Research Statement

Ross A. Knepper

Robots hold the promise of helping us solve many of society's greatest challenges, from feeding the planet to caring for an aging population. Robots have seemingly been on the verge of pervasive deployment among people for decades, and yet there are virtually no commercially viable personal robots on the market. To achieve such deployment, we need to gain an understanding of the **science of robot systems**. Robots are complex, engineered, intelligent systems built from many interacting components. From these interactions arise emergent behaviors that must be studied and opportunities for building inherently robust robot systems. My research addresses both the science and engineering of robot systems using algorithms for planning, perception, control, manipulation, learning, and human-robot interaction.

1 Defining the Robot Systems Problem

The study of *robot systems* considers trade-offs that arise when components are combined to build complex, intelligent, embodied systems. I have long led by example in illustrating the importance of systems-level thinking in robotics research [1–12]. In contrast, much of the robotics research community focuses on either a particular application or a component in isolation, such as planning, sensing, mapping, localization, and control. Often neglected are the emergent phenomena that occur when putting components together. These phenomena should be understood at a level that transcends particular applications. Scientific questions surrounding these phenomena comprise one major goal of my research.

On the engineering side, I am interested in building robot systems that can function for prolonged periods of time, whether working for humans or with them. These systems must provide a variety of assurances in order to be trustworthy, including interpretability, responsiveness, and interactivity. Superficially, a focus on systems appears to compound the many research problems faced by those who study the components in isolation, but a more sophisticated view reveals synergies that simplify rather than complicate the picture. As Eisenhower noted, *if you can't solve a problem, make it bigger*. I discuss two areas of recent work:

- **Robots working for people.** When a robot acts on behalf of people, it must be robust to unpredictable and changing circumstances. Contemporary machine learning approaches promise to allow robots to capture and generalize complex patterns from training data, but they come without assurances that instill trust. In Section 2, I describe robots that learn to generalize reliably by hybridizing learning techniques with traditional methodologies.
- **Robots working with people.** In order for a robot to work effectively on a team with people, it must understand the innate human practice of implicitly conveying information. Implicit communication is so innate that people assume robot teammates use it as humans do [13]. In Section 3, I discuss robot systems that can participate fully and responsively to the needs of the team using the ability to recognize and generate implicit communication.

2 Systems That Act on Humans' Behalf Must Provide Assurances

Robotics engineers have long enjoyed the ability to construct whatever internal algorithms and representations best accomplish the mission. These traditionally hand-designed systems were grounded in logical principles that reassured us that they would function predictably, but the assumptions that were built into these systems limited their generalizability in the complex, messy real world. Nowadays, many powerful machine learning techniques learn subtleties of real phenomena like physics and human behavior, yet modern learning-based techniques make no assurances of performance and interpretability. Most future autonomous robot systems will need to operate out in the world without a trained human expert standing by to take control if something goes wrong. I build systems that combine modern machine learning with traditional methods in order to provide necessary assurances such as interpretability and interactivity.

Hybrid autonomy. A powerful way to mitigate the limitations of autonomy is to leverage assistance from human bystanders. To achieve this, robots must learn to communicate like people. I built a distributed, robotic assembly system called IkeaBot that accomplishes this. Besides being the **first autonomous robot system to build Ikea furniture**, IkeaBot's importance stems from its adaptability. The system intelligently reconfigures itself and can build a new product in minutes, whereas the same transition for traditional factory robots could take experts a week to complete. IkeaBot's generality gives rise to a nearly endless list of very low-probability failures. IkeaBot therefore augments its autonomy by delegating resolution of unhandled failures to people [5]. To resolve these failures, the system makes a spoken help request. My team built a learned semantics model that converts a symbolic internal representation of a system failure into natural language that maximizes the chance of a human correctly interpreting the help request. The failure resolution system was the first to be able to acquire a variety of sophisticated help from untrained, situationally-unaware humans. Previous systems, in contrast, place cognitive load on people and require expertise to diagnose and repair failures. The IkeaBot system paper was a **finalist for the Best Automation Paper award** at ICRA 2013 [4], and the failure resolution work received the **Best Paper award** at RSS 2014 [14]. I delivered the IkeaBot system software to my sponsor, Boeing, for deployment on their airplane assembly line.

Figure 1: DeepMPC masters the complex dynamics of food cutting by solving the dynamics and control problems separately [8].

Hybrid methodology. One way to extend the generalizability of machine learning is to hybridize it with established robotics techniques. My team pioneered a method called DeepMPC to perform control on systems with complex dynamics [8]. The approach combines a learned representation of the dynamics model with a standard model-predictive controller. Many researchers are trying to develop learned direct control policies, but such approaches tend to generalize poorly. Our approach, in contrast, is much more versatile since it decouples the learned dynamics model from the well-studied classic control problem. In this work, we focused on the dynamics of food cutting (Figure 1). Our method has an **order of magnitude lower sample complexity** than competing approaches. By varying cutting speed and force, the robot can adapt to dozens of different food consistencies that it was not trained on. **The controller adapts online by discovering manipulation strategies**, such as slicing, sawing, and chopping. I coadvised Ph.D. student Ian Lenz on this project and oversaw the completion of Lenz's thesis in 2015. He became a Postdoctoral Researcher with Prof. Andrea Thomaz (ECE, University of Texas at Austin).

Hybrid structure. The end-to-end deep learning approach to building robot systems circumvents many hard engineering decisions involving structure and intermediate representation. In its pure form, this approach obviates the traditional computational modules used by robots, such as perception, mapping, planning, and control. Recognizing that there is wisdom embedded in traditional robot architectures, my team has replicated this global structure within a neural network. We have developed a robot system capable of visuomotor navigation (from pixels to control inputs) of a quadcopter guided by natural language commands [15, 16]. The network is structured as a series of maps that are used for perception, planning, and control. A consequence of this architecture is that the network is interpretable. It is straightforward to see by inspection of the network where objects of interest are located in the map as well as motion plans and goals. These traits enable us to articulate and test meaningful properties of the neural network that would be impossible with an opaque end-to-end system. My team has also demonstrated sim-to-real domain transfer with this system, which enables the robot to **train almost entirely in simulation** (thousands of trials) **with only tens of real quadcopter training examples** [12]. We achieved this result by exploiting the system architecture to isolate real robot training to a small functional unit within the overall network. As a result, the system is capable of generalizing effectively to new environments and landmarks. This work was supported by a Young Investigator Program award from AFOSR as well as several Amazon research awards.

Future work. We expect robots to generalize appropriately beyond their training data even when we cannot precisely specify what correct behavior looks like. Often, this reduces to the Potter Stewart test, "I know it when I see it." Yet machine learning methods face a danger of fooling us into overestimating their efficacy [17], particularly in robotics. Benchmarks compare competing algorithms but also optimize them to excel on the benchmark itself, sometimes at the loss of overall performance. This highlights the danger of relying solely on empirical validation that is specific to a single task. I am a leader in deploying the full range of validation tools for robot systems, like proofs [18–20], formal verification [21], scalability [22], ablation studies [12, 15], and human subject studies of the integrated system [5, 10]. Using the instruction-following quadcopter, I will explore how these tools can be leveraged holistically to validate generalization with concrete assurances by hybridizing these techniques with data-driven benchmarking methods.

3 Communication and Action Within a Team Are Entangled

An important class of systems is human-robot teams. In building robots that work effectively with humans, the fundamental challenge is to perform *joint computation between robots and humans*. People already are capable of performing joint computation with one another: we call it teamwork. I seek to understand the essential properties of teamwork and apply them to robots. Achieving this goal requires not only new robot-human interfaces but also new robot algorithms and representations that mirror human computation.

One prominent property of human teamwork is the entanglement of action and communication. Any observable action performed within a team communicates to the team, regardless of whether its intent is primarily communicative or functional. Human teams routinely leverage this communication for situational awareness and efficiency (Figure 2). In robotics, communication and action are typically approached as separate problems, but robots will not be accepted as peers until they can understand the deeper significance of human actions and also choose their own actions carefully to communicate the correct messages to human teammates. I am generating major advances in a robot's ability to reason holistically about action and communication.

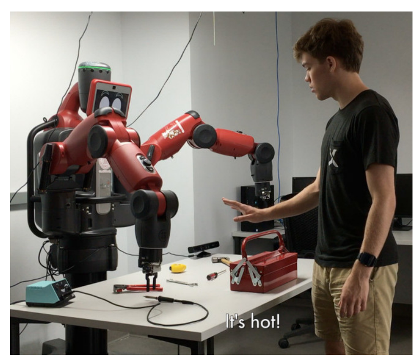

Figure 2: A human user gestures *stop* while saying "It's hot!". The robot uses its knowledge of implicit communication to understand the meaning that it should not grasp the soldering iron by its tip [9].

Implicit communication framework. People perform actions both for communicative and functional purposes, but the line separating the two is decidedly blurry. I was the PI on a $1M grant from ONR that studied the entanglement of action and communication in a human-robot team. I described and formalized the phenomenon of *implicit communication* [9], which describes the method by which people intuitively encode and decode communicative meaning within actions. Humans are so accustomed to communicating by implicit means that they are often later unable to recall whether they learned a particular fact implicitly or explicitly [23]. Since robots traditionally communicate only explicitly, they miss half of the conversation. My research provides a **mathematical framework that unifies diverse results** from the fields of robotics, linguistics, and psychology to form a single computational mechanism for implicit communication that has far-reaching consequences across many communities. This result will empower a social intelligence for robots that generalizes across diverse tasks and robot designs to a much greater degree than the state of the art. This research was a **finalist for the Best Technical Paper award** at HRI 2017 [9].

Social navigation. Navigating a mobile robot in a crowd of people has commonly been thought of as a collision avoidance problem: humans move how they want, and robots must react appropriately. Yet this perspective neglects the degree to which the participants are interdependent. I am the first to recognize and study robotic social navigation as a first-class example of teamwork, and these insights have led to the first socially-competent robot for navigation among pedestrians. To support this work, I am the computational co-PI on three different grants from the NSF bringing in over $3M to study social robot navigation.

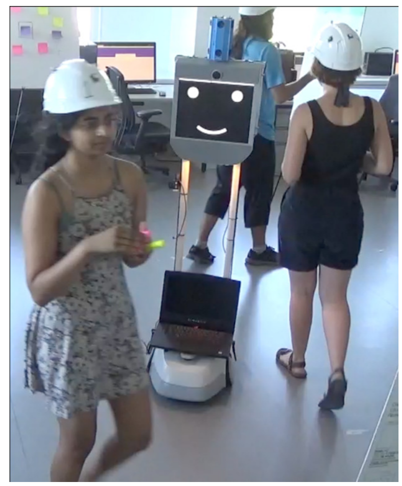

Figure 3: Our socially-competent robot navigates among people by modeling the group's motion topologically [10, 24–29].

Recent results in my lab show the link between action and communication during social navigation. My team has focused on structures from algebraic topology, particularly the *braid group*, to represent possible motions of individual human or robot agents in a group. These topological representations allow a robot to generate motions that balance the objectives of collision avoidance and communication of intentions to and from human pedestrians [24, 28] (Figure 3). This pedestrian navigation algorithm is the **first to consider the interaction of the robot with an arbitrary, scalable number of people moving intuitively in real time**. The reactive control problem for social navigation requires that a robot be highly reactive by both interpreting human intentions and expressing its own intentions in real time [10, 26]. We formulated a Hamiltonian dynamical system that conserves energy by following a specified braid [27, 29]. Compared to gradient-descent-based optimizers, this method is **orders of magnitude more efficient and is capable of following arbitrary braids correctly** while also optimizing for other path-shape criteria like social signaling. I graduated Ph.D. student Christoforos Mavrogiannis on this research in March 2019. He is now a Postdoctoral Research Associate with Prof. Siddhartha Srinivasa (CSE, University of Washington).

Action as delegation. Building on the insights from these results, I began to explore the general role of *actionable implicit communication*, in which teammates implicitly delegate actions for others to perform. I found a simple model of this phenomenon in the collaborative card game Hanabi [11]. My team showed that using an actionable implicit communication strategy can increase task performance by about 50% over a pure information strategy. Despite this gain, our web-based human user study paradoxically revealed no benefit to the implicit strategy in live gameplay with an AI. The web interface supports the full game rules, but its lack of embodied copresence appears to block effective teamwork. This result suggests the fast-growing research effort in pure-software Hanabi AIs [30–32] is barking up the wrong tree: embodiment is a vital component of human Hanabi play, which motivates the study of Hanabi strategy using robots. This paper won **Best Paper Honorable Mention** at CHI 2019 [11].

Future work. My lab recently conducted an exploratory study of how Hanabi players use implicit social cues to guide their partner's decisions. We found that people bend the rules of the game, sometimes unknowingly, by subtly "leaking" implicit information over a variety of natural modalities. Although this is technically cheating in the game, it is good teamwork in any other context. We are currently building an embodied Hanabi-playing robot system that will employ the same tactics to leak implicit information to human players. In a controlled laboratory environment, we will measure the effect of these subtle behaviors on teamwork quality. We will fold these lessons back into a broader theory of communication and action in teams. Then we will build domain-independent robot skills that leverage intuitive human understanding of a situation to implicitly communicate task status, delegate activities, and coordinate goals.

4 Future Directions

In my three years at MIT, I exposed and solved several important robot systems problems. My five years at Cornell culminated in robot systems advances that enable effective coordination and collaboration with people. I plan to pursue two new areas in the next five years.

Mental models for human-robot teamwork. I am the PI on a recent ONR grant worth $2.6M to build computational models that robots can use to understand human teammates. I am partnering with Prof. Melissa Ferguson (Psychology, Cornell) and Prof. Julie Shah (Aeronautics and Astronautics, MIT) to advance the psychology of impression formation and develop algorithms that enable robots to learn mental models of human teammates. Using these models, we will investigate how robots can better calibrate human expectations to improve team performance.

Figure 4: The Baxter Deli system can take verbal orders and serve food to customers. By manipulating success and failure, we use it as a testbed for how people form mental models of robots.

In preliminary work, we developed new methods for the use of robots in human experiments to study the psychology of impression formation [33, 34]. We **overturned a conventional belief** which states that a person's implicit impressions, once formed, are durable except in the face of strong counter-evidence [35]. Thus, people update impressions of robots differently than of other people (Figure 4). This result suggests the need for further study in psychology but also bolsters the notion that robots can calibrate a human's mental model of them over time.

New architectures for robust robot systems. The unexamined assumptions built into most robots make implicit trade-offs that frequently go unstudied. I recently articulated some scientific best practices of good robotics systems research [36] that will help us make informed trade-offs. Many aspects of the traditional robot software architecture have hardly changed in fifty years [37]. The rigid assumptions baked into the prevailing architecture can lead to brittleness of performance caused by the system's sensitivity to unmodeled parameters. Some of these assumptions are that the robot's environment is static, that sensors reliably describe objects, and that scenes are not excessively cluttered with objects. These assumptions are violated in spaces shared with people. I will revisit the components and interfaces of a robot system architecture in order to reduce sensitivity to the many sources of uncertainty in the world.

References

[1] Ross A. Knepper, Siddhartha S. Srinivasa, and Matthew T. Mason. An equivalence relation for local path sets. In *Proceedings of the Workshop on the Algorithmic Foundations of Robotics*, Singapore, December 2010.

[2] Ross A. Knepper and Matthew T. Mason. Realtime informed path sampling for motion planning search. In *Proceedings of the International Symposium on Robotics Research*, Flagstaff, USA, August 2011.

[3] Siddhartha S. Srinivasa, Dmitry Berenson, Maya Cakmak, Alvaro Collet, Mehmet R. Dogar, Anca D. Dragan, Ross A. Knepper, Tim Niemueller, Kyle Strabala, Mike Vande Weghe, and Julius Ziegler. HERB 2.0: Lessons learned from developing a mobile manipulator for the home. *Proceedings of the IEEE*, 100(8):2410–2428, August 2012.

[4] Ross A. Knepper, Todd Layton, John Romanishin, and Daniela Rus. Ikeabot: An autonomous multi-robot coordinated furniture assembly system. In *Proceedings of the IEEE International Conference on Robotics and Automation*, Karlsruhe, Germany, May 2013.

[5] Ross A. Knepper, Stefanie Tellex, Adrian Li, Nicholas Roy, and Daniela Rus. Recovering from failure by asking for help. *Autonomous Robots*, 39(3):347–362, October 2015.

[6] Mehmet Dogar, Ross A. Knepper, Andrew Spielberg, Changhyun Choi, Henrik I. Christensen, and Daniela Rus. Multi-scale assembly with robot teams. *International Journal of Robotics Research*, 34(13), November 2015.

[7] Javier Alonso-Mora, Ross A. Knepper, Roland Siegwart, and Daniela Rus. Local motion planning for collaborative multi-robot manipulation of deformable objects. In *Proceedings of the IEEE International Conference on Robotics and Automation*, Seattle, USA, May 2015.

[8] Ian Lenz, Ross A. Knepper, and Ashutosh Saxena. DeepMPC: Learning deep latent features for model predictive control. In *Proceedings of the Robotics Science and Systems Conference*, Rome, Italy, July 2015.

[9] Ross A. Knepper, Christoforos I. Mavrogiannis, Julia Proft, and Claire Liang. Implicit communication in a joint action. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, Vienna, Austria, March 2017.

[10] Christoforos I. Mavrogiannis, Alena Hutchinson, John Macdonald, Patrícia Alves-Oliveira, and Ross A. Knepper. Effects of distinct robot navigation strategies on human behavior in a crowded environment. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, Daegu, South Korea, March 2019.

[11] Claire Liang, Julia Proft, Erik Andersen, and Ross A. Knepper. Implicit communication of actionable information in human-AI teams. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, Glasgow, United Kingdom, May 2019.

[12] Valts Blukis, Yannick Terme, Eyvind Niklasson, Ross A. Knepper, and Yoav Artzi. Learning to map natural language instructions to physical quadcopter control using simulated flight. In *Proceedings of the Conference on Robot Learning*, Osaka, Japan, October 2019.

[13] Cynthia Breazeal, Cory D Kidd, Andrea Lockerd Thomaz, Guy Hoffman, and Matt Berlin. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *2005 IEEE/RSJ international conference on intelligent robots and systems*, pages 708–713. IEEE, 2005.

[14] Stefanie Tellex, Ross A. Knepper, Adrian Li, Daniela Rus, and Nicholas Roy. Asking for help using inverse semantics. In *Proceedings of the Robotics Science and Systems Conference*, Berkeley, USA, July 2014.

[15] Valts Blukis, Dipendra Misra, Ross A. Knepper, and Yoav Artzi. Mapping navigation instructions to actions with position visitation prediction. In *Proceedings of the Conference on Robot Learning*, Zurich, Switzerland, October 2018.

[16] Valts Blukis, Nataly Brukhim, Andrew Bennett, Ross A. Knepper, and Yoav Artzi. Following high-level navigation instructions on a simulated quadcopter with imitation learning. In *Proceedings of the Robotics Science and Systems Conference*, Pittsburgh, USA, June 2018.

[17] Atticus Geiger, Ignacio Cases, Lauri Karttunen, and Chris Potts. Posing fair generalization tasks for natural language inference. *arXiv preprint arXiv:1911.00811*, 2019.

[18] Ross A. Knepper, Siddhartha S. Srinivasa, and Matthew T. Mason. Toward a deeper understanding of motion alternatives via an equivalence relation on local paths. *International Journal of Robotics Research*, 31(2):168–187, February 2012.

[19] Ross A. Knepper and Daniela Rus. On the completeness of ensembles of motion planners for decentralized planning. In *Proceedings of the IEEE International Conference on Robotics and Automation*, Karlsruhe, Germany, May 2013.

[20] Abhishek Anand and Ross Knepper. Roscoq: Robots powered by constructive reals. In *The 6th Conference on Interactive Theorem Proving*, Nanjing, China, August 2015.

[21] William Thomason, Abhishek Anand, Greg Morrisett, and Ross A. Knepper. Automatic distributed multi-agent coordination of single-agent robot controllers, in submission 2020.

[22] Ross A. Knepper and Daniela Rus. Pedestrian-inspired sampling-based multi-robot collision avoidance. In *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication*, Paris, France, September 2012.

[23] Steve Nicolle and Billy Clark. Experimental pragmatics and what is said: a response to Gibbs and Moise. *Cognition*, 69(3):337–354, 1999.

[24] Christoforos I. Mavrogiannis and Ross A. Knepper. Decentralized multi-agent navigation planning with braids. In *Proceedings of the Workshop on the Algorithmic Foundations of Robotics*, San Francisco, USA, December 2016.

[25] Christoforos I. Mavrogiannis, Valts Blukis, and Ross A. Knepper. Socially competent navigation planning by deep learning of multi-agent path topologies. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vancouver, Canada, September 2017.

[26] Christoforos I. Mavrogiannis, Wil Thomason, and Ross A. Knepper. Social momentum: A framework for legible navigation in dynamic multi-agent environments. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, Chicago, USA, March 2018.

[27] Christoforos I. Mavrogiannis and Ross A. Knepper. Multi-agent trajectory prediction and generation with topological invariants enforced by hamiltonian dynamics. In *Proceedings of the Workshop on the Algorithmic Foundations of Robotics*, Mérida, Mexico, December 2018.

[28] Christoforos I. Mavrogiannis and Ross A. Knepper. Multi-agent path topology in support of socially competent navigation planning. *International Journal of Robotics Research*, 2018.

[29] Christoforos Mavrogiannis and Ross A. Knepper. Decentralized navigation planning with topologically robust multi-agent trajectory prediction. *International Journal of Robotics Research*, In submission.

[30] Hirotaka Osawa. Solving Hanabi: Estimating hands by opponent's actions in cooperative game with incomplete information. In *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[31] Bruno Bouzy. Playing Hanabi near-optimally. In *Advances in Computer Games*, pages 51–62. Springer, 2017.

[32] Nolan Bard, Jakob N Foerster, Sarath Chandar, Neil Burch, Marc Lanctot, H Francis Song, Emilio Parisotto, Vincent Dumoulin, Subhodeep Moitra, Edward Hughes, et al. The Hanabi challenge: A new frontier for ai research. *arXiv preprint arXiv:1902.00506*, 2019.

[33] Minae Kwon, Malte F. Jung, and Ross A. Knepper. Human expectations of social robots. In *Late Breaking Report at the ACM/IEEE International Conference on Human-Robot Interaction*, Christchurch, New Zealand, March 2016.

[34] Minae Kwon, Thomas Mann, Melissa Ferguson, and Ross A. Knepper. An exploration of attitudes towards robots using implicit and explicit measures. In preparation.

[35] Melissa J Ferguson, Thomas C Mann, Jeremy Cone, and Xi Shen. When and how implicit first impressions can be updated. *Current Directions in Psychological Science*, pages 331–336, 2019.

[36] Soham Sankaran and Ross A. Knepper. Usability squared: Principles for doing good systems research in robotics. *arXiv preprint arXiv:1906.06775*, 2019.

[37] Benjamin Kuipers, Edward A. Feigenbaum, Peter E. Hart, and Nils J. Nilsson. Shakey: From Conception to History. 38(1):88–103, 2017.