Collaborative AI

We are interested in discussing the design and implementation of AI systems able to interact among themselves or with humans through representations that are understandable and meaningful to humans. We believe this will require unpacking the black box to provide a compositional structure that can 'explain' or 'justify' the opinion expressed by the AI system. Such representations might be structured in terms of so-called 'narratives' that can link elements of a situation to previously encountered themes or schemas, but could also be simpler structures possibly based on natural language. While this is not the traditional approach of explainable AI (XAI), it clearly places XAI as an enabler of collaborative AI, a focus that gives a different emphasis and aim to XAI. We are also interested in discussing new theoretical models for multi-agent collaboration and/or compositionality of learning systems. These could include ways of representing and using uncertainty estimates in order to combine evidence for a particular hypothesis that is communicable. Collaboration of any kind requires communication between agents, so they can coordinate their roles and the division of tasks and convey information pertinent to the execution of those tasks. While not restricted to this topic, the discussion will focus on how we might be able to advance our understanding of how artificial systems can coordinate their actions among themselves and/or with humans in a human understandable fashion.

Organisers

Nicolò Cesa-Bianchi, University of Milan (Ital) and John Shawe-Taylor, University College London (UK).

Panel

James Crowley, Antti Oulasvirta, John Shawe-Taylor, Nicolò Cesa-Bianchi

Statement of Prof. James Crowley

Transformers provide a revolutionary new technology for multimodal perception, action, and interaction that can be used to build new classes of intelligent systems that interact with people and enhance quality of life. Examples include devices that detect and evoke emotions (affectors), devices that enhance perception (media), devices that provide advice (advisors) and devices that interact and collaborate (collaborators). I will briefly discuss these classes, and then focus on a new approach for collaborative AI using a hierarchical framework that has emerged from the AI4EU T7.3 working group on Collaborative AI, and is currently part of the new Humane AI Net Strategic Research Agenda.

This hierarchical framework for collaborative AI organises research challenges based on the nature of the collaborative activity and the information that must be shared, with each level building on capabilities provided by lower levels. These levels concern reactive, situational, operational, praxical and creative collaboration. For each level I will provide a definition, describe examples, and discuss open research challenges. A summary of the research challenges raised by each level reveals three common research problems that occur in different forms at each level:

Comprehension, Explanation, and Learning. I will discuss how these three problems provide a grand challenge for Collaborative Artificial Intelligence

Statement of Prof. Antti Oulasvirta

In my statement to this panel, I will make a single point: In order to collaborate with humans, AI needs representations that allows it to reason about them by reference to psychological constructs known to be important in human-human social interaction.

Background: Social intelligence in humans refers to the 'common sense' we exhibit when interacting with other people, including the ability to make sense of them and to behave appropriately in a given situation [Kihlstrom2000). Social intelligence, as an intellectual capability, includes both knowledge aspects (knowing what) and functional aspects (knowing how). The basic cognitive processes underpinning such abilities include the perception of social stimuli, the integration of perceptions with contextual knowledge, and the representation of possible responses to a situation.

Our goal: To build "artificial social intelligence", we need to study the foundations of a novel type of AI that could better collaborate with people by reasoning about factors behind people's observed behavior, in particular their beliefs, capabilities, motivations, and attitudes towards others and towards cooperation; in other words AI that can reason about social cognition.

What we do: To this end, our group is studying a key enabling technology: simulator-based machine learning that combines (1) psychologically plausible simulator models with (2) a powerful machine learning approach suitable for interactive settings with humans. In particular, we will study a novel class of computationally rational models and efficient Bayesian inference methods, respectively. These methods complement the prevailing methods in interactive AI, especially supervised and reinforcement learning, by enabling more accurate, data-efficient and human-like reasoning about human partners.

James L. Crowley is Professor Emeritus at the Institut Polytechnique de Grenoble, and holds the Chair of Collaborative Intelligent Systems at the Univ. Grenoble Alpes Multidisciplinary Institute for Artificial Intelligence (Miai@UGA). His current research combines multi-modal interaction with cognitive modelling to explore new forms of interaction with intelligent systems.

Prior to October 2021, James L. Crowley directed the Pervasive Interaction research team at the Inria Research Centre of the. Univ Grenoble Alpes. He was awarded a Ph.D. from Carnegie Mellon University in 1982, with a thesis on invariant representations for Computer Vision. In Sept. 2011, he was appointed as Senior Member of the l'Institut Universitaire de France (IUF). In Mar. 2014, he was named Chevalier de l'Ordre National du Mérite. A complete CV, selected publications and recordings of lectures and invited talks may be found at http://crowley-coutaz.fr/jlc/jlc.html.

Nicolò Cesa-Bianchi is professor of Computer Science at the University of Milan, Italy. His main research interests are the design and analysis of machine learning algorithms for statistical and online learning, multi-armed bandit problems, and graph analytics. He is co-author of the monographs "Prediction, Learning, and Games" and "Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems". He served as President of the Association for

Computational Learning and co-chaired the program committees of some of the most important machine learning conferences, including NeurIPS, COLT, and ALT. He is the recipient of a Google Research Award, a Xerox Foundation Award, a Criteo Faculty Award, a Google Focused Award, and a IBM Academic Award. He is member of the ELLIS management board and co-director of the ELLIS program on Interactive Learning and Interventional Representations.