Project Specification for exploration in foraging

Background

Humans and other animals make innumerable stay-or-leave decisions, from choosing to leave one shop for another, switching TV programs, or seeking berries to eat. These are foraging decisions, to stay and consume or to leave and look for better resources elsewhere. To get a better understanding of how we make these decisions, researchers of decision-making are increasingly turning to foraging paradigms for their experiments..

Ecologists' classic Marginal Value Theorem (MVT) gives the optimal condition for when an agent should leave a patch to forage elsewhere and maximise its average reward intake. The optimum is to leave the patch when the agent's estimate of its reward rate in the patch falls to its estimate of the average reward rate for the whole environment. This decision rule is deterministic and based on perfect information; by contrast, humans and animals make stochastic choices based on imperfect information. Consequently, this theorem is not a biologically plausible model of the brain mechanisms underlying decisions to leave, nor can it offer insight into why people may vary in their choices.

This project will study if a model that makes stochastic choices can provide a better account of how humans and other animals make stay-or-leave decisions. In particular, such stay-or-leave decisions are suboptimal in humans and other animals, as they stay longer than optimal in one location as its resources deplete. We will test the hypothesis that this "overharvesting" is driven by the trade-off between exploiting known resources and the risk of exploring the unknown elsewhere. We will then generate predictions for how a foraging agent will behave under different foraging problems and under different forms of stochastic choice.

Patch-foraging notes:
Patches are typically modelled as some having a resource that decays with each harvest; the initial yield of the resource on entering the patch defines the patch "type". Agent's job is simply to harvest and decide when to leave: on leaving, it arrives at the next patch. A collection of patch types define the environment, which has a defined average reward rate (reward per unit time).

Aims

1. Understand how patch leaving decision depends on stochastic choice to leave: when and how does an agent overharvest?
2. Predict how an agent will behave under different stochastic selection models (and verify predictions on data)
3. Predict how an agent will behave under different world models: resource depletion models (exp decay, linear decay, late-breaking drop in r(t))
4. [Stretch] Determine if there are order effects of the environment

5. [Stretch] Understand how agents choose with imperfect information - Directed exploration

Steps

1. Implement a simple patch-foraging model (draft notes on model to be shared for implementation)
    a. Specify a world*:* *P* patch types in environment *E*; *N* patches per environment, sequence randomly drawn from the *P* types
    b. Patch reward model: exponential decay
    c. Agent's choice model: softmax with intercept
2. Run agents on the world model:
    a. Compute normative expected leaving time for each patch (assuming MVT): compute background reward rate for environment
    b. Compute agents' performance:
        i. Mean leaving time: E(Leave)
        ii. Variation in leaving time VAR(Leave)
3. Test implementation vs MVT predictions for E(leave):
    a. Does increasing patch yield = longer staying time?
    b. Does a richer environment = shorter staying time?
    c. Does the agent overharvest? If not, what would make it overharvest?
4. Predict effects of stochastic action selection algorithm:
    a. Show e-greedy cannot be what foragers use
    b. Try mellowmax: https://arxiv.org/abs/1612.05628
    c. Find other random selection algorithms (https://openreview.net/forum?id=wzzrs5QH5k) ?
5. [Stretch] Check predictions of different algorithms on foraging data
    a. Analyse basic properties of LeHeron task foraging data (expected leaving time, VAR(leave), changes over patch and environment etc)
    b. Fit basic softmax algorithm to LeHeron data per participant (options here include minimising error in E(leave) and VAR(leave))
    c. Fit other selection algorithms to LeHeron data
    d. Compare fits - which model is best for each participant? (e.g. AIC/BIC; or cross-validation)
6. Predict how foragers would behave under different world models: resource depletion in patches
    a. Linear depletion
    b. Late-breaking depletion

7. [Stretch] Predict order effects: poor->rich vs rich->poor. What does model predict will happen for different orders of experienced environments?
    a. Implement an agent that adapts exploration over time:
        i. Computes average reward rate
        ii. Converts that to exploration

    b. Study how it behaves as environments are switched: What effects on E(leave) and VAR(Leave)? How long do the effects last?
    c.
8. [Stretch] Predictions for directed exploration: what if the agent is uncertain about its reward?
    a. Sketch models for how the agent might represent uncertainty. For example
        i. Uncertainty in foreground and/or background reward rate
        ii. Assumes decaying reward, but uncertainty in the decay parameter: so sample from distribution to get current estimate of foreground rate, then use deterministic rule (MVT). i.e. **question is:** does uncertainty in estimate of input propagate to decision to result in stochastic choice? (i.e. can we capture stochasticity of behaviour using this?)

Notes on potential further points:

Using value as input to foraging decision:
https://www.biorxiv.org/content/10.1101/2024.07.08.602539v1.full

What patch parameters (yield, decay) lead to immediate rejection of that patch? (leave time of 0)

Agent using change in reward as basis for decision?

Suggested reading **(* essential project reading)**

Constantino, Sara M. / Daw, Nathaniel D.
Learning the opportunity cost of time in a patch-foraging task. (2015)
Cognitive, affective & behavioral neuroscience, 15, 837-853

* Le Heron, Campbell / Kolling, Nils / Plant, Olivia / Kienast, Annika / Janska, Rebecca / Ang, Yuen-Siang / Fallon, Sean / Husain, Masud / Apps, Matthew A. J. (2020)
Dopamine Modulates Dynamic Decision-Making during Foraging.
J Neurosci, 40, 5273-5282

Mobbs, Dean / Trimmer, Pete C. / Blumstein, Daniel T. / Dayan, Peter
Foraging for foundations in decision neuroscience: insights from ethology (2018)
Nature reviews. Neuroscience, 19, 419-427

Schulz, Eric / Gershman, Samuel J.
The algorithmic architecture of exploration in the human brain. (2019)
Current opinion in neurobiology, 55, 7-14