A normative theory of approach-avoidance conflicts during dynamic foraging in humans

Arthur Guez1, Ritwik Niyogi2, Dominik Bach3, Marc Guitart-Masip4, Raymond J. Dolan1, Peter Dayan1
1 Gatsby Computational Neuroscience Unit, UCL 2 Zurich University Hospital for Psychiatry 3 Karolinska Institutet 4 Welcome Trust Centre for Neuroimaging, UCL

Introduction

We propose a normative model of the behaviour of human subjects playing a dynamic foraging game containing a time-stochastic threat. The game is intended to capture the essence of the conflict between approach and avoidance. The realistic nature of the task makes planning challenging; we therefore rely on recent innovations in model-based methods to approximate the optimal policy, and on Approximate Bayesian Computation to fit our models.

The Task

Transform an animal paradigm to study approach-avoidance issues in humans.

- Human player is foraging for tokens on a 24x16 landscape grid.
- Tokens move randomly to different locations at regular intervals.
- Captured tokens are valuable at the end of the game (approach motivation).
- Sleeping robber wakes up at random & chases player (avoidance motivation).
- Player can only escape the robber at the safe place. Loses all tokens if caught.
- 3 threat level conditions (low, med, high) correspond to prob. of robber waking up.

Methodology: Modeling

Computational Level

- Discrete episodic MDP. Each time-step corresponds to 200ms, episode \( \leq 15s \).
- State: position of agent/robber, positions of tokens, wake-up state of robber, token tally, time. Assume transition model known.
- Reward function: 1 for each token, \( -\tau \cdot \beta \) for getting caught.

Algorithmic Level

- Huge combinatorial state space and stochastic transitions.
- Optimal policy not computationally tractable \( \implies \) look for approximations.
- Many approximations perform well, but they are not all good match for the data!
- Consider heuristic planning and variants of model-based, forward-search, planning algorithms:

- Greedy Heuristic
- Monte-Carlo Tree Search (MCTS)

Monte-Carlo Tree Search with macro steps (MCTS-MS). Plan using MCTS using macro actions. One macro action to go to each of the token and a macro action to return to safe place. A macro action is interrupted if robber wakes up or if target token disappears.

Monte-Carlo Tree Search (MCTS): Planning 1 macro step ahead followed by a value estimate. Values are learned using TD(\( \delta \)) in a linear architecture. Features include distances (to robber, safe place, tokens) and timing information.

Methodology: Fitting

With a complex model and task, we cannot directly compute \( P(\text{data}|\text{model}) \). Instead we rely on a likelihood-free method for model estimation:

Approximate Bayesian Computation (ABC)

- Use form of approximate rejection sampling.
- Define \( m = \text{model}, \theta = \text{params}, D = \text{data} \).
- Want posterior \( P(m, \theta | D) \).
- Use features \( \phi = f(D) \) as summary for data.

Features (summary statistics)

- Commonly used features in animal literature on risk approach-avoidance plus feature revealing about planning mechanism (see right figure). Greediness: went for nearest token / collected tokens.

ABC rejection sampling alg.

1. Sample from prior \( m, \theta \sim P(m, \theta) \).
2. Simulate with \( m, \theta \) to obtain \( D \).
3. Compute features from simulation: \( \hat{\phi} = f(D) \).
4. Accept sample if \( \epsilon = ||\phi - \hat{\phi}|| < \epsilon \).

Methodology: Modeling (Cont.)

Models: MCTS, MCTS-MS+GreedyRollouts

1. Sample from prior \( m, \theta \sim P(m, \theta) \).
2. Simulate with \( m, \theta \) to obtain \( D \).
3. Compute features from simulation: \( \hat{\phi} = f(D) \).
4. Accept sample if \( \epsilon = ||\phi - \hat{\phi}|| < \epsilon \).

Preliminary Results

Different planning models/parameters supported by data for the different group/threat levels.

Risk and Behavior

Risk affects behavior in different ways. Multiple systems at play:
- Modify/bias loss function (e.g. be sensitive to variance).
- Pavlovian responding, pre-encoded behaviors (PAG).
- Hippocampal lesions in rodents have some anxiolytic characteristics. Associated with threat level in this task.

References


This work was funded by the Gatsby Charitable Foundation and the Natural Sciences and Engineering Research Council of Canada.

Contact: aueurz@gatsby.ucl.ac.uk