

Injecting Machine Learning into the Apprentice Learner Architecture, Project Milestone Report 1 15-300, Fall 2019

Cayden Codel (ccodel)
Project website: www.contrib.andrew.cmu.edu/~ccodel

December 15, 2019

1 Major Changes

As mentioned in my project proposal, I had a choice of working with either Daniel Weitekamp or Napol Rachatasumrit. Danny's work focuses on authoring, whereas Napol will work on injecting reinforcement learning theory into the way feedback is given to simulated agents, as agents currently receive feedback much more often than humans do.

Because of my reinforcement learning background, I felt I would have a better research experience with Napol, even though he's a first year and has just started on the project. After speaking with Professor Vasilescu, I confirmed my decision, and so I plan on working with Napol next semester. Therefore, some of the goals in the 15-400 milestones no longer apply. See **Revisions** for details.

2 What I've accomplished so far

As per the proposal, my literature review and review of the Apprentice Learner codebase is well underway, and I expect to make good progress over break. I have not implemented my own agent at all, but plan on starting when the spring semester begins.

3 Revisions to the milestones

Because I plan on transitioning to working with Napol, the milestones for my project need to reflect the change in work. Of course, in January, when I begin regularly meeting with Napol and Dr. Koedinger, I'll learn more about where my contribution will lie.

Napol's work centers on TRESTLE, an incremental hierarchical algorithm that helps simulated learning agents develop the *when-learning* mechanisms that tell them when to apply their set of operations to solve a problem. Currently, TRESTLE successfully trains agents, but only when given positive or negative feedback per portion of an answer, rather than the entire answer itself (for example, getting feedback per digit of a multiple-digit multiplication problem). The hope is to improve TRESTLE by incorporating reinforcement learning theory to decrease the amount of feedback necessary and to possibly incorporate delayed feedback mechanisms to allow for learning on batches of problems.

The first four to eight weeks will likely be exploratory, examining how TRESTLE responds to RL-like algorithmic changes. After that time period, we will likely have a feel for what works and what doesn't, and begin to train agents on the new method. Updates will come as the project continues.

4 Resources needed

Currently, I have all of the resources I need, aside from research literature that I will acquire as the need arises.