

Emergent Behaviors in Human-Robot Systems

RSS 2020 Workshop

July 12, 2020

Oregon State University at Corvallis, Oregon, USA **Virtual Event**

Key Facts

Event: RSS 2020 Workshop

Web: iliad.stanford.edu/rss-workshop

Date: July 12, 2020 (Full Day)

Contact: ebiyik@stanford.edu

Confirmed Speakers

Brenna Argall

Northwestern University

Anca Dragan

University of California, Berkeley

Judith Fan

University of California, San Diego

Jakob Foerster

Facebook AI & University of Toronto

Robert D. Hawkins

Princeton University

Maja Matarić

University of Southern California

Igor Mordatch

Google Brain

Harold Soh

National University of Singapore

Mac Schwager

Stanford University

Organizing Committee

Erdem Biyik

Stanford University

Minae Kwon

Stanford University

Dylan Losey

Stanford University

Noah Goodman

Stanford University

Stefanos Nikolaidis

University of Southern California

Dorsa Sadigh

Stanford University

Call for Papers

The workshop is soliciting papers for 5-minute spotlight talks. To participate, please submit a paper title and a 2 page abstract + references (using [RSS paper format](#)) via email with the subject "[RSS 2020 Abstract]" to: rss-workshop@cs.stanford.edu by **June 12, 2020 at 23.59 (Anywhere on Earth)**. All contributions will undergo a brief review by the organizers, and the authors will be notified of acceptance by **June 19, 2020**.

Accepted papers and eventual supplementary material will be made available on the workshop website. However, this does not constitute an archival publication and no formal workshop proceedings will be made available, meaning contributors are free to publish their work in archival journals or conference.

Abstract

Robots are increasingly becoming members of our everyday community. Self-driving cars, robot teams, and social and assistive robots operate alongside human end-users to carry out various tasks. Similar to the conventions between humans, which are the low dimensional representations that capture the interaction and can change over time, emergent behaviors form as a result of repeated long-term interactions in multi-agent systems or human-robot teams. Unfortunately, these emergent behaviors are still not well understood. For instance, the robotics community has observed that many different, and often surprising, robot behaviors can emerge when robots are equipped with artificial intelligence and machine learning techniques. While some of these emergent behaviors are simply undesirable side-effects of misspecified objectives, many of them significantly contribute to the performance in the task and influence other agents in the environment. These behaviors can further lead to developing conventions and adaptation of other agents, who are possibly humans, by encouraging them to approach the task differently.

Goal. We want to investigate how complex and/or unexpected robot behaviors emerge in human-robot systems, and to understand how we can minimize their risks and maximize their benefits. This workshop promotes a discussion on

- How emergent behaviors form in human-robot, human-human, and robot-robot settings
- How we can predict emergent behaviors
- What types of emergent behaviors we would like to see in real-world applications
- How their negative consequences and risks might be alleviated
- How they can be better utilized for more efficient human-robot or multi-agent interaction