Exploring Mixed-Initiative Interaction for Learning with Situated Instruction in Cognitive Agents

Shiwali Mohan and John E. Laird
Computer Science and Engineering, University of Michigan
{shiwali, laird}@umich.edu

Human-agent interaction for learning with instruction can be viewed on a continuum of instructor/agent control. At one extreme are systems that learn by instructor-driven interactions, such as learning by demonstration, examples, or imitation. The other extreme of the continuum is occupied by systems where instructor interaction is limited to responding to the questions posed by the agent or to providing feedback on agent's performance.

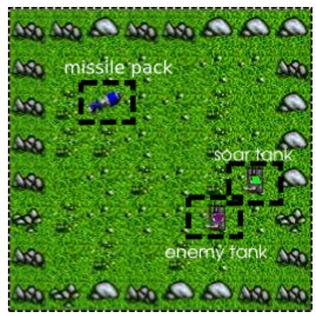
There are advantages to an approach that explores a mixed-initiative instructional dialog. First, an agent that can assume control can guide its own learning by requesting clarifications, asking for missing information, and correcting instructor's response based on its own understanding of the state. Second, an agent that can relinquish control can take advantage of the instructor's knowledge of task structuring and goal decomposition. In a mixed control setting, the instructor can ask the agent for information regarding its state and the environment, verify an agent's learning by questioning the agent, and provide corrections.

To be able to maintain the state of interactions with the instructor while acting in the environment, and to be able to learn from these instructions in the context they were provided in, the agent needs a model of task-oriented interaction. Such a model is required to support the properties described below.

1. Both the instructor and the agent can assume control of the interactions at any time.
2. The interaction model provides a context for instructor's elicitation, allowing the agent to take relevant actions.
3. The interactions by the agent should be informed by agent's reasoning, learning and acting mechanisms.
4. The interaction model and the sequence of interactions should inform agent's learning.

A Motivating Example

To demonstrate our implementation of the interaction model and how it informs agent's learning, we use a scenario in a grid-world type domain, TankSoar (Figure 1). The figure also describes the state representation used and actions known to the agent. The sequence of actions/subtasks required to successfully complete the required task depends on whether the agent is carrying a missile. If the agent is carrying a missile, the task execution would involve pointing the tank in at the enemy tank (`align-tank`), and firing the missile. However, if the agent is not carrying a missile, the task involves picking up the missile (`pick-up-missile`), aligning the tank and firing. The environment is *partially observable* to the instructor and the task is *unknown* to the agent, necessitating mixed initiative, bi-directional information transfer.




```
World state:
missile:(<mx, my>)
enemytank:(<ex,ey>)
tank:(<x,y> <n/s/e/w>)
Agent state:
has_missile: no/yes
Known actions/subtasks:
Actions: move, turn,
fire-missile
Tasks: pick-up-missile,
align-tank
```

Figure 1: TankSoar Domain

Agent Design

Our agents are instantiated in Soar (Laird, 2008), a symbolic, cognitive architecture based on the problem-space hypothesis. A Soar agent's current state is derived from its perceptions, its beliefs about the world and knowledge in its long-term memories and is held in its working memory. Behaviors and action in Soar are represented as production rules in its procedural memory. Soar's episodic memory (Derbinsky and Laird, 2009) is a context dependent memory that records the agent's experience during its lifetime by taking snapshots of working memory and storing them in chronological fashion.

Interaction Model

The interaction model we describe has been adapted from Rich and Sidner (1998). It captures the state of task-oriented interaction between the agent and the instructor. To formalize the state of interaction, we introduce (1) *events* that change the state of interaction; these include dialog utterances, actions and learning, (2) *segments* that establish a relationship between contiguous events, and (3) a *focus-stack* that represents the current focus of interaction.

In accordance with the discourse interpretation algorithm described by Rich and Sidner (1998), each event is changes the focus-stack by, (i) starting a new segment whose purpose contributes to the current purpose (and thus, pushing a new segment with a related purpose on the focus stack), (ii) continuing the current segment by contributing to the current purpose, (iii) completing the current purpose (and thus eventually popping the focus stack) or (iv) starting a new segment whose purpose does not contribute to the current purpose (and thus pushing a new, interrupting segment on the focus-stack, changing the purpose of the interaction). An event contributes to a segment, if (i) it directly achieves the purpose, and (ii) it is a step in achieving the purpose.

Applying the Interaction Model to TankSoar

If the agent does not know which tasks can be executed in the environment, it initiates an interaction with the instructor by pushing a segment (s1) on the focus-stack. The purpose of this segment is to acquire the next task to undertake. The instructor replies with a task, attack-tank. The purpose of segment (s1) is achieved, and it is removed from the focus stack. A new segment (s2) is introduced, whose purpose is for the agent to execute the task in the environment.

To execute it, the agent proposes an operator corresponding to the task. However, since the agent is executing this task the first time, it does not know how to proceed further. An impasse occurs, which causes the agent to introduce another segment (s3) on the focus stack with a purpose to get an example execution of the task. The agent then initiates a segment (s4) to prompt the instructor for a next action to execute. The selection of next action depends on if the agent is carrying a missile. However, the instructor cannot observe what the agent's state it. If the interaction was constrained to be agent-initiated only, the instructor would be forced to take an uninformed guess. In case the instructor suggests an action that is not available to the agent, it would not be able to progress in the task.

In mixed-initiative interaction settings, the instructor initiates a segment (s5) with a purpose to get more information about the state of the agent.

Segment (s5) interrupts the current purpose of interaction by introducing a momentary distraction. The agent replies with information about its state, and removes segment (s5) from the focus-stack, reverting the status of interaction to what it was earlier. The agent continues to prompt the instructor for next action until the instructor indicates that the agent has successfully completed the task. Segment (s3) is removed from the focus-stack. The agent, then, initiates a segment (s6) whose purpose is to inquire about the goal of the task. The instructor replies with the goal predicates, which achieves the purpose of the agent-initiated segment (s6). The agent performs situated explanation through introspective recall, and on arriving at successful explanation removes (s1).

Situated Instructional Learning

The goal of learning with human instruction to be able to acquire general task execution knowledge, such that it is applicable to not only the situations that the instructions were provided in, but also in similar, analogous situations. Huffman and Laird (1995) introduced general task learning by situated explanation in which the agent internally projects the effects of instruction starting from the state the new task was first suggested. If the projection successfully achieves the termination conditions of the new operator the agent learns general task execution rules. In our implementation, the episodic memory of the agent is instrumental in reconstructing the state where the task was first executed in, and the focus-stack is instrumental in projecting the instructed actions in the correct context.

Discussion

Previous work on mixed-initiative instructional learning has assumed that the instructor can completely observe the environment, and therefore, limit the interactions initiated by the instructor (Huffman and Laird 1995); has focused on learning by demonstration (Allen et al. 2007) or has concentrated on acquiring task specification knowledge for non-situated knowledge-bases (Boicu et al. 2001). We are interested in learning task execution knowledge situated in agent's experience in the environment and aided by mixed-initiative, bi-directional transfer of information between the agent and the instructor.

In our representation of the interaction state, *initiative* can be defined in terms of the ability of a participant to introduce segments on the focus stack. Events from both the instructor and the agent cause modifications in the stack, allowing both to introduce segments with specific purposes in the interaction (requirement 1). The focus-stack provide context to instructor elicitation (requirement 2) and allows the agent to progress in the environment. Agent-initiation depends on it detecting its incapability to progress further in the task, signified by an impasse (requirement 3). The focus-stack also facilitates in retrieving the correct episode from episodic memory (requirement 4), allowing the agent to learn generalized task application knowledge.

There are several avenues for further exploration and study. An agent operating in a novel environment is required to not only acquire task-application knowledge, but also to learn recognition and semantic categorization of new objects, spatial relationships between objects and hierarchical task decomposition. We are interested in exploring if the interaction model described here can be used effectively in these learning scenarios

References

Allen, J.; Chambers, N.; Ferguson, G.; and Galescu, L. 2007. Plow: A Collaborative Task Learning Agent. *In Proceedings of AAAI.*

Boicu, M., Tecuci, G., and Stanescu, B. 2001. Mixed-Initiative Agent Teaching and Learning, *In Proceedings of ICAI.*

Derbinsky, N., and Laird, J. 2009. Efficiently Implementing Episodic Memory. In *Proceedings of the 8th ICCBR*.

Huffman, S. B., and Laird, J. E. 1995. Flexibly Instructable Agents. *JAIR*.

Laird, J. 2008. Extending the Soar Cognitive Architecture. In *AGI*.

Rich, C., and Sidner, C. 1998. COLLAGEN: A Collaboration Manager for Software Interface Agents. *User Modeling and User- Adapted Interaction*.