1

Looking Forward to Interactive Task Learning

Kevin A. Gluck and John E. Laird

Human learning has been the subject of extensive research in multiple areas of science. People are always learning, from whatever sources of knowledge are available. Language and other forms of natural communication enable us to master novel tasks quickly; once we do, we often share the resultant knowledge with others. In just a few minutes, we can grasp how to play a new game, use a new device (e.g., smart phones, industrial machinery), or assist a disabled family member in meeting specific challenges. Importantly, as we learn and hone performance on a task, we adapt in real time to emergent needs—sometimes figuring things out for ourselves, sometimes interacting with others to gain efficiencies or address any problems we encounter.

Contemporary artificial agents, by contrast, are bound to the specific tasks for which they were originally programmed. Even systems designed to acquire knowledge and expertise can learn only a single task at a time (e.g., Chess, Go, video games), becoming idiot savants with amazing capabilities, but without any abilities beyond that narrow specialization. Without doubt, advances in artificial intelligence, cognitive science, and robotics point to future systems with sufficient cognitive and physical capabilities to perform a wide variety of diverse tasks. But how will they learn tasks that arise unexpectedly—tasks that cannot be anticipated and therefore preprogrammed or trained for? How can agents pursue a task when there is insufficient prior knowledge or time for exploration to guide learning?

Interactive task learning (ITL) attempts to answer those questions by providing a conceptual framework for agents to learn not only how to perform tasks better, but also to learn new tasks from scratch through natural, real-time interactions with others. ITL involves interactions between an agent (human or machine), its world, and, crucially, other agents in the world. ITL is a bidirectional process between teacher and learner (both of which can be humans or machines) that results in collaboration and knowledge creation.

The catalyzing idea behind ITL is as follows: for artificial systems to learn from and teach us entirely new things at any given moment, we must advance beyond the traditional approach of creating specialized AI agents for single,

From "Interactive Task Learning: Humans, Robots, and Agents Acquiring New Tasks through Natural Interactions," edited by K. A. Gluck and J. E. Laird. Strüngmann Forum Reports, vol. 26, J. R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-03882-9.

predetermined niche purposes and instead incorporate a rich set of natural interaction and learning mechanisms into our systems. This involves two crucial requisites. First, the way in which artificial systems learn and teach new tasks must be natural for people, not constrained by traditional programming and digital forensics. Second, learning and performing multiple tasks can only be bounded by physical and informational limits, not by design, implementation, or optimization for single task performance.

The concept of ITL is controversial, in that it disrupts the status quo and involves diverse challenges. ITL requires theoretical and practical advances in the integration of a broad range of capabilities associated with cognition, including extracting task-relevant meaning from perception, task-relevant action, grounded language processing, dialogue and interaction management, integrated knowledge-rich relational reasoning, problem solving, learning, and metacognition. This integration contrasts with the general trend toward increasing fragmentation and focus on narrow capabilities and problems. Beyond these challenges, ITL creates an opportunity to rethink the fundamental nature of our most advanced and capable artifacts. How can we move beyond artifacts that are designed for a single use or purpose, to ones that can be dynamically adapted to our changing needs, increasing the rate of our progress and the quality of our lives as individuals and societies? Isolated efforts to develop more intelligent agents and robots are already underway in areas such as healthcare, in-home assistance, education, and transportation. We propose the missing link among them is the unifying vision of ITL.

Our optimism that the time is right for a coordinated and concerted push toward ITL is grounded in our assessment that despite shortcomings, gaps, and challenges, the research community has made progress on important component capabilities and their integration. To shape R&D investments in a way that advances ITL in artificial systems requires the identification of broad organizing themes. *Pace*, *persistence*, and *partnering* are core characteristics that constitute research challenges around which we can rally our science and technology investments.

Pace, Persistence, and Partnering

The human capacity for rapid, nearly instantaneous learning of entirely new tasks on the basis of brief communications and one, two, or a few demonstrations sets a *pace* requirement for ITL. The pace of the interaction, the pace of the teaching and learning, and the pace of task completion must all occur on timescales aligned with and amenable to real-time human experience.

Humans are engines of creation. From the imaginative play of early childhood, to the generative nature of language, to scientific discovery and technological innovation, we are constantly creating new constructs, concepts, and capabilities. Our *persistence* throughout these activities requires us to both

From "Interactive Task Learning: Humans, Robots, and Agents Acquiring New Tasks through Natural Interactions," edited by K. A. Gluck and J. E. Laird. Strüngmann Forum Reports, vol. 26, J. R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-03882-9.

assimilate and accommodate newly gained knowledge and existing knowledge. As we work to create intelligent artifacts that interact, we must recognize that it will never be possible to anticipate and represent, in advance, all the knowledge and skill that may be required in the future. ITL agents must be able to learn and adapt continually with robust success, over a long period of time in environments that are dynamic, nonstationary, and boundlessly novel.

Human beings help each other. It is what we do. We organize in ways that support joint objectives and goals. Our most valued relationships are with family, friends, partners, and teammates. These relationships develop over time out of shared experiences in which we demonstrate an ability and willingness to be there for each other in times of need. We are at our best when we take the initiative to assist or compensate without being asked, simply because we know it will be helpful. By contrast, contemporary machine artifacts do none of this. They function merely as tools, responding as designed, reactive but not proactive. They are unable to engage in true *partnering*. ITL systems need to be more like partners or teammates, and not merely tools.

Each of these core characteristics has received some attention from isolated subsets of the research community. To achieve ITL in future agents, we must find a way to integrate these characteristics into systems. This will not be easy, for at the core of each characteristic and their integration is the challenge of understanding.

The Challenge of Understanding

Perhaps the most important limitation of our contemporary intelligent machines is that they are not capable of understanding with the depth and breadth found in humans. Many impressive accomplishments have been achieved in the cognitive and computational sciences in recent decades. Most of those are best known to isolated subcommunities of researchers toiling away on issues with great scientific merit. A precious few have captured the imagination of the public due to high profile events, demonstrations, and competitions. Algorithmic advances, blazing fast processors, and massive amounts of training data make it possible to show that silicon-based computation can classify objects, learn well-defined games, and answer some types of questions as well as or better than people can. Less well hyped is the characteristic fragility of these systems. When they are wrong, they are often wrong in ways that are surprising and confusing to people. This is because people understand the questions, images, and activities within the broad context of not just a single task, but within the myriad of tasks, experiences, and relationships they develop over time in ways that the algorithms do not.

At least as troubling as the lack of understanding in our most advanced artificially intelligent machine learners is the fact that we often don't understand them. This is certainly true for the general public, who tend to ascribe

From "Interactive Task Learning: Humans, Robots, and Agents Acquiring New Tasks through Natural Interactions," edited by K. A. Gluck and J. E. Laird. Strüngmann Forum Reports, vol. 26, J. R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-03882-9.

assortments of sophisticated humanlike intellectual capacities to computational systems where it is not warranted. It is often also true for the developers of some of our most impressive learning machines. Among those working at the leading edge of science and technology, the issue is not one of unjustified anthropomorphism. Rather, it is the reality of human cognitive limitations running up against complex, hybrid computational systems. The emphasis on powerful learning mechanisms scaled for use on big data sources has abandoned transparency and left even the innovators of these capabilities scratching their heads and asking, "Why and how is it doing that? What did it learn?" Generally those questions can be answered by engaging in some committed digital forensics, but the time and energy required for those analyses far exceed what would be tolerable in the context of ITL. Queryability, explainability, and transparency must be baked into these systems in order to foster natural, efficient understanding.

Finally, in a recursive descent into scientific and technological challenges, as a research community we must face the reality that the root cause of our machines' poor understanding and of our poor understanding of complex learning machines is the fact that we simply do not understand the concept of understanding. There is, in effect, no scientific consensus about what understanding actually is, despite an abundance of work by philosophers, psychologists, neuroscientists, and computer scientists. Indicative of this absence of agreement is a great deal of ambiguity regarding how to assess understanding. This should come as no surprise, given the inconsistent and haphazard manner in which we, as individuals, evaluate the understanding of other people in our daily lives. We tend to assume a great deal of understanding in the minds of others. Sometimes those assumptions are valid and supported by social cues, prior experience, or knowledge of the other, which makes these assumptions defensible. Other times they are simply efficient conveniences. Rarely do we bother to rigorously evaluate the extent to which another person understands.

Up to now, we have been able to overlook our ignorance regarding the fundamental nature of understanding, our poor understanding of complex learning systems, and the absence of understanding in our machines. We have been satisfied with the traditional approach of implementing systems for pre-determined niche purposes. However, the vision for ITL in artificial systems creates a forcing function to address these issues. The development, improvement, and evaluation of understanding in humans, robots, and agents is critical to the creation of ITL.

Moving Forward: A Multidisciplinary Challenge

Clearly, we are enthusiastic about the potential societal benefits that ITL systems could bring. Nonetheless, we appreciate that the challenges are daunting. To even begin, experts from multiple areas of science and technology must be

From "Interactive Task Learning: Humans, Robots, and Agents Acquiring New Tasks through Natural Interactions," edited by K. A. Gluck and J. E. Laird. Strüngmann Forum Reports, vol. 26, J. R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-03882-9.

able to communicate, find common ground, and implement novel capabilities across disciplinary divides.

With the support of the Ernst Strüngmann Forum we sought to initiate a dialogue among experts from robotics, cognitive modeling, computer science, artificial intelligence, and developmental and comparative psychology. This discourse aimed to analyze how humans and artificial agents acquire new tasks through natural interactions as well as to define ITL from various perspectives, in an effort to establish a foundational reference and organizing framework. The results of this multifaceted dialogue are captured in this volume. Organized around the following primary topics, each contribution explores key aspects of ITL:

1. *Knowledge*: In Chapter 3, Robert Wray III et al. discuss the functional roles of knowledge in ITL, examine central challenges that must be overcome, and pose research questions to direct future research. From a formal, computational perspective, Christian Lebiere (Chapter 4) presents different forms of knowledge and skills involved in ITL. Through an examination of the collaborative interactions inherent in learning and teaching, Charles Rich (Chapter 5) analyzes the abstract form, nature, and organization of task knowledge. Concluding this section, Niels Taatgen (Chapter 6) explores what is needed to construct a cognitive architecture capable of supporting flexible knowledge and skills.
2. *Interaction*: In Chapter 7, Andrea Thomaz et al. consider which qualities of human interaction and learning will be most effective and natural to incorporate into an ITL agent; central to this is the alignment of common ground between a teacher agent and a learner agent. In his analysis of natural forms of purposeful interaction among humans, Stephen Levinson (Chapter 8) delineates the basic organization of interactive language use and discusses the challenge of incorporating the predictive nature of human comprehension into an ITL agent. In Chapter 9, Joyce Chai et al. outline the different types of knowledge that can be transferred between agents and discuss the perception, action, and coordination capabilities that enable teaching–learning interactions; in addition, they consider challenges and research opportunities associated with enabling *natural* interaction in artificial agents. To conclude this section, Wayne Gray et al. (Chapter 10) explore how experimental psychology, machine learning, and advanced statistical analyses can be used to understand the complexity of interactive performance in complex tasks involving single or multiple interactive agents in dynamic environments.
3. *Instruction*: In Chapter 11, Julie Shah et al. present frameworks, models, and methods for task instruction, broadly connecting structural and adaptive improvements to instruction, historical developments in programming, and the extraordinary challenge that fluid, flexible,

From "Interactive Task Learning: Humans, Robots, and Agents Acquiring New Tasks through Natural Interactions," edited by K. A. Gluck and J. E. Laird. Strüngmann Forum Reports, vol. 26, J. R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-03882-9.

co-constructive task instruction and learning places on the vision for ITL. In Chapter 12, Kurt VanLehn looks at prototypical human tutoring behavior, analyzing what exceptional tutors sometimes do (but most tutors do not) and comparing the effectiveness of human versus computer tutors. In Chapter 13, Katrin Beuls et al. examine what type of general architecture is needed to construct artificial agents that can assume the role of teacher (by carrying out teaching strategies) or the role of learner (by carrying out learning strategies that benefit from these teaching strategies); they argue that a meta-layer is necessary to understand and implement strategies and point to operational examples in the domain of second language teaching. In Chapter 14, Arthur Still et al. explore the concept of creativity and its relationship to the development of education theory, focusing on what is necessary to inform teaching practice and development of education technology.

4. *Learning*: Summarizing their discussions at the Forum, Dario Salvucci et al. explore in Chapter 15 the learning of task knowledge through interaction, the capabilities that facilitate learning, aspects of interaction that relate closely to learning, as well as evaluation dimensions and metrics for ITL systems. Based on knowledge of preexisting capabilities that appear early in human development, Franklin Chang (Chapter 16) introduces a world-state prediction model—one that can learn detailed physical regularities in the environment and develop representations for predicting the actions and goals of animate agents—to suggest that prediction and prediction error are capabilities that could improve ITL systems. In Chapter 17, using an existing agent, Rosie, to illustrate how an ITL agent can learn many tasks in a variety of domains, John Laird et al. present characteristics of the learning problem and examine how these influence underlying learning algorithms; learning approaches are discussed that respond to the unique challenges of ITL.

Throughout this process, the open exchange of ideas and perspectives—hallmarks of the Ernst Strüngmann Forum—was bolstered by our own inclination toward asking lots of questions, especially the hard ones, and to accept disagreement, countervailing opinions, and inevitable failures on the path of progress. As one might imagine, many questions surfaced and, where appropriate, ways of pursuing these have been highlighted. Two priorities that emerged, however, have been given special attention. The first, a common reference frame to guide future discussion, is presented by Tom Mitchell et al. in Chapter 2. The second is an appreciation that we must commence with ethical considerations now, even as we debate the nature, viability, and path toward ITL. There will be valid security and privacy concerns, and it is certain that people with malicious intent will attempt to repurpose ITL for harm. Thus, now is the time to think about and take action on these matters. To that end, in Chapter 18, Matthias Scheutz examines different ethical aspects of ITL.

From "Interactive Task Learning: Humans, Robots, and Agents Acquiring New Tasks through Natural Interactions," edited by K. A. Gluck and J. E. Laird. Strüngmann Forum Reports, vol. 26, J. R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-03882-9.

Despite all the challenges, we believe ITL offers great potential for humanity and hope this volume will inspire the international research community to pursue the necessary science and technology. We look forward to working with a global community of researchers to realize this vision.

Acknowledgments

We thank the Ernst Strüngmann Foundation for its extraordinary commitment to the exploration of multidisciplinary scientific challenges, and especially for approving and funding this Forum on Interactive Task Learning. However, organizations are only as good as the people within them, and it is the collective contributions of the individuals leading the Ernst Strungmann Forum that make it the world-class, impactful experience that it is. Our sincere appreciation to Julia Lupp, its Director, for her deep immersion and commitment, unwavering support, and impressive patience, and to Aimée Ducey-Gessner, Marina Turner, and Catherine Stephen for their excellent professional support throughout this process.

Participating in an Ernst Strungmann Forum is a significant investment of time and intellectual energy. We thank all of the participants for setting aside their many other existing commitments to join us in this endeavor. Special appreciation is due to our Program Advisory Committee (Ken Ford, Elena Lieven, Julia Lupp, Luc Steels, and Niels Taatgen), who worked with us to shape new ideas and rough intentions into a more complete and concrete plan for the Forum. We must also recognize the impressive work of our rapporteurs (Dario Salvucci, Julie Shah, Andrea Thomaz, and Bob Wray) who toiled diligently throughout and following the Forum to summarize, represent, and organize diverse discussion points into the group reports introducing each section of the book.

Finally, although we were as comprehensive and inclusive as possible, there is no way to include everyone who is doing important and relevant work in a single event such as this. Thus we thank our many additional colleagues and collaborators from the cognitive, computing, social, and psychological sciences who are chipping away at the barriers, making progress on the challenges, and choosing to travel the path toward ITL. You are inspiring, and we look forward to learning with you in future interactions.

From "Interactive Task Learning: Humans, Robots, and Agents Acquiring New Tasks through Natural Interactions," edited by K. A. Gluck and J. E. Laird. Strüngmann Forum Reports, vol. 26, J. R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-03882-9.

From “Interactive Task Learning: Humans, Robots, and Agents Acquiring New Tasks through Natural Interactions,”
edited by K. A. Gluck and J. E. Laird. Strüngmann Forum Reports, vol. 26,
J. R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-03882-9.