Mouselab-MDP: A new paradigm for tracing how people plan

Frederick Callaway, Falk Lieder, Paul M. Krueger, and Thomas L. Griffiths

University of California, Berkeley

Motivation

Planning is a latent cognitive process that cannot be observed directly. This makes it difficult to study how people plan. To address this challenge, we propose a new paradigm that provides experimenters with a timecourse of participant attention to information in the task environment.

Background

Many problems that people face require planning many steps into the future. Brute force search quickly becomes computationally intractable, thus people must rely on **approximate planning strategies**. Characterizing these strategies is difficult becaues we cannot directly observe the search process. For example, although Huys et al. (2015) successfully modeled a **pruning** strategy as discounting rewards after large penalties, they could not provide a detailed process model based on beahvioral data alone.

Process tracing paradigms such as think-aloud protocols and eye tracking externalize some aspect of a cognitive process. Payne, Bettman, and Johnson (1988) developed one such methodology for studying multi-alternative risky choice: In the "Mouselab" paradigm, a participant clicks on the cells of a matrix to inform a betting decision.

Usage

This CofeeScript code generates Fig. 1a

```
trial =
  type: 'mouselab-mdp'  # use the jspsych plugin
  graph:  # defines transition and reward functions
    B:
      up: [5, 'A']
      down: [-5, 'C']
    A: {}
    C: {}
  layout:  # defines position of states
    A: [1, 1]
    B: [1, 2]
    C: [1, 3]
  initial: 'B'  # initial state of player
  stateLabels: {A: 'A', B: 'B', C: 'C'}
  stateDisplay: 'always'  # options: never, hover, click, always
  edgeLabels: 'reward'  # can also be an arbitrary mapping
  edgeDisplay: 'click'  # options: never, hover, click, always
  edgeClickCost: 1  # subtracted from score every time an edge is clicked
```

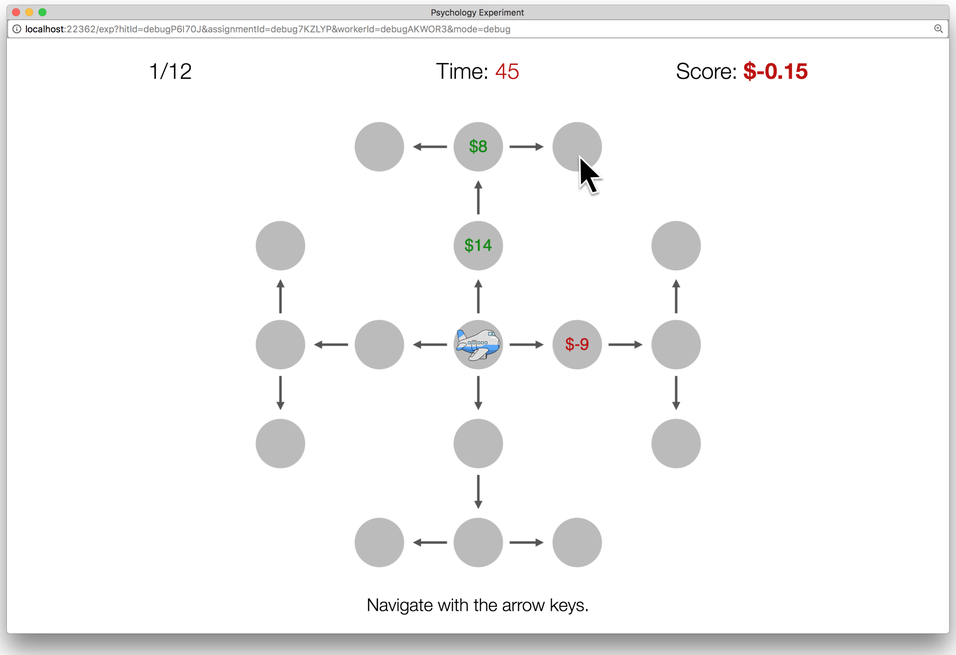

Figure 1: (a) example in **Usage**; (b) a programatically generated layout; (c) rewards displayed on hover.

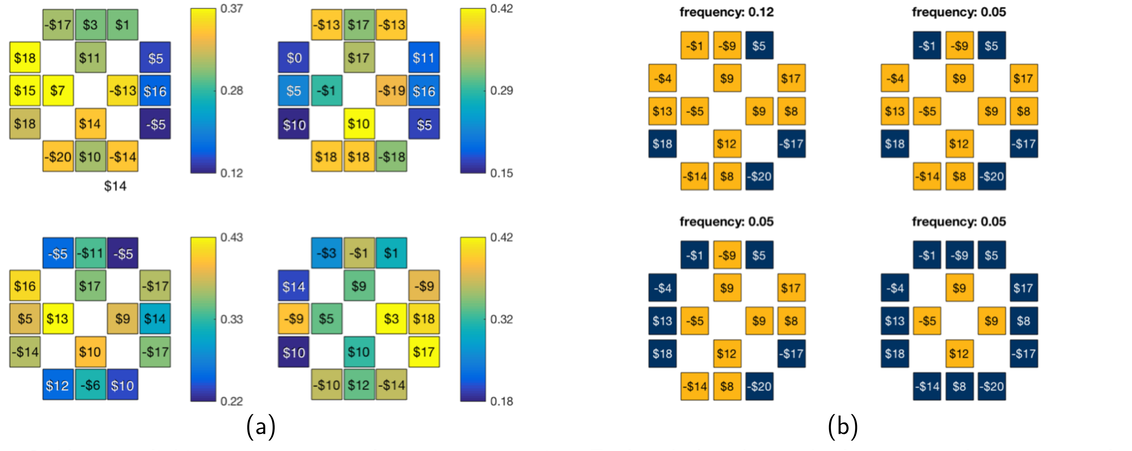

Figure 2: Human clicking patterns reveal pruning strategies. Each subplot shows the location and monetary value of each state. (a) Colors indicate how often each state was inspected before the first move, for four MDPs. b) The four most common click sets for one MDP. Inspected states are gold.

Experiment

To demonstrate the paradigm in action, we designed a layout that encourages pruning (center image). 31 participants completed the online experiment. We set `stateLabels` to be the rewards associated with the edge leading to each state. We set `stateDisplay` to 'click' and `stateClickCost` to 0.10; thus participants could click on a state to reveal the reward for traveling to that state, at the price of $0.10. Participants were required to spend at least 45 seconds on every trial to prevent time cost from discouraging participants from clicking and planning.

Our results provide **direct evidence of pruning**: participants were less likely to click on states after large rewards (Fig. 2). This effect **varies smoothly with reward** (Fig. 3).

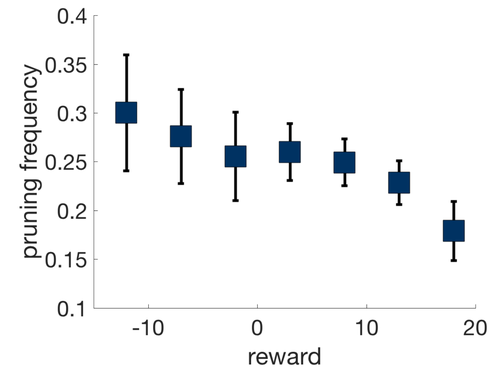

Figure 3: Frequency of pruning ±1SEM given the reward at the stem of the branch (horizontal axis). Pruning was defined as inspecting none of a branch's outer states (before the second move) after having inspected the reward at its stem (before the first move).

References

Huys, Q. J., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., ... Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, *112*(10), 3098–3103.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 534.