Unlearning the bias: An agent-based simulation for increasing diverse representation through leadership emergence

Andria Smith
Max Planck Institute for Intelligent Systems, Stuttgart, Germany

Simon Heuschkel
University of Tübingen, Tübingen, Germany

Ksenia Keplinger
Max Planck Institute for Intelligent Systems, Stuttgart, Germany

Charley Wu
University of Tübingen, Tübingen, Germany

Abstract

Despite increased interest in creating more diverse and inclusive organizational environments, bias exists in how we choose leaders, who we interact with, and who we consider influential. Drawing from leadership emergence theory, we investigate potential interventions that support diverse leaders. Using agent-based simulations, we model a collective search process on a fitness landscape. Agents combine individual and social learning, and are represented as a feature vector blending relevant (e.g., individual learning characteristics) and irrelevant (e.g., race or gender) features. Agents use rational principles of learning to estimate feature weights on the basis of performance predictions, which are used to dynamically define social influence in their network. We show how biases arise based on historic privilege, but can be drastically reduced through the use of an intervention (e.g. mentorship). This framework allows us to test interventions best suited for unlearning bias in favor of performance-relevant traits.

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*