Learning the Mental Map of a Fuzzy Decision-Maker

Andrew Buck and James Keller

Computational Intelligence Research Lab – University of Missouri

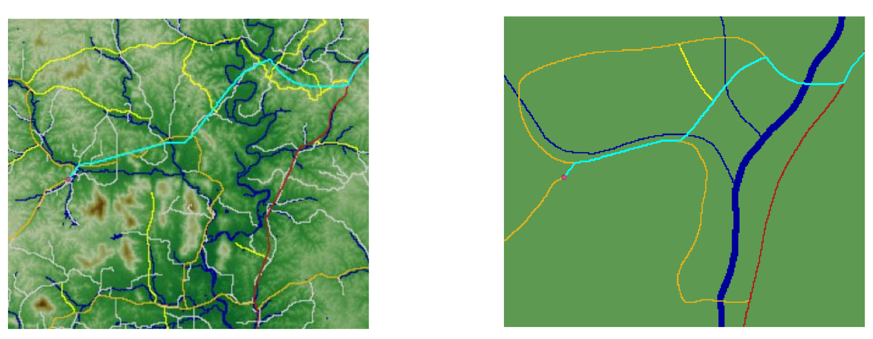
Abstract

We can represent many real-world tasks such as navigation through an unknown environment as sequential decision processes through some state space. Decision-makers in these domains are faced with the problem of acting on incomplete information in order to maximize some reward or minimize the total cost. We model this problem using fuzzy weighted graphs and the principle of bounded rationality to infer how a decision-making agent might form a "mental map" of the environment and develop a policy. By observing the actions of the agent, we can gain insight into the agent's beliefs and desires, allowing us to understand what the agent's mental map looks like. We draw upon recent work in the field of inverse reinforcement learning and show several examples demonstrating these ideas, including some novel visualizations and preliminary experimental results for this ongoing work.

Visualizing Mental Maps

What does a person's mental map look like?
Why is one route preferred over another?

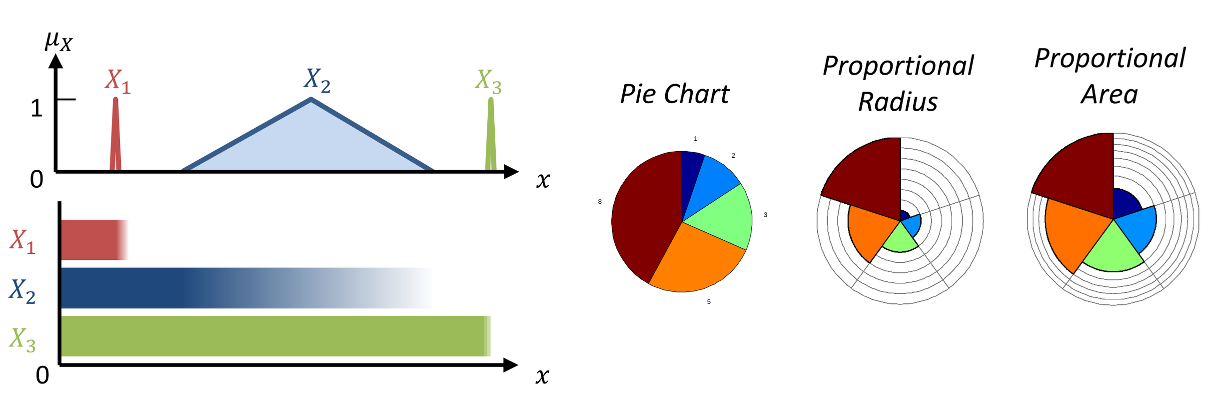
Humans are expert decision-makers, but it's not always clear what is being optimized. We use bounded rationality to model human decision-makers.



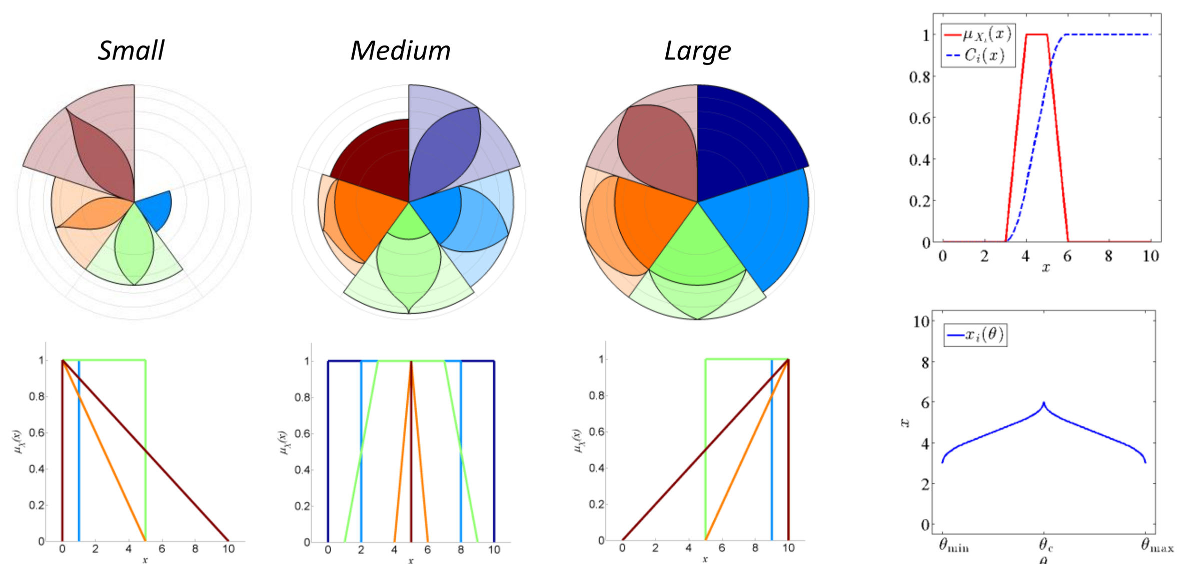
Ground Truth Mental Map

Fuzzy Rose Diagrams

How can we visualize vectors of fuzzy numbers or graphs with imprecise weights? Graphics should be easy and intuitive to interpret, without misrepresenting the data. The principle of perceptual proportionality suggests that the amount of ink or color used to indicate a value should be proportional to its magnitude. We use a fuzzy weighted graph, which has vectors of fuzzy numbers attributed to the nodes and/or edges, to model the mental map of a decision-making agent.

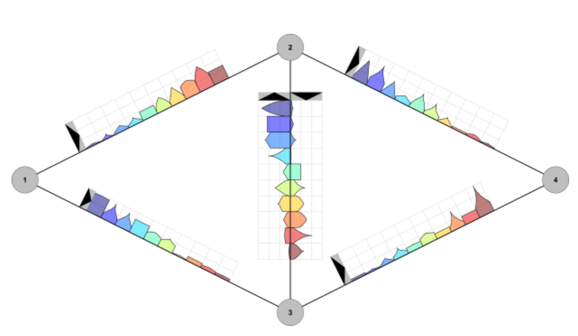


Petal shape is defined by the cumulative function.



Fuzzy Weighted Graph

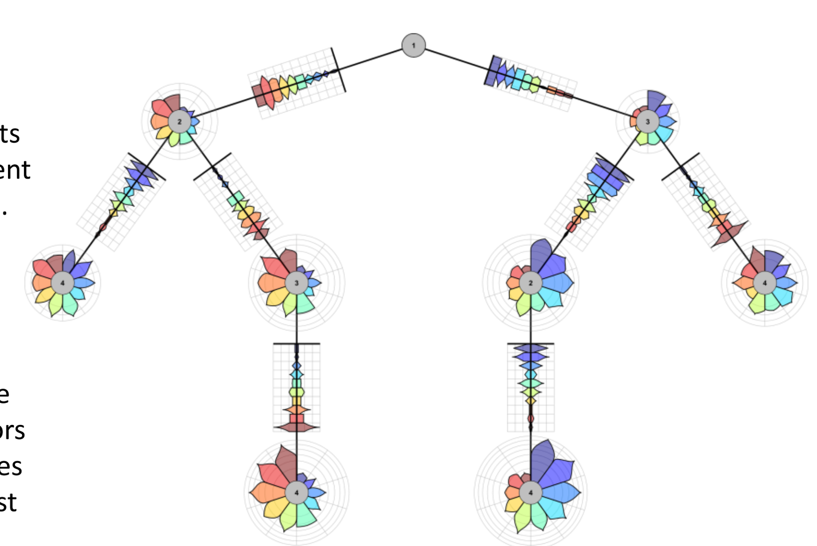
A navigation problem might be to plan a route from node 1 to node 4.



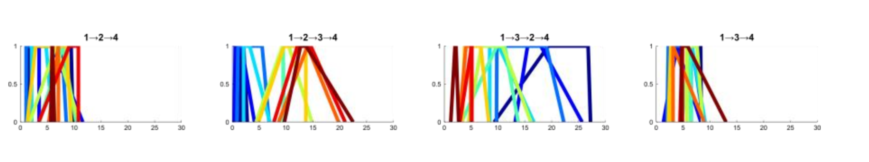
Decision Tree

Each leaf node represents the total cost for a different path through the graph.

The agent compares the aggregated feature vectors for each path and chooses the one with the greatest reward or least cost.

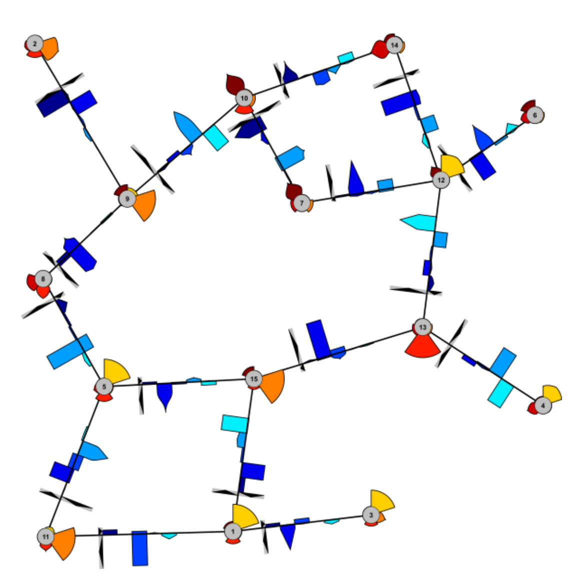


Aggregated Path Features



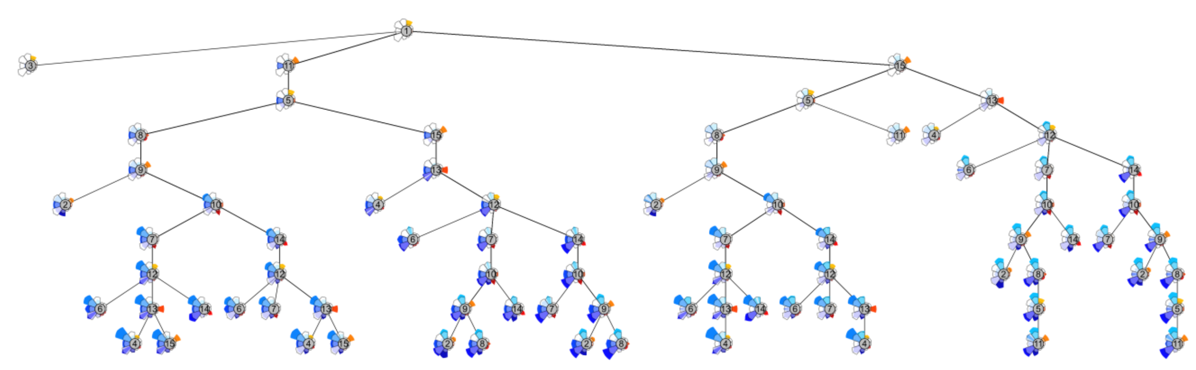
Stochastic Decision-Making

Mental map as a fuzzy weighted graph. Nodes have rewards (red) and edges have costs (blue).

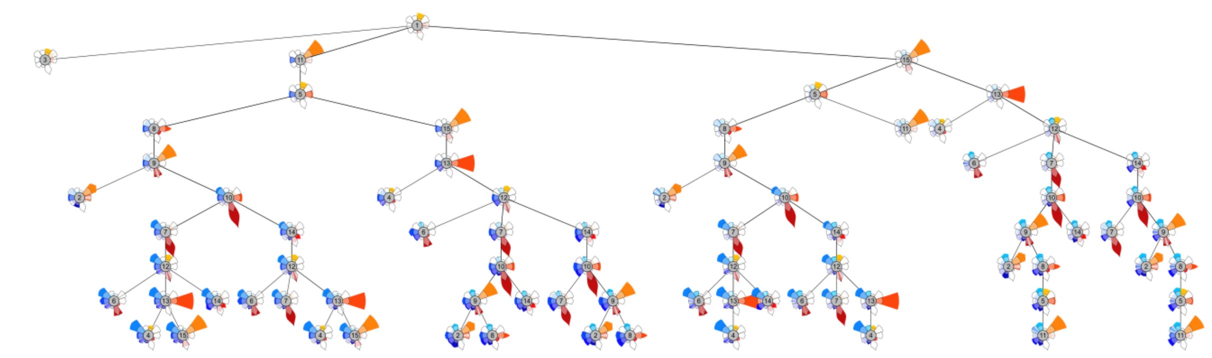


All possible destination/path pairs from the agent's current location.

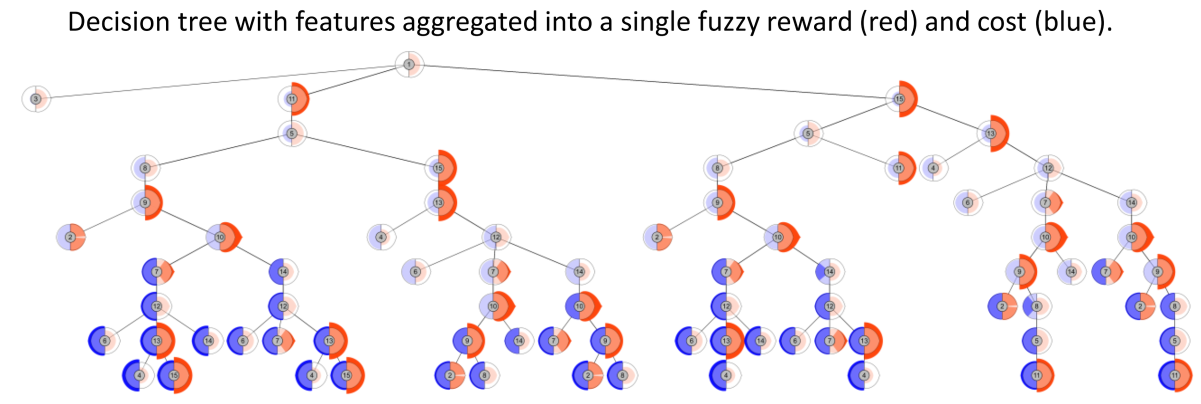
Unbiased decision tree. Each node represents a unique destination and path.



Decision tree biased according to the agent's preferences.

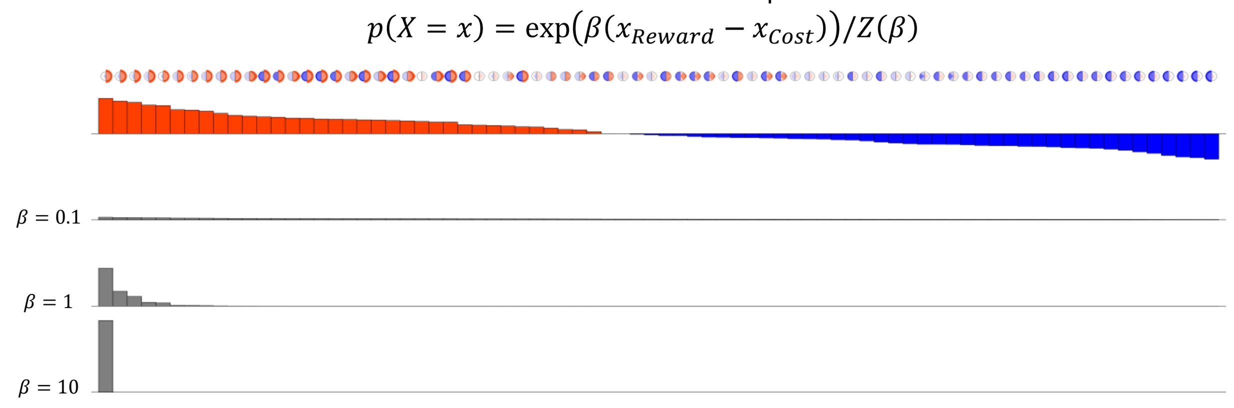


Decision tree with features aggregated into a single fuzzy reward (red) and cost (blue).



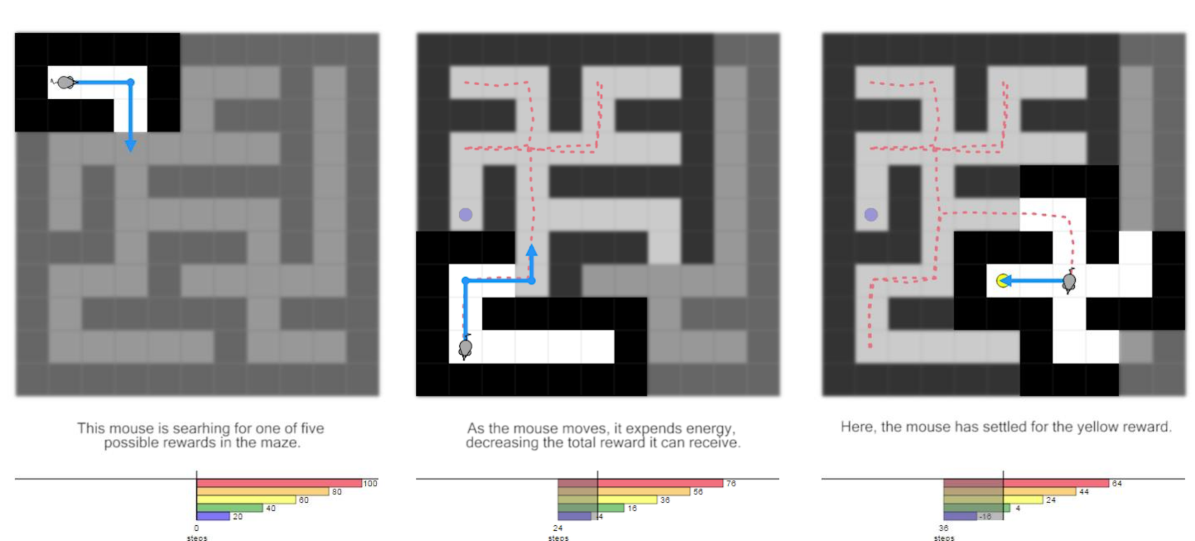
A Gibbs distribution is used to sample a decision.

$$p(X = x) = \exp\left(\beta(x_{Reward} - x_{Cost})\right)/Z(\beta)$$



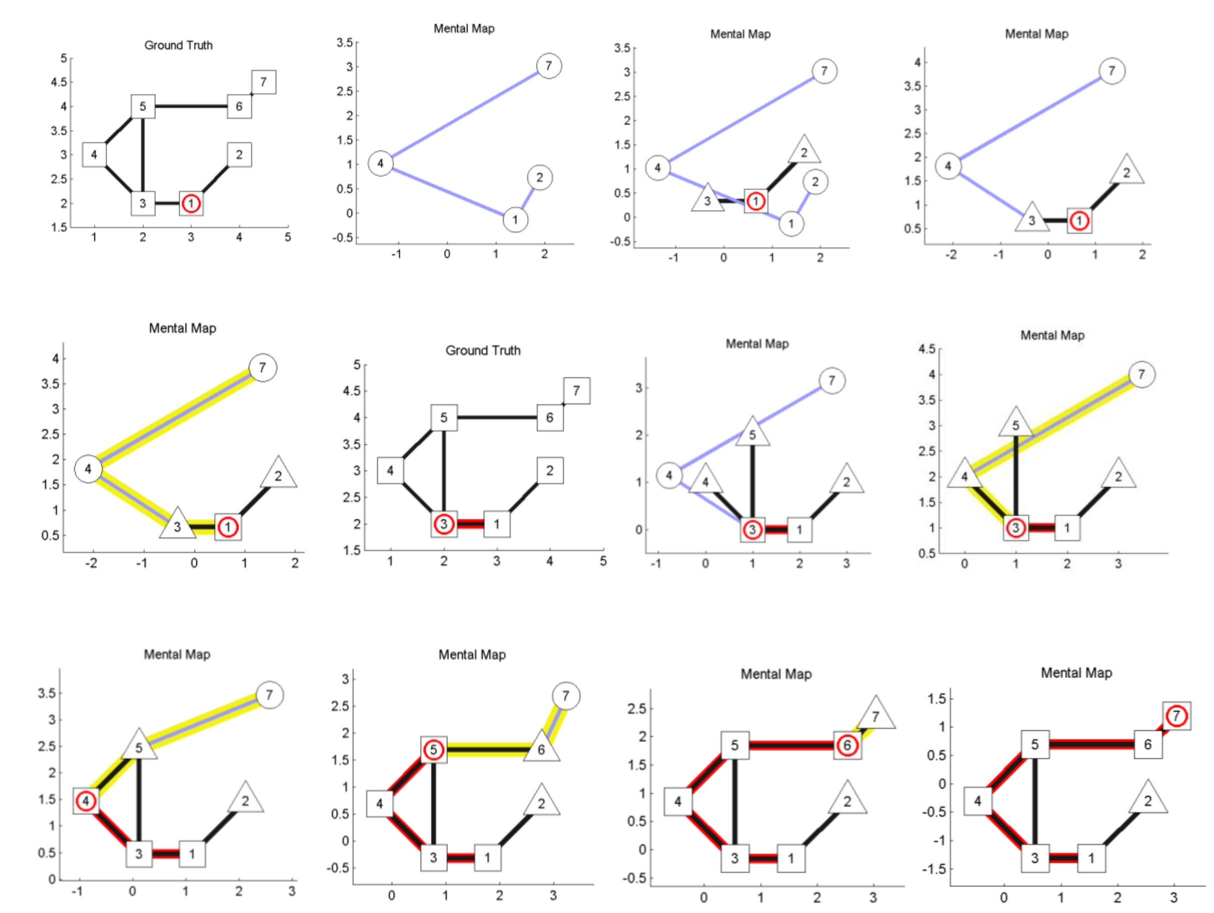
Bayesian Theory of Mind

Can we infer an agent's reward structure and mental map just by observing its actions in the environment?



Conflation with Ground Truth

An agent's mental map may not be an accurate representation of the environment. As the agent moves through the environment, it observes new ground truth and conflates this new information with its existing mental map. If the ground truth is significantly different from the agent's expectation, the agent creates a new plan using the information in the updated mental map.



Inverse Reinforcement Learning

Inverse Reinforcement Learning is defined as follows:

Given

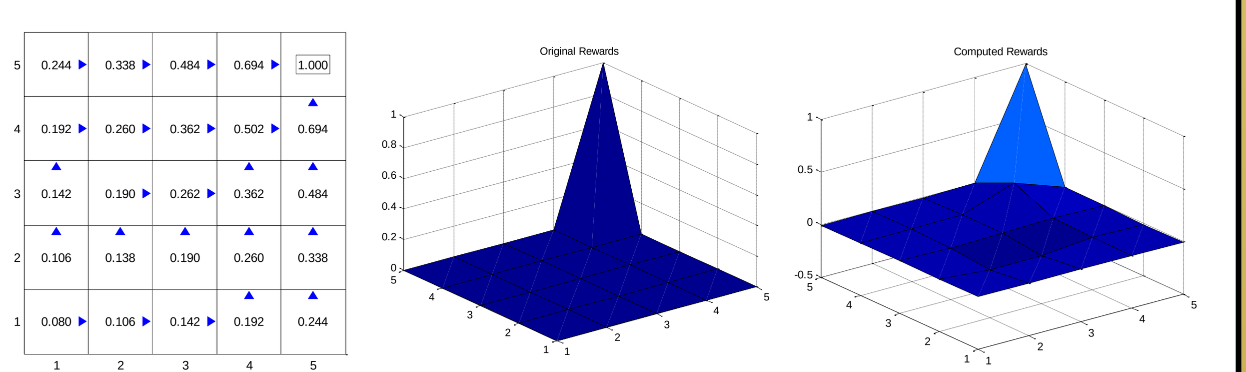
1. Measurements of an agent's behavior over time, in a variety of circumstances
2. If needed, measurements of the sensory inputs to that agent
3. If available, a model of the environment

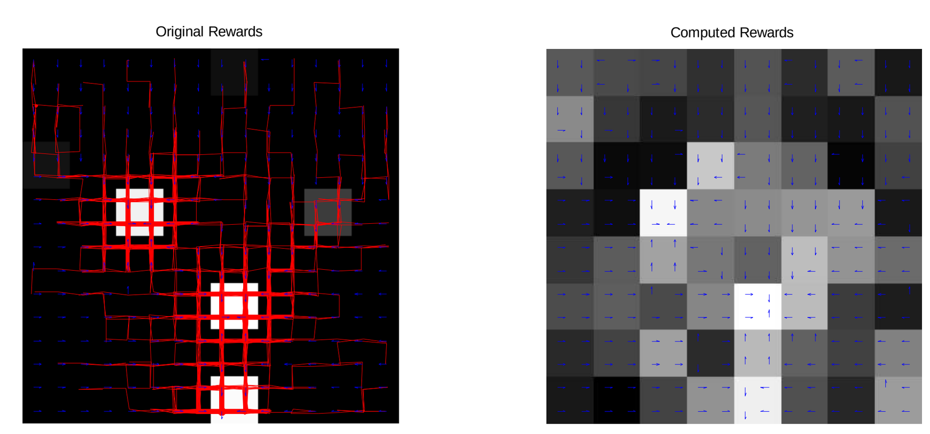
Determine

The reward function being optimized

Experimental setup:

1. Generate a gridworld environment and place rewards.
2. Compute an optimal policy.
3. Given only the policy and the transition function, or observed trajectories through the environment, determine the reward being optimized.
4. The reward may be a function of the features defined for each state.





References

[1] A. R. Buck and J. M. Keller, "Visualizing uncertainty with fuzzy rose diagrams," in Computational Intelligence for Engineering Solutions (CIES), 2014 IEEE Symposium on, 2014, pp. 30–36.

[2] A. R. Buck, J. M. Keller, and M. Popescu, "An alpha-level OWA implementation of bounded rationality for fuzzy route selection," in World Conference on Soft Computing, 2013.

[3] C. L. Baker and J. B. Tenenbaum, "Modeling human plan recognition using Bayesian theory of mind," in *Plan, activity and intent recognition: Theory and practice*, Morgan Kaufmann, 2014.

[4] A. Ng and S. Russell, "Algorithms for inverse reinforcement learning," in 17th International Conference on Machine Learning (ICML), 2000, pp. 663-670.

[5] P. Abbeel and A. Ng, "Apprenticeship learning via inverse reinforcement learning," in 21st International Conference on Machine Learning (ICML), 2004.