concept of *price of fairness* introduced by the same authors, this paper considers fairness-efficiency trade-offs that a decision-maker faces in the allocation of scarce resources under the α*-fairness* scheme.

(8) Manshadi, V., Niazadeh, R., and Rodilitz, S. 2021. Fair dynamic rationing. In *Proceedings of the 22nd ACM Conference on Economics and Computation.* 694–695

This paper considers the problem of allocating goods to sequential arriving agents with varying levels of need in an efficient and equitable manner, contributing to less explored areas such as *dynamic fairness*.

(9) Kasy, M. and Abebe, R. 2021. Fairness, equality, and power in algorithmic decision-making. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.* 576–586.

This paper presents a series of limitations to observational definitions of fairness through the lens of social welfare and theories of justice, and shifts the focus towards hidden concepts of within-group heterogeneity and merit-based inequity.

(10) Barocas, S. and Selbst, A. D. 2016. Big data's disparate impact. *California Law Review 104*, 671.

As EC researchers are puzzled with contradicting notions of fairness and the legal limitations of their proposed technical solutions, this essay offers a law perspective to how data-driven algorithmic techniques can lead to disparities and highlights open questions in the intersection of law and computation.