Spring 2020

Internal Badger Workshop

Session 6: Dreaming and Deliberating in Badger

Moderator: Nicholas Guttenberg

- Future stages of Badger - what would it take to have an architecture that benefits from 'being able to sit and think', 'make a plan for how to solve a task', etc.? What do inputs and outputs look like? What kind of internal structure? What is the 'outer loop policy' here?
- Twitter post from Marek

Pre-discussion Comments & Resources

Things which benefit from more run-time, even without additional inputs:

- Search (MCTS for example); related: planning, path-finding
- Optimization (explicit optimization algorithms in inner loop?)
- Counterfactuals (what-if questions, internal speculation - sort of like getting new 'information' but it's self-generated rather than external)
- Computationally-bottlenecked calculations where accuracy/extent scales with resources (imagine something like an ODE/PDE solver, versus timestep size or grid resolution)

How should we do planning in high-dimensional, highly abstract state spaces that our own networks construct? That might be too hard to think of at once, so lets break things down into related questions:

- Can we make state graphs/representations which are convenient to plan on (soft planning, something like world models, ???)
- What kinds of planning are appropriate for big, open-ended state spaces? Is there some kind of open-ended pathfinding we can use? MCTS/Fractal AI? Also, are there differentiable versions of these so we could directly optimize for 'plannability'?
- Maybe consider a program synthesis approach, where we 'plan' the route that information/messages are passed through a sequence of experts. How are the experts trained in this case (outer loop, inner loop)?

Discussion Notes

- Running an RNN, trained on a language task, for longer does not bring any benefits
- But some processes (mentioned in text above) do benefit from longer computation
- These are usually some forms of search methodologies where time allows for more thorough exploration
- Question - can we learn to deliberate / benefit from prolonged computation?
- How do we pose planning as a loss function to result in the behaviour we desire
- Counterfactuals are an important aspect of this topic
- It might be necessary for the inner loop to be generative in some sense so as to allow for even having a chance at getting some benefit from prolonged computation?
- Relation to cultural evolution and dual inheritance
- Existing meta-learning (learning to learn) methods in actual fact don't learn learning algorithms but simply perform inference with meta-learned priors.
    - can we put search algorithms in the inner loop?
- The multi-agent nature of badger on its own might warrant a deliberation in order to adapt to evolving nature and to retain feasibility of continual learning and adaptation
- Mathematics as an abstract form of deliberation without the need to interact with the environment
    - Objections to the lack of need of environmental impact?
- Even though experts in badger are inside the mind of an agent, how should the topology and the internal structure of a badger agent be perceived. As environment or something else or nothing in particular?
- Desire to focus on incorporation of planning/search as part of the inner loop

References/Researchers mentioned during discussion

- Multi-Agent Reinforcement Learning with Multi-Step Generative Models
- Improving Policies via Search in Cooperative Partially Observable Games
- Doing more with less: meta-reasoning and meta-learning in humans and machines
- Analogues of mental simulation and imagination in deep learning
    - and related special issue of Current Opinion in Behavioral Sciences journal - SI: 29: Artificial Intelligence (2019)