The Origins of Common Sense in Humans and Machines

Kevin A Smith,[1,2] **Eliza Kosoy,**[3] **Alison Gopnik,**[3] **Deepak Pathak,**[4,5] **Alan Fern,**[6] **Joshua B Tenenbaum,**[1,2] **& Tomer Ullman**[2,7]
1. MIT BCS, 2. CBMM, 3. UC Berkeley Psychology, 4. CMU Robotics, 5. FAIR, 6. OSU Computer Science, 7. Harvard Psychology

Keywords: commonsense reasoning; core knowledge; intuitive theories; machine learning

Overview

People impose structure on raw percepts, filling the world with objects, agents, events, and properties. This reasoning develops early: by their first birthday, infants can determine features of objects such as number and motion (Spelke, 1990; Wynn, 1992), construe agents' actions as goal-directed (Woodward, 1998; Gergely & Csibra, 2003), and distinguish helpers from hinderers (Hamlin, Wynn, & Bloom, 2007). Agents, objects, events, properties – these are the building blocks of meaning and common sense that allow even young children to rapidly understand and interact with novel scenes.

Despite recent great strides in several domains (e.g., Devlin, Chang, Lee, & Toutanova, 2019; Silver et al., 2016), artificial systems cannot yet understand and act on the world in a human-like manner, and do not yet perform basic commonsense reasoning at the level of even young children (Lake, Ullman, Tenenbaum, & Gershman, 2017). What are these AI systems missing? Is it certain learning mechanisms, initial primitives, motivations, or the right environments and curricula?

We focus on young children's common sense as a step towards building general machine common sense: by understanding the origins of common sense in humans, we hope to understand how to recapitulate it in machines. In turn, by looking at the successes and failures of machines, we can make scientific progress towards understanding the initial state and learning mechanisms of human intelligence. This approach has spurred much interest across various fields of cognitive science and artificial intelligence, and investment from agencies that hope to design artificial systems with common sense. The organizers of this workshop are all members of a large, multi-site DARPA grant aimed at developing machine common sense inspired by the capabilities of young children, and represent a diverse set of teams and approaches to meeting this challenge.

The origin, development, and engineering of common sense is a broad problem that is studied in many different ways by many different members of the Cognitive Science community. We aim to use this workshop as a forum to discuss this problem with the broader community by considering theories and approaches for understand and building common sense, presenting experimental research that probes the foundations of common sense in people, and reporting on progress made by the DARPA project teams and others on building artificial agents with infant-like commonsense reasoning capabilities. We have brought together leading researchers from cognitive psychology, developmental psychology, machine learning, neuroscience, and robotics, all of whom share a strong background and interest in the cognitive systems underlying common sense in humans, or in the effort to engineer commonsense reasoning into artificial agents in ways inspired by human reasoning.

More specifically, the workshop will be a forum for addressing the following inter-related key questions:

- What commonsense knowledge is "built-in" in infants, and what must be learned?
- What systems or strategies do infants have to learn common sense efficiently?
- How unitary or diverse are the cognitive systems underlying commonsense reasoning?
- What structures and learning mechanisms are missing from current artificial systems that would be required for common sense?
- How would we evaluate whether an artificial system has "common sense"?

The workshop website is at https://ocs2020.github.io

Target audience

This workshop is very closely tied to this year's theme of "Developing a mind: Learning in humans, animals, and machines", as we plan to discuss how common sense develops in both humans and machines. The workshop should be of interest to a broad audience that would include most of the CogSci conference, including cognitive psychologists, developmental psychologists, neuroscientists, roboticists, machine learning researchers, and philosophers.

Organizers and presenters

Kevin Smith (Organizer) is a Research Scientist at MIT, working with Josh Tenenbaum. He studies how people use physical reasoning for commonsense tasks such as prediction, learning about object properties, and action planning.

Eliza Kosoy (Organizer) is a graduate student at UC Berkeley advised by Alison Gopnik. She is interested in understanding how children learn new concepts quickly and with little data, and studies these questions by using tools from artificial intelligence and developmental psychology using techniques such as placing kids into environments traditionally created for testing artificial agents and having children explore them.

Alison Gopnik (Organizer) is a Professor of Psychology and Philosophy at UC Berkeley. She studies how children explore and learn about the causal structure of the world, and how this learning allows them to develop intuitive theories.

Deepak Pathak (Organizer) is an Assistant Professor of Machine Learning & Robotics at CMU and Researcher at Facebook AI Research (FAIR). He researches how to design systems with a human-like ability to generalize in real and diverse environment, by continually developing knowledge and acquiring new skills from raw sensory data.

Alan Fern (Organizer) is a Professor of Computer Science at Oregon State. His research is focused on making machines smarter by integrating machine learning with planning and control. He is especially interested in building AI systems that can learn models of the environment from experience that can be used for effective planning of behavior.

Joshua Tenenbaum (Organizer) is a Professor of Cognitive Science at MIT. His research sits at the intersection of cognitive science and machine learning, and is focused on understanding the distinctively human ability to learn about the world rapidly and flexibly from little data.

Tomer Ullman (Organizer) is an Assistant Professor of Psychology at Harvard. He is interested in common-sense reasoning, and building computational models of how children and adults form intuitive theories of agents and objects.

Moira Dillon is an Assistant Professor of Psychology at NYU. Dillon's lab uses cognitive, developmental, and computational approaches to understand how experience with the objects, agents, and places of everyday life might form the foundation of human symbolic and abstract thought.

Stanislas Dehaene is a Professor and Chair of Experimental Cognitive Psychology at the Collège de France. He studies the neural bases of cognitive processes such as reading, numerical cognition, and consciousness.

Jessica Sommerville is a Professor of Psychology at the University of Toronto. She studies cognitive and social development from infancy through preschool with an emphasis on social and moral cognition and behavior, early persistence and action/perception relations.

Shari Liu is a graduate student at Harvard advised by Elizabeth Spelke. She studies the origins of social intelligence: the representations and computations supporting our earliest understanding of the beliefs, goals, and costs motivating other people's actions.

Ori Ossmy is a postdoctoral scholar with Karen Adolph at NYU. He is interested in building systems that learn to solve everyday problems using an integration of theory and methods from child development, AI, and cognitive neuroscience.

Matthew Botvinick is the Director of Neuroscience Research at DeepMind and an honorary professor of neuroscience at UCL. He is interested in using tools from neuroscience to build more advanced models of artificial intelligence, and using models from machine learning to better understand brain circuitry.

Tucker Hermans is an Assistant Professor in the School of Computing at the University of Utah. His research is focused on autonomous learning, planning, and perception for robot manipulation; in particular he is interested in enabling robots to autonomously discover and manipulate objects with which they have no previous knowledge or experience.

Workshop structure

This workshop would span a full day, split into three sessions. The first session, *Frameworks for Common Sense*, will focus on theoretical accounts of the systems required for common-sense reasoning. In the second session, *Common Sense in Humans*, presenters will discuss recent advances in studying the origins of common sense in people. The third session, *Common Sense in Machines*, will focus on recent attempts to build artificial agents that perform commonsense tasks like human infants. We will end the workshop with a panel discussion that will tie together and contrast various theories of common sense, as well as cover future directions for studying and reverse engineering common sense.

Presenter	Topic
Frameworks for Common Sense	
Alison Gopnik	Children are MESS-y: Model-building, Exploratory, Social learning systems allow the emergence of common sense
Joshua Tenenbaum	Tools for reverse engineering common sense
Alan Fern	A model-based reinforcement learning perspective on common sense learning
Moira Dillon	Cognitive Artificial Intelligence: Building better machines... and babies!
Common Sense in Humans	
Stanislas Dehaene	Advances in understanding human geometrical intuition
Jessica Sommerville	The extents and limits of infants' socio-moral cognition and behavior
Eliza Kosoy	Curiosity and exploration in children's maze playing using DeepMind Lab
Shari Liu	Origins of social intelligence in human infants
Ori Ossmy	A behavioral approach to the development of common sense
Common Sense in Machines	
Matthew Botvinick	Learning intuitive physics (almost) from scratch
Deepak Pathak	Learning to generalize via self-supervised exploration
Kevin Smith	Building models of infants' physical understanding
Tucker Hermans	Can common sense guide autonomous robot exploration?

References

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805 [cs]*. (arXiv: 1810.04805)

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: the naıve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, e253.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., … Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, *14*(1), 29–56.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749–750.