links: Research projects MOC

#project #research #RL #max-entropy #DRA

DRA vs Max Ent RL

To do

- ◐ Narrative
  - ☑ Write a brief sketch of the process vs normative models
- ◐ Simulation results
  - ☑ Gather all existing task schematics and results
  - ◐ Write a short discussion
- ◐ Theory
  - ☑ Go through the derivation again
  - □ Compare and contrast with others, esp. SVPG
- □ Implications for neuroscientists
  - □ TBD

Memory resource allocation & entropy regularization in RL as two sides of the same coin

Key questions

1. Are they behaviorally identical?
2. Can we cast them in the same framework to analyze the differences analytically?
3. What are the implications for neuroscientists?

Threads to explore and hopefully merge

Narrative

- Process vs normative modeling
- Encoding noisy values + sampling them to act vs acting according to a heuristic
  - In the theoretical analyses, we will show how the heuristic process model is equivalent to encoding "soft" values and acting greedily wrt these

Behavior

- Simulations DRA vs max-entropy RL
- MaxEntRL frequency vs stakes
  - Intuition behind why MaxEntRL yields the same behavior as DRA
- differences DRA maxEnt

Theory

- Shining the Inference lamp on DRA:
    - Luigi Alex DRA as inference
    - Regularization in RL
- DRA analytical gradient for 2 options
    - Can we use this to flesh out the differences in a toy task? Is it worth the effort?

Implications for neuroscientists

TBD, but 2nd point of the narrative does go into it a little bit.

- When decoding neural activity, what should we look for?
- should we look for "soft" values or real values or something completely different?
- We see neurons that correlate with values in a lot of places all over the brain, but except for perceptual decision-making, we don't really see signatures of evidence accumulation or values being encoded somewhere in the brain.
- Suhaimi, ..., Makino 2022 find neurons encoding values/policy in mouse PPC. They use Advantage Actor Critic with an entropy term in its objective, making it similar to "soft" values.

Appendix

Code

- DRA vs maxEnt RL code

Backlinks

- Alex meeting DRA vs maxent
- Alternative to softmax
- code DRA vs maxEnt RL original
- code DRA vs maxEnt RL refactored v0
- code DRA vs maxEnt RL refactored v0.1
- code refactoring DRA vs maxEnt RL
- differences DRA maxEnt
- DRA analytical gradient for 2 options
- DRA vs maxEnt RL code
- Ideas for hypocampus++
- Inference lamp
- Luigi Alex DRA as inference
- MaxEntRL frequency vs stakes
- Regularization in RL
- Research projects MOC
- Simulations DRA vs max-entropy RL
- Trajectory vs time-step