Human-artificial-intelligence hybrid learning systems

Seth HERD [a,1], Geoffrey URLAND [b] and Brian MINGUS [a] and Randall O'REILLY [a]

[a] *Department of Psychology and Neuroscience,*
University of Colorado,
345 UCB,
Boulder, Colorado 80309

[b] *Toravner,*
637 S 40th St,
Boulder, Colorado 80305

Abstract.
The only known path to general intelligence is that taken by humans. Adapting elements of this path to achieving artificial general intelligence (AGI) has become a common area of interest. We address the role of human teachers in this process, using the concept of the zone of proximal development (ZPD). We explore the range of possible human-teacher interactions, including those modeled closely on humans, those involving accessing and changing the AGI learners internal representations, and tighter integrations amounting to human-AI hybrid learning system (HAIHLS). In such a system, a human teacher scaffolds an untrained subsystem by producing the outputs desired from a fully trained version. Those outputs both train that subsystem and provide more useful information to the remainder of the cognitive system. This aid enables all subsystems to learn within the context of the richer behavior and cognition possible with the aid of the human subsystem.

Keywords. Motivation, Learning, Neural Network

Introduction

There is only one known path to a generally intelligent system; that taken by humans. An outstanding question in artificial intelligence research is how to capitalize on our knowledge of that path, while taking any available shortcuts. Embodied and brain-mimetic systems have been of increasing interest to AI researchers [1,2], but the path also includes human self-selection of learning experiences (Self-Directed Learning, [3]) and the intervention of human teachers. Here we discuss a number of ways in which a human teacher could help a properly designed machine learning system to achieve general intelligence.

[1]Corresponding Author: E-mail: seth.herd@colorado.edu

The general idea of having humans teach AGI learners is far from new. Many existing systems are trained in a supervised style, using human-labeled data. This obvious application of human teachers is useful, but suffers from inflexibility and time inefficiency (human teachers must spend time before training to classify or be quickly available during training to provide classifications). Even with adequate resources, it is especially impractical for an embodied learner to have individual sense-inputs classified by humans; the system would have to pause to allow humans to classify its inputs (imagine that while navigating a novel landscape an embodied learner has to stop and wait for its human teacher to answer whenever it cannot identify a new object!). But humans can classify a few crucial situations that will be particularly helpful to the learner, if those can be identified by either the learner or its teacher(s).

Human teachers offer undeniable benefits to human learners. Much human teaching involves pushing learning into the Zone of Proximal Development (ZPD), the space of problems that can be solved by a learner when aided by a teacher [4]. We follow the generalization of this proposal, that all types of learning are more effective when a teacher helps a learner to expand their abilities. This assumption is implicitly shared across all (human) educational systems. Existing approaches to human-aided machine learning all constitute some variety of generalized ZPD, in that the human teacher in some way extends the abilities of the machine learner. These include variations based on conditioning (such as reinforcement learning [5], shaping [6], and active learning [7]), demonstration [8], and explicit human direction [9,10].

1. Humans Teaching Artificial Agents

Beyond simply following the model of human teacher-learner interactions, many other routes are available to a human in teaching a machine. We focus here on the idea that humans could aid machines by aiding or actually playing the role of specific subcomponents of their cognitive system. We term this type of interaction Human-Artificial-Intelligence Hybrid Learning Systems (HAIHLS). The human component could either serve to scaffold that system by having it learn from the humans contributions, or simply enable other parts of the system to learn more effectively by pushing it further into a zone of proximal development. We focus here on the example of a human aiding and/or standing in for elements of the reward-prediction element of a machines motivational component, but the idea could be applied to any cognitive subsystem.

Another (non-exclusive) strategy is to have a human directly observe and/or change the learners internal representations and knowledge structures. A properly designed interface could allow a teacher to both know what the learner is thinking, and to influence that thinking much more reliably than is possible with human learners.

Human student-teacher interactions and their adaptation to teaching AGI learners In the human learning environment, teachers fill a variety of roles. Many of these can be applied in a straightforward way to human teachers helping AGI learners [11]. Existing work has also adapted animal training techniques to teaching an artificial agent [13]. We will skip to some less obvious adaptations of human student-teacher relationships.

One subtle role of a human teacher involves gauging when a learner could use information that it does not yet know to ask for. Human instructors can gauge, by gaze and more subtle action patterns, what is currently puzzling to the learner, and supply crucial conceptual information to fill gaps. With AGI learners, a variety of internal variables can be made available to help the teacher gauge what information to supply. The teacher can then provide that information, perhaps in conjunction with guiding the learners attention, through external (gestures or spoken labels) or internal (directly providing sensory inputs and/or guiding its sensory apparatus to objects of importance).

The Socratic method, in which a teacher asks questions that are carefully chosen to clarify a learners conceptual framework, also has a direct applicability to teaching machines. The teacher can focus the machines learning efforts on crucial questions (e.g., what are you? or what is important?) and through further questioning guide the learner to the desired outcome while helping the learner build a conceptual framework for itself.

The potential of accessing and affecting a machine's internal representations in real time affords a number of variants on human student-teacher interactions. These could prove useful alternatives to either a pure programming or pure learning approach to AGI development. We focus here on perhaps the most extreme variant: treating a human as one component of a cognitive architecture.

2. Human-Artificial-Intelligence Hybrid Learning Systems

Progressing from narrow AI to artificial general intelligence (AGI) will require the integration of many cognitive subsystems into a functional whole. Little work to date has addressed the new challenges inherent in doing so. One novel application of Vygotsky's concept of scaffolding a learner into a zone of proximal development is to stand in for cognitive systems that are relatively less capable. We call a system incorporating human teachers as cognitive subsystems a Human-Artificial Intelligence Hybrid Learning System (HAIHLS).

This approach has two advantages: first, it allows the other subsystems to learn in the context of a more complete, highly functional whole. As such, that system can learn in a context more like its intended functional setting. Second, a human working as a "subsystems" can serve to train its replacement.

In this approach, an untrained machine learning subcomponent rides along and learns from how the human performs their computational role, in the context of the whole, functioning system. As machine learning systems become more reliable and flexible, this approach could circumvent the need for carefully engineering systems, and even circumvent the need for understanding what training signals and learning criteria are used by the analogous system in the human brain. For instance, rather than discovering that human infants are intrinsically motivated to move and so learn by motor babbling, [14], the system could be trained directly to make productive motor movements by learning based on control signals are sent to its actuators by a skilled and goal-aware human. This type of learning goes far beyond reinforcement learning, by providing a rich (vector) training signal appropriate to a cognitive and sensory situation.

2.1. Example: Human as Object Recognition System

A human observer could receive the visual input received by an embodied AGI learner (in some partly-filtered form) and classify the object. Instead of doing so blindly, the observer could track goal and conceptual representations to provide the most useful possible classification. A banana could be classified as fruit, food, or a toy depending on the systems current questions and concerns. This context-sensitive human object classification would then serve to train the object-recognition system, allowing it to eventually give relevant classifications in similar contexts.

The object recognition subsystem is thus trained in precisely the information environment in which it must perform. The HAIHLSs behavior can be as rich as though it had a fully functioning visual system, and the information supplied from other cognitive subsystems is precisely as it would be once the human is removed from the system. Similarly, the other cognitive subsystems can learn in the context of a fully functional object recognition subsystem; while the behavior of the trained machine subsystem will not be identical to that of the human subsystem, it should be similar enough to provide substantial learning advantages.

2.2. Example: Human as Reward Prediction System

We work within the Intelligent Adaptive Curiosity (IAC) framework [15] and a biologically-motivated adaptation of the same ideas as Self-Directed Learning (SDL) [3]. The reward prediction system in SDL, or the prediction-prediction system within the IAC framework, serves to direct the learner to spend time in areas of behavioral space in which learning is relatively rapid. This is the essence of self-directed learning: learn about what you can learn about, do not waste time on what is, at least currently, unlearnable. In SDL, success at an arbitrary task is rewarding, as it is for humans [22]. The system seeks to keep doing things with rapidly increasing (or possibly just intermediate) levels of reward prediction; anything that never supplies reward is frustrating, while anything that always provides success and therefore reward is boring. In IAC, a rapidly increasing predictability of behavior plays the same role; no change in predictability indicates that the task is either currently unlearnable, or already well-learned. To engineer a complex, successful self-directed learner from the ground up, we would need to discover or deduce what is intrinsically motivating to human learners, and so causes them to choose learning situations adequate to enable general intelligence. However, if a human reward predictor directs the nascent machine learning prediction system toward useful learning situations, this need can be bypassed entirely. For instance, the human subsystem might predict reward whenever the machine is looking at a human whose eyes move toward the machine, but only when the currently active goal is getting attention. In this situation, we need not know what a human infant finds rewarding about getting human attention; the human trains the system, merely by using their own skills and knowledge of what high-level situations humans do find rewarding. The machine reward-prediction system can then generalize from its own sensory apparatus, and over time develop a suitable representation of what sensory information signals human attention. Human attention will, after that learning, activate the reward prediction system (and thereby the dopamine reward signal) triggering the system to both learn, and to remain attentionally tied to that situation in hopes of more learning opportunities [3].

Inversely, having a human as part of the motivational system could train the system away from useless or dangerous behavioral domains. Energetically banging an actuator into a wall or body could simply be marked by a button press as no fun, and so ceased in favor of new learning. By looking at the relevant goal and context representations, the human as a reward prediction system can perform a much more useful role than simply providing reinforcement signals based on some inflexible criteria (e.g., pain).

These two examples of a human as part of a larger system should illustrate the possibilities inherent in such hybrid systems. A human serving this role is very much acting to scaffold the agents skills, as discussed in expansions on Vygotskys ZPD framework [17,18]. The human functions very much as a scaffolding does for a building; it supports the structure as it is built, and is removed when the edifice is complete and can stand on its own.

3. Caveats and conclusions

It bears more than a casual mention that the sort of self-directed learning system we expand upon here has drawn severe criticism, for reasons that we find highly convincing [19,20]. There is a real possibility that a successful AGI learning system will achieve its goals very successfully. By this logic, the motivational structure of our AGIs becomes extremely important. And even the seemingly benign goal of constant learning implicit in both the IAC and SDL approaches could prove disastrous when taken to the extreme.

In sum, it seems likely that some variant of using humans as subsystems that can stand in for and train parts of an AGI learners cognitive architecture will prove useful. We have suggested a range of ideas; we now await specific implementations to provide specificity and see which approaches bear the most fruit.

References

[1] M. C. Anderson, Rethinking interference theory: Executive control and the mechanisms of forgetting, Journal of Memory and Language 49 (2003) 415–445.

[2] Brunette, E.S., Flemmer, R.C., Flemmer, C.L.: A Review of Artificial Intelligence. In: Proceedings of the Fourth International Conference on Autonomous Robots and Agents, (2009) 385–392

[3] Herd, S.A., Mingus, B., O'Reilly, R.C.: Dopamine and Self-directed Learning. In: Biologically Inspired Cognitive Architectures 2010: Proceedings of the First Annual Meeting of the BICA Society,(2010) 58–63. IOS Press, Fairfax, VA

[4] L. S. Vygotsky, Mind in society: The development of higher psychological processes, Harvard University Press, Cambridge, MA, 1978.

[5] A. L. Thomaz, C. Breazeal, Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance.
[6] Knox, W.B., Fasel, I., Stone, P.: Design Principles for Creating Human- shapable Agents. In: AAAI Spring 2009 Symposium on Agents that Learn from Human Teachers (2009)
[7] Settles, B.: Active Learning Literature Survey. Technical Report, University of WisconsinMadison (2009)
[8] A. Y. Ng, Reinforcement learning and apprenticeship learning for robotic control. In: Balczar, J., Long, P., Stephan, F. (eds.) Algorithmic Learning Theory. LNCS, vol. 4264 (2006) 29–31. Springer, Heidelberg
[9] Chernova, S., Veloso, M.: Teaching Multi-Robot Coordination Using Demonstration of Communication and State Sharing. In: Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS), 1183–1186 (2008)
[10] Thomaz A.L., Breazeal, C.: Experiments in socially guided exploration: lessons learned in building robots that learn with and without human teachers. Connect. Sci., 20 (2008) 91–110
[11] Thomaz, A.L., Cakmak, M.: Learning About Objects with Human Teachers. In: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction. (2009) 15–22. ACM, New York, NY
[12] Zhu, X., Goldberg, A.: Introduction to Semi-Supervised Learning. In: Brachman, R., Dietterich, T. (eds.) Synthesis Lectures on Artificial Intelligence and Machine Learning, 3 (2009) 1–130. Morgan & Claypool Publishers, San Rafael, CA
[13] Goertzel, B., Pennachin, C., Geissweiller, N., Looks, M., Senna, A., Silva, W., Heljakka, A., Lopes, C.: An Integrative Methodology for Teaching Embodied Non-Linguistic Agents, Applied to Virtual Animals in Second Life. In: Proceedings of the First Artificial General Intelligence Conference (2008) IOS Press, Amsterdam, The Netherlands
[14] Demiris, Y., Dearden, A.: From Motor Babbling to Hierarchical Learning by Imitation: A Robot Developmental Pathway. In: Proceedings of the Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems (2005) 31–37
[15] Oudeyer, P.Y., Baranes, A., Kaplan, F.: Intrinsically Motivated Exploration for Developmental and Active Sensorimotor Learning. In: Sigaud, O., Peters, J. (eds.) From Motor Learning to Interaction Learning in Robots: Studies in Computational Intelligence (2010) 264, pp. 107–146. Springer, Heidelberg
[16] Aron, A.R., Shohamy, D., Clark, J., Myers, C., Gluck, M.A., Poldrack R.A.: Human Midbrain Sensitivity to Cognitive Feedback and Uncertainty During Classification Learning. J. Neurophis., 92, 1144–1152 (2004)
[17] Conner, D. B., Cross, D. R.: Longitudinal Analysis of the Presence, Efficacy and Stability of Maternal Scaffolding During Informal Problem-Solving Interactions. B. J. Dev. Psych., 21 (2003) 315–334
[18] Kermani, H., Brenner, M.E.: Maternal Scaffolding in the Child's Zone of Proximal Development across Tasks: Cross-Cultural Perspectives. J. Rsch. Chil. Ed., 15, (2000) 30-52
[19] Mijic, R. Bootstrapping Safe AGI Goal Systems: CEV and Variants Thereof. In: Proceedings of the Third Conference on Artificial General Intelligence. Atlantis Press (2010)
[20] Yudkowsky, E.: Artificial Intelligence as a Positive and Negative Factor in Global Risk. In: Bostrom, N., Cirkovic, M. (eds.) Global Catastrophic Risks. Oxford University Press, New Yoprk, NY (2008)