INTELLIGENT TUTORING SYSTEM

Final Project Proposal

Georgia Tech

Spring 2021

Instructors: Prof. James McClellan & Greg Krudysz

Hazel John
hjohn8@gatech.edu
gtid: 903355078

Project Description: Adaptive ITS (ITS-RL)

The focus of this project will be to identify optimal learning strategies that can be incorporated into the Intelligent Tutoring System. Currently, the system displays random questions from the collections associated with the concept to the student, however a more intelligent model would be able to adapt to the students' needs by understanding and serving questions based on their level of mastery of the concept. Reinforcement learning is uniquely suited to this as it can not only suggest the optimal policy for the current steps, but also for subsequent steps leading to best longer-term reward.
What have some data around the question (question type, associated concept, and some tags). We also student data like the duration, score and rating. We currently don't have sequence data for the questions, so we can't currently infer if the questions reviewed in the past had any impact on the student performance for the question being attempted. However, in most learning scenarios, we would want to student to progress from fundamental concepts to more advanced concepts, we can use this inherent assumption for incorporating some sequential relationships between questions both within book chapters and across chapters.

The current design is to model the system as a Markov Decision Process. Modeling the question progression in a way to allow intelligent policy creation will be key. The MDP model will have the following.

State: Each question along with context (chapter) and type (multiple choice, matching, short answer …) will form the basis for a state representation. One option to prevent presenting the students with the same question multiple times could be by storing completed questions as part of the state.
Action: The actions could be failure and multiple success levels based on duration and score (split into intervals).
Reward: Reward can be modeled using the score potentially assigning negative score instead of zero for failure.
Transition probability: Transition probability can be inferred from the distribution of scores and duration, along with some domain knowledge about textbook and chapter/section interdependencies). If the state is modeled not to include completed questions, the transition probability will change dynamically as we would modify transition probabilities into a completed question to zero on completion. However explicit transition probability is only required for a model-based algorithm. For model-free algorithm, I would still utilize domain knowledge to identify list of potential next states but will not need to explicitly define transition probability.

For the initial model, I will use a model free approach, implementing Q-learning along with epsilon decay to ensure sufficient exploration in the beginning and more exploitation as the learning progresses.

A maximum review time or number of steps constraint along with higher rewards for more complex chapters will be added to ensure that the optimal policy learnt by the system will take the students through all the concepts. Otherwise, it is possible that the learnt policy would just have the students remain with simple questions and attain top rewards.

Project Plan

1. **Vision:** Adaptive Learning based on student performance

2. **Areas of Exploration:** State modeling, State transition modeling, Reward shaping, Reinforcement learning algorithms (Model based (policy/value iteration), or model free (Q-Learning).

3. **Assumptions:** Duration taken to answer question and score will be a good reflection of student's performance in the next logically sequenced concept.

4. **Plan:** Start with the simplest model and iterate.

5. **GitHub Repository:** https://github.gatech.edu/VIP-ITS/IRS-RL. Once the initial structure is created, I expect to make weekly check-ins (at minimum) to visibly track progress.

6. **Tools/Technology:** Python, IPython Notebook, Optional: OpenAI Gym (if time allows) though Gym comes with a set of toy environments, creating an environment from scratch might not be the best use of time, I can do what I need with just Python.

7. **Potential Issues/Pitfalls:** The number of states in the system will be large. Visual representation of the policy will not be possible. I also don't have sufficient domain knowledge about the book to incorporate domain specific relationships for inter-concepts transitions, instead I will model a simple serial progression of chapters.

   In order to get the best value from Reinforcement learning, it is important to model the system in a way to best answer the problem to be solved. Designing the state/actions/reward/transitions can require multiple iterations and are part of the plan.

8. **Policy Validation:** A/B testing of the user experience with the intelligent policy vs random selection will used to evaluate performance and validate correctness/efficacy of the adaptive policy.

9. **What is not covered:** Integrating the intelligent policy into the ITS system.

10. **Common Topics with Other Projects:** The database layer and associated functionality can be built as a common resource that can be used across the different projects (TutorJS and Chatbot). I will try to partner with the backend teams to see if a common module or API can be built so as to be used by all teams.

Group Membership

Name	Skills	Tools	Contact information

Hazel John	Python, Database, Machine Learning	OpenAI Gym, Python Notebook (visualization of policy), MySQL	hjohn8@gatech.edu

Milestones

Milestone name	Target date	Status
Project Plan Created	Feb 5th, 2021	On Track
Environment Setup Completed	Feb 12th, 2021	TBD
Finalization of initial model	Feb 15th, 2021	TBD
Initial check in with stubbed Code	Feb 22nd, 2021	TBD
Initial model/policy creation completion	March 15th, 2021	TBD
A/B Testing of the model performance	March 22nd, 2021	TBD
Policy Storage Schema Design	March 29th, 2021	TBD
Database Creation/Update	April 5th, 2021	TBD
Timeboxed improvements	April 19th, 2021	TBD
Final Code Improvements and Checkins	April 26th, 2021	TBD
Project Cleanup and final documentation	April 26th, 2021	TBD
Project Presentation	Before End of Semester	TBD

Resources

Reinforcement Learning Lectures:

1. OMSCS Udacity Course with Dr. Isbell and Dr. Littman

   https://www.udacity.com/course/reinforcement-learning--ud600 (Dr. Isbell and Dr. Littman)

2. DeepMind's David Silver lectures

   Introduction to Reinforcement learning:
   https://www.youtube.com/watch?v=2pWv7GOvuf0&list=PLqYmG7hTraZBiG_XpjnPrSNw-1XQaM_gB&index=2&t=0s

   Markov Decision Processes:
   https://www.youtube.com/watch?v=lfHX2hHRMVQ&list=PLqYmG7hTraZBiG_XpjnPrSNw-1XQaM_gB&index=2
   Exploration and Exploitation:
   https://www.youtube.com/watch?v=sGuiWX07sKw&list=PLqYmG7hTraZBiG_XpjnPrSNw-1XQaM_gB&index=9

3. Other Tutorials

   Q-Learning: https://deeplizard.com/learn/video/qhRNvCVVJaA

Exploration vs Exploitation: https://deeplizard.com/learn/video/mo96Nqlo1L8

Some of the papers relevant to this topic include:

1. AgentX: Using Reinforcement Learning to Improve the Effectiveness of Intelligent Tutoring Systems: Kimberly N. Martin and Ivon Arroyo
2. Learning teaching strategies in an Adaptive and Intelligent Educational System through Reinforcement Learning Ana Iglesias · Paloma Martínez · Ricardo Aler · Fernando Fernández

OpenAI Gym exploration:

https://www.oreilly.com/radar/introduction-to-reinforcement-learning-and-openai-gym/