Expanding Human Student-Teacher Interactions to Create Human-AI Hybrid Learning Systems (HAIHLS)

Seth Herd¹, Geoffrey R. Urland², Brian Mingus¹, Heather Underwood¹, and Randy O’Reilly¹

¹Computational Cognitive Neuroscience Lab
Department of Psychology
University of Colorado at Boulder
Muenzinger D244, 345 UCB
Boulder, Co, 80309, USA
²Toravner
637 South 40th St.
Boulder, CO, 80305, USA
{seth.herd, brian.mingus, heather.underwood, randy.oreilly}@colorado.edu
http://grey.colorado.edu/ccnlab

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 proximal zone of development (ZPD). We explore the range of possible human-teacher interactions, including those modeled closely on humans, those involving accessing and changing the AGI learner’s 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: learning, teacher, subsystems, cognitive architecture

1 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][3], but the path also includes human self-selection of learning experiences (“Self-Directed Learning,” [10]) 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.

The 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. 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 [22]. We follow the generalization of this proposal, that all types of learning are more effective when a teacher helps and directs a learner. 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 and directs the abilities of the machine learner. These include variations based on conditioning (such as reinforcement learning [19], shaping [13], and active learning [17]), demonstration [14], and explicit human direction [5] [20].

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 guiding 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 serve to scaffold that system by having it learn from the human’s contributions, and 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 examples of a human aiding and/or standing in for elements of the reward-prediction element of a machine’s motivational component, and as a replacement and trainer for the object recognition system, 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 learner’s 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.

The remainder of the paper proceeds as follows. We first broadly classify
some styles of human student-learner interactions, and suggest that these could be applied fairly directly to interactions between a human teacher and an appropriate AI learner. Next, we explore some additional avenues of instruction that are not available with human learners; these involve a more direct interaction with the machine’s internal representations. Finally, we address the possibility of a true HAIHLS system, using examples to flesh out the core idea.

2 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 [21]. We start by briefly listing some examples, including some more subtle teaching methods.

Perhaps the most obvious role of a human teacher is to provide answers when a learner needs them. In some cases, the learner asks for help. This approach is used (or perhaps over-used) in the variety of schemes for supervised learning, especially active learning [17]. Hybrid (or semi-supervised) approaches, in which the learner bootstraps from a limited number of classified examples, expand on this model [24]. The ability of these approaches to ask for classification on a few problematic examples leverages the contributions of human teachers with limited time.

A more 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 learner’s attention, through external (gestures or spoken labels) or internal (directly providing sensory inputs and/or guiding its sensory apparatus to objects of importance).

Other methods of teaching are also applicable to instructing machines. Learning by example (or “apprenticeship learning”) allows learners to understand a skill simply by watching the teacher perform it and has already been applied to robotic navigators [15]. Applying this technique to a broader range of skills will require careful engineering of the learner’s ability to imitate a human instructor. It can also be combined with other learning techniques (such as reinforcement learning) to create a more powerful multi-stage learning process [9].

The Socratic method, in which a teacher asks questions that are carefully chosen to clarify a learner’s conceptual framework, also has a direct
applicability to teaching machines. The teacher can focus the machine’s 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.

This style of teaching relates to another of Vygotsky’s ideas about the development of intelligence: the importance of language [22]). He suggested with regard to humans, and it has been proposed [6][18] with regard to artificial general intelligence, that language provides conceptual representations, crucial to our level of intelligence. However, it has been suggested [4] that existing natural language processing approaches are inadequate for developing rich representations that can support a general intelligence, and that embodiment provides a necessary semantic anchor to the physical world. Interaction with human teachers may be particularly important in developing language representations with rich semantic underpinnings. A human instructor can ask questions and supply concepts that emphasize different and contextually specific meanings of an individual word.

3 Self-Directed Learning

While direction by human teachers is important, its benefits may be greatest when it is applied to self-directed learning. A framework called Intelligent Adaptive Curiosity has been proposed and implemented by Oudeyer and colleagues [16][11]. This setup drives the learner to zones in which its actions are becoming more predictable, indicating rapid learning and the potential for even more rapid gains.

We have previously discussed [10] a possible role of the dopamine system in performing a similar function in guiding human attention to situations that provide new learning. In brief, the human reward system is known to direct learning toward imperfectly predicted rewards. This function ensures that learning is directed toward new topics over time, rather than remaining focused on highly predictable, already learned predictors of (and opportunities to gain) reward. In addition, the successful performance of arbitrary laboratory tasks is known to activate the human reward system. Taken together, these facts suggest that humans self-direct their learning toward situations that provide new successes at an arbitrary “task,” as conceptualized by the learner.

A teacher providing direction a self-directed learning system could provide a massive efficiency boost in teacher time: the learner then merely explores behavioral space surrounding a novel goal provided by the instructor.

4 Beyond Human Teaching: Changing AGI Learner’s Minds

All of the above teaching styles are attempts to usefully adapt human student-
teacher interactions to human teacher and AGI learner interactions. But there are more possible ways to interact with artificial systems than with closed-brain biologicals. A teacher may directly access and influence the learner’s internal representations in a variety of ways.

Accessing sensory representations offers little benefit beyond a teacher’s ability to monitor a student’s gaze. But a larger benefit could come from accessing a machine’s representation of context and goals. For instance, knowing that an embodied AGI learner is “hungry” for power and looking at its power dock would elicit a very different teacher response from seeing that gaze while knowing that the machine is fully-powered and has a goal of solving another task.

Directly influencing a machine’s sensory representations allows an expansion on a human teacher’s ability to draw attention to relevant items. For instance, the AGI learner in the above example could be directly presented with an image of its power dock (when it is “hungry”) without physically moving its position or cameras.

Directly influencing deeper representations, those of contexts, goals, or reward values, starts to edge across the (blurry) line between human teaching and a hybrid human-AI system. When the human clicks the ‘reward’ button while enforcing an internal representation of the power dock, they are teaching the reward system to expect reward in conjunction with that image. In this particular case, the human teacher serves to “shape” [13] the system’s behavior, as a human might do by administering food or praise to an animal during training. The direct interaction with internal representations expands on the animal shaping case by allowing a more reliable means to know to what representations that reward signal “attaches.” Further, to the extent that the “reward” button press affects the current behavior of the system, the human is, to a limited extent, acting as part of the cognitive system. A more direct integration of the human teacher into the cognitive system, treating the human as an actual subsystem, is also possible.

5 Human-Artificial-Intelligence Hybrid Learning Systems

Having humans substitute for part of a cognitive architecture sounds like an admission of failure to successfully engineer that subsystem. One common, but subtle example of substituting human intelligence for working subsystems is the tendency to engineer systems to solve very specific problems. In this case, the programmer is effectively acting as the surrounding cognitive system, supplying information in a usable form, and supplying motivation that would in a full system come from other subsystems. This approach allows just one subsystem of interest to be built in detail. As we progress from narrow AI to AGI, cognitive systems need to consist of more working subsystems.
Despite the (justifiable) criticisms of building a system specifically for one limited task, (and so avoiding the need for many other subsystems), working on some subsystems while postponing progress on others has obvious benefits. One alternative to severely limiting the task to be approached (and hand-coding for success at that task) is to incorporate humans into a cognitive architecture as an online, (near) real-time component. This approach can enable the benefits of an embodied, explorative agent, while avoiding the need to simultaneously engineer all of the necessary subsystems.

In addition to allowing some components of the system to function as if part of a more complete, highly functional set, this approach can serve to train the missing piece(s). When a learning approach is possible, the untrained 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,” [8], 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.

5.1 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 system’s 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 HAIHLS’s 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.

5.2 Example: Human as Reward Prediction System

Roughly the same ideas apply within the Intelligent Adaptive Curiosity (IAC) [16] and Self-Directed Learning (SDL) [10] frameworks. We will address both. 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 [2]. 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 [10].

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 agent’s skills, as discussed in expansions on Vygotsky’s ZPD framework [7][12]. 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.

6 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 [14][23]. There is no good reason to suppose that a learning system that achieves a human level of intelligence will stop or even pause there before becoming at least weakly superhuman. It seems that we should accept at least the possibility that a successful AGI learning system will do whatever it wants. 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.

Our conclusion is that, while our specific examples may prove to be flawed, that some aspects of using humans as to stand in for and train parts of AGI learner’s cognitive architecture will likely 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

Netherlands (2008)


19. Thomaz, A.L., Breazeal, C. Reinforcement Learning with Human Teachers:


