A Natural Language Argumentation Interface for Explanation Generation in Markov Decision Processes

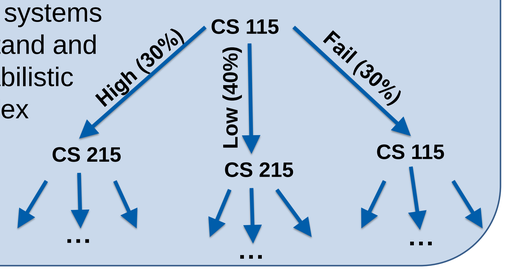
Thomas Dodson, **Nicholas Mattei**, and **Judy Goldsmith**
University of Kentucky, Lexington, KY, USA

Objective:

- Build a robust, scalable, real-time system that convinces the user to take the action recommended by the optimal policy.
- Leverage human psychology and behavioral decision making studies to select the most convincing set of arguments from the available information in the model, policy, and case-base.
- Design the system from the ground up to be portable between application domains and incrementally upgradeable.

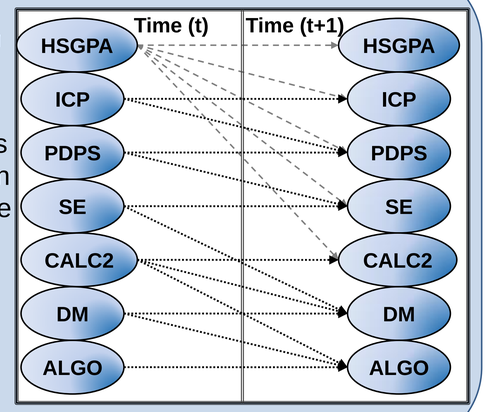
Motivation:

- Markov Decision Processes (MDPs) are a mathematical formalism to used to reason optimally in uncertain domains.
- Policies created from MDPs are complex objects that are hard to understand by direct examination.
- Existing MDP-based systems assume users understand and can reason with probabilistic information and complex policies.



Decision Model:

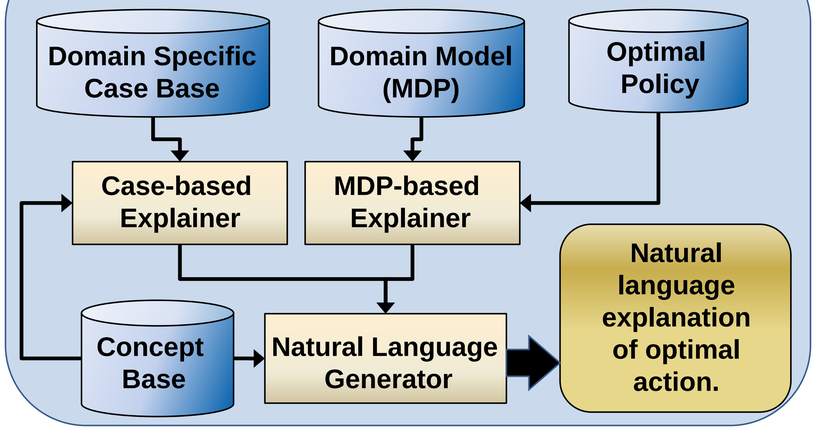
- Uses MDPs to reason in the domain of academic advising.
- Each action corresponds to taking a class in a given semester and each course has a conditional probability of success.
- Uses optimal policy, mapping from states to actions that maximizes discounted future reward.



Case-Based Explainer:

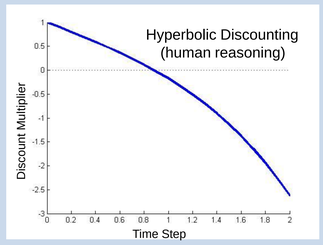
- Uses corpus of past student transcript data to generate convincing arguments from similar student cases.
- Leverages human tendency to reason by analogy through providing examples of past students in similar situations (similar transcripts).
- Highlights differences between the naive distribution over the results of the recommended action and distributions conditioned on current state.
- Optional component, is used as additional evidence and is not required for the rest of the system to function.

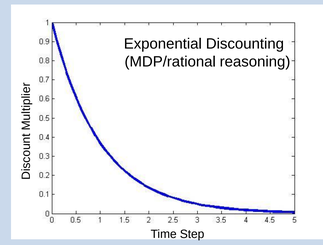
System Overview:



MDP-Based Explainer:

- Uses information present in the reasoning model and policy to convince user to take the optimal action
- Justifies recommendation with arguments over the short term future benefits (hyperbolic discounting).

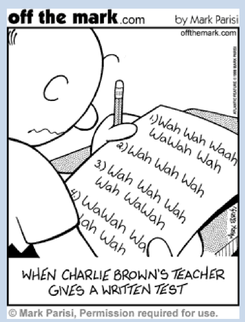




Natural Language Generation:

- Uses common verb phrases for actions that are appropriate for target audience.
- Uses common nouns for state variables.
- Uses adverb phrases such as “very likely” instead of “greater than 85% chance.”

This helps make probabilistic reasoning more accessible to target users.



Example Output:

The recommended action is taking *Introduction to Program Design and Problem Solving*, generated by examining possible future courses. It is the optimal course with regards to your current grades and the courses available to you. Our model indicates that this action will best prepare you for taking *Introduction to Software Engineering* and taking *Discrete Mathematics* in the future. Additionally, it will prepare you for taking *Algorithm Design and Analysis*. Our database indicates that with either a grade of A or B in *Introductory Computer Programming* or a grade of A or B in *Calculus II*, you are more likely to receive a grade of A or B in *Introduction to Program Design and Problem Solving*, the recommended course.

Conclusions and Future Work:

- System is operational and works through a text based interface.
- Users interact with the system in real time; as they define their courses and grades text explanations are created using all the elements of the system.
- We have recently received IRB approval for a multi-stage study looking at advising from the expert (advisor) point of view as well as the student (user) point of view.
- We hope to learn what makes for useful advice in our domain and compare the quality of our “most convincing” heuristic against the “most coverage” heuristics of other systems.

•This work is supported by NSF EAGER grant CCF-1049360