Deep Reinforcement Learning and Its Neuroscientific Implications

Botvinick et al.
Presented by: Matthew Pisano

Overview

Introduction

Vanguard Studies

Future Directions

Challenges

Conclusions

Questions and Discussion

Introduction

Review and Applications

What does this review do?

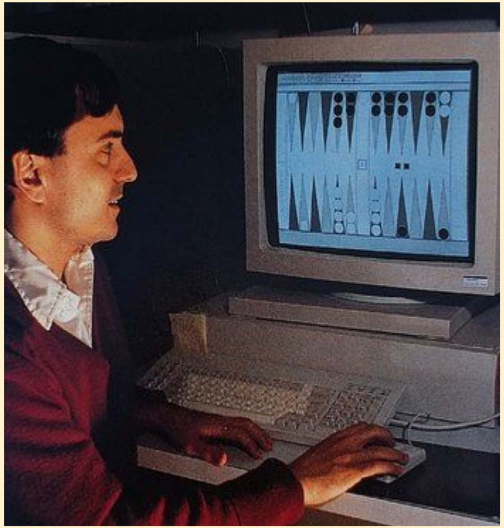
- A brief summary of recent RL advances.
- How RL has been implemented with deep learning.
- Future directions for deep RL.
  - What can we learn about the brain from this?

Why?

- Deep RL works well and imitates brain structure and behavior.
- Authors suggest that this can inspire useful insight into the brain itself.
- "Virtuous cycle" between AI and Cognitive Science

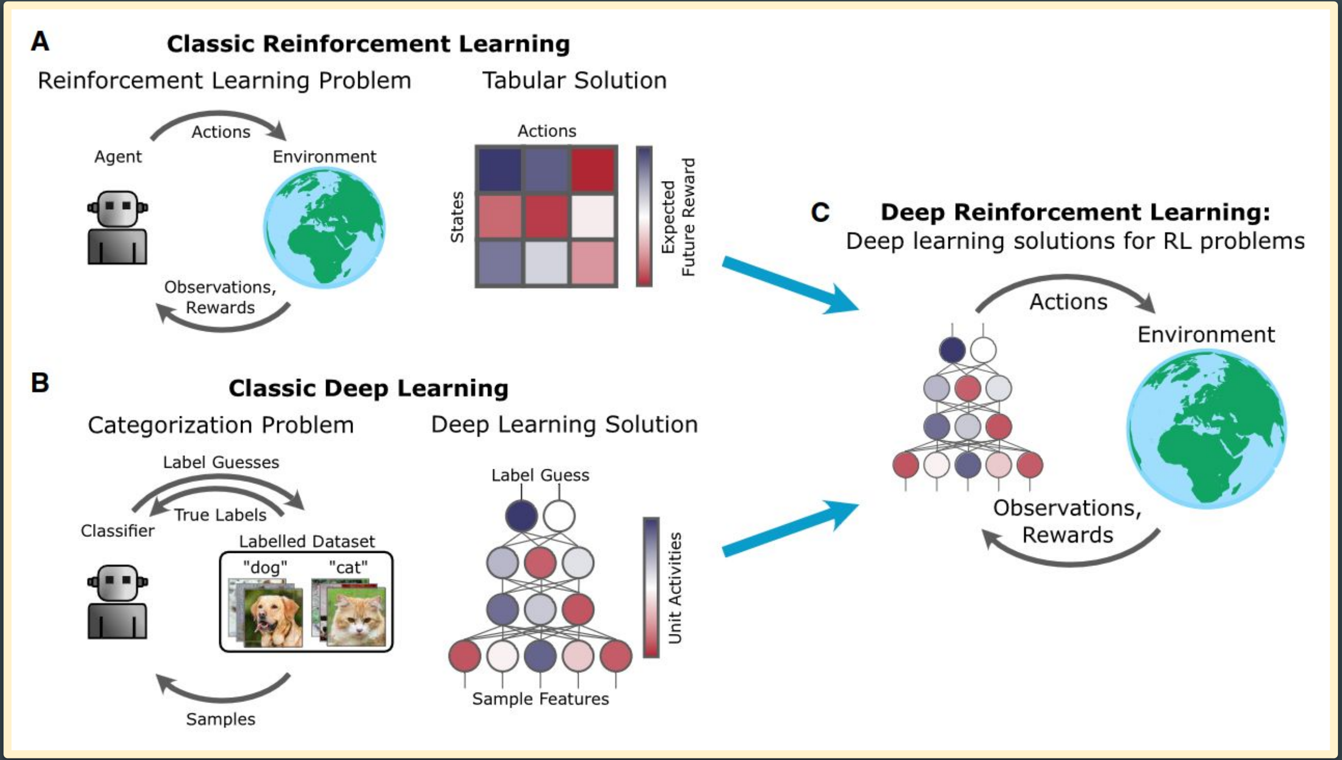
Concept Review

- What is RL?
  - A learning strategy, not an architecture.
  - Policies used to map states to actions.
    - Can be represented as non-linear functions.
- What is Deep learning?
  - Usage of neural network for learning.
  - Non-linear function approximators.
- What is Deep RL?
  - RL optimal policies approximated by deep networks.
  - Can learn extremely complex policies.
  - TD-Gammon: an early pioneer.
  - DQN: a modern exemplar.



G. Tesauro and TD-Gammon
(IBM Research)

Components of Deep RL



Deep Reinforcement Learning and Its Neuroscientific Implications (Botvinick et al., 2020)

Unique Angles

- What is different between this review and *Reinforcement learning, fast and slow*?
  - Previous goes over more details of techniques.
  - How deep RL works and why.
  - More of a broad overview or introduction.
- What does this review do differently?
  - Focus on neuroscience.
  - Many specific examples here of similarities.
  - Particular studies and implications for neuroscience.
  - Details challenges and differences from neuroscience.
  - Future direction and room for improvement.

Introduction Check-in

Questions?

Comments?

Implications for Neuroscience

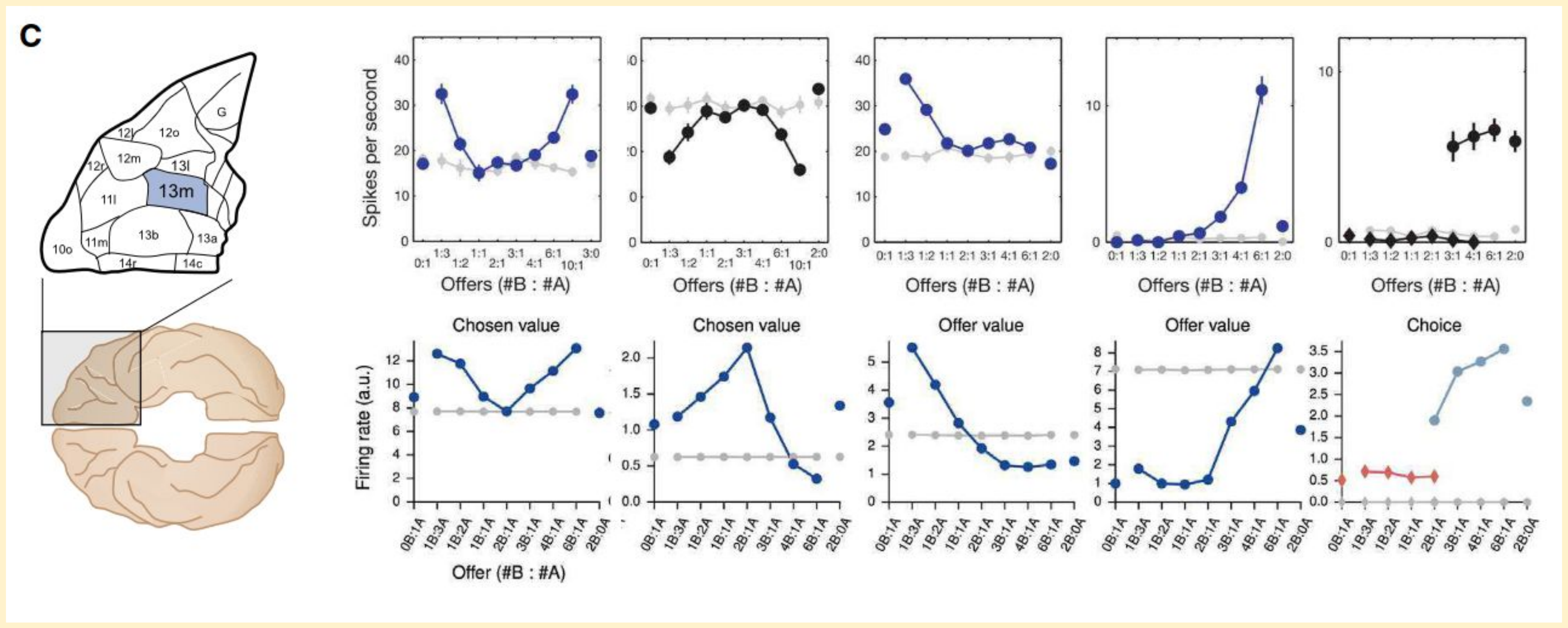
- Can mimic processes within the brain.
  - Relation to dopamine neurons.
- Mimics a way in which the brain really learns.
- Architecture encourages emergent long-term planning.
  - Complexity of networks allows for complex planning.
- Very effective usage in the realm of games.
- Demonstrates abilities not found in RL or deep learning separately.
  - More than the sum of its parts

Vanguard Studies

Ties Between Deep RL and The Brain

- Song et al. (2017) demonstrates that recurrent RL networks can mimic real neuron activity.
  - Observations of juice preference in monkeys.
  - Real neurons indicated juice presence.
  - RL network used to mimic this behavior when exposed to stimuli.
- Banino et al. (2018) shows how deep RL mirrors cells during environment navigation.
  - Grid cells in entorhinal cortex aid in an animal's navigation.
  - Representing this structure improves navigational abilities in RL agents.

RL to Mimic Monkeys' Preferences

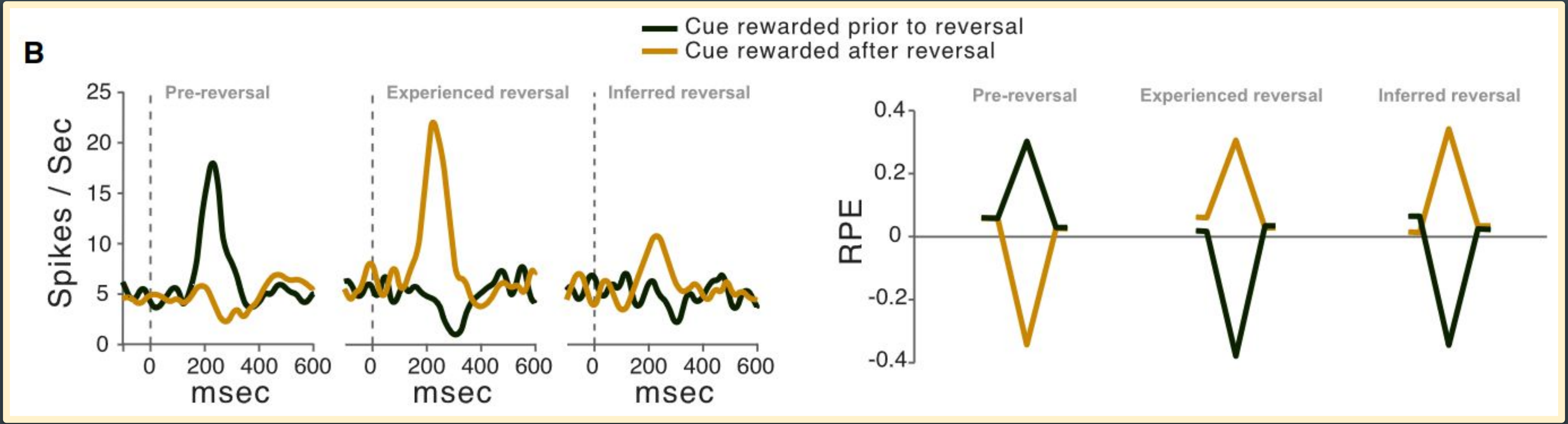


(Song et al., 2017, Padoa-Schioppa and Assad, 2006, Stalnaker et al., 2015)

Emergent Phenomena

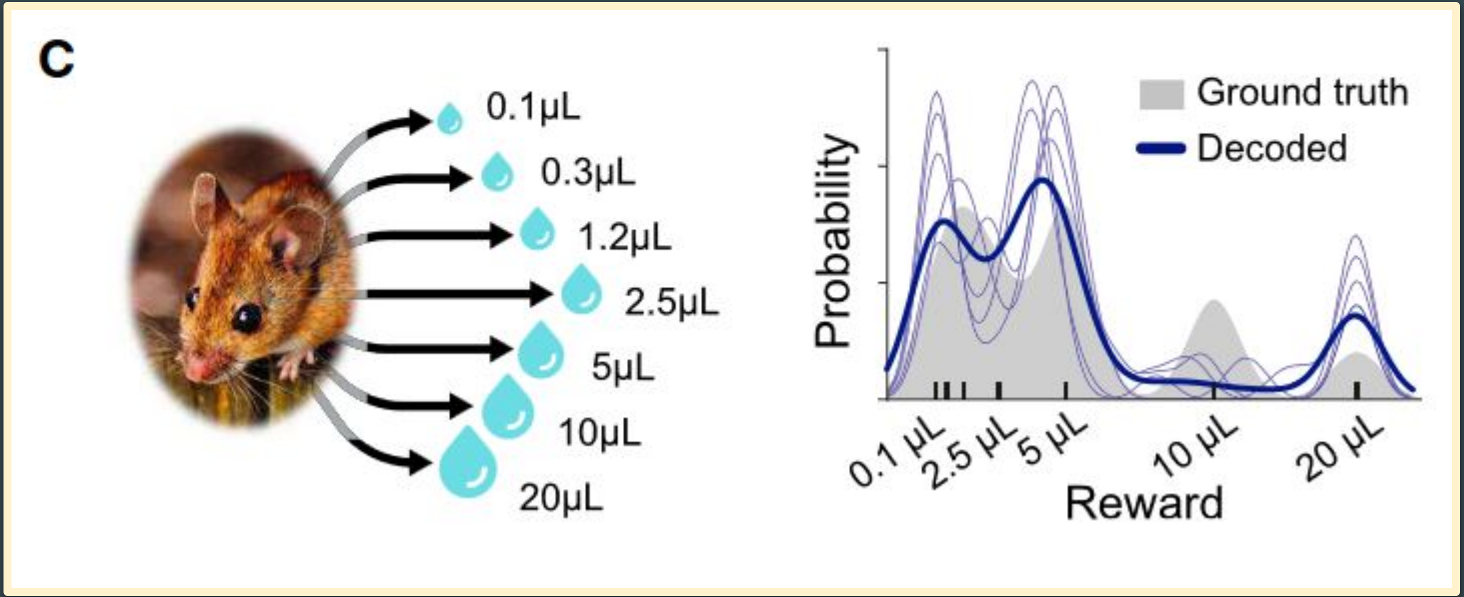
- Wang et al. (2018) shows that deep RL can generalize tasks.
  - Agents trained on similar tasks but different reward structures.
  - Agents can generalize to other, similar tasks without further learning.
  - Meta-RL similar to meta learning in real organisms.
- Dabney et al. (2020) utilizes distributional RL to predict dopamine spikes.
  - Used to predict perceived reward from a stimuli.
  - Distributions learned by model predict real dopaminergic signals.
  - Suggests that these signals are distributional in real brains, rather than scalar.

Meta-RL Predicting Dopaminergic Activity



(Bromberg-Martin et al., 2010, Wang et al., 2018)

Distributional RL Predicting Dopaminergic Activity



A distributional code for value in dopamine-based reinforcement learning (Dabney et al. (2020)

Vanguard Studies Check-in

Questions?

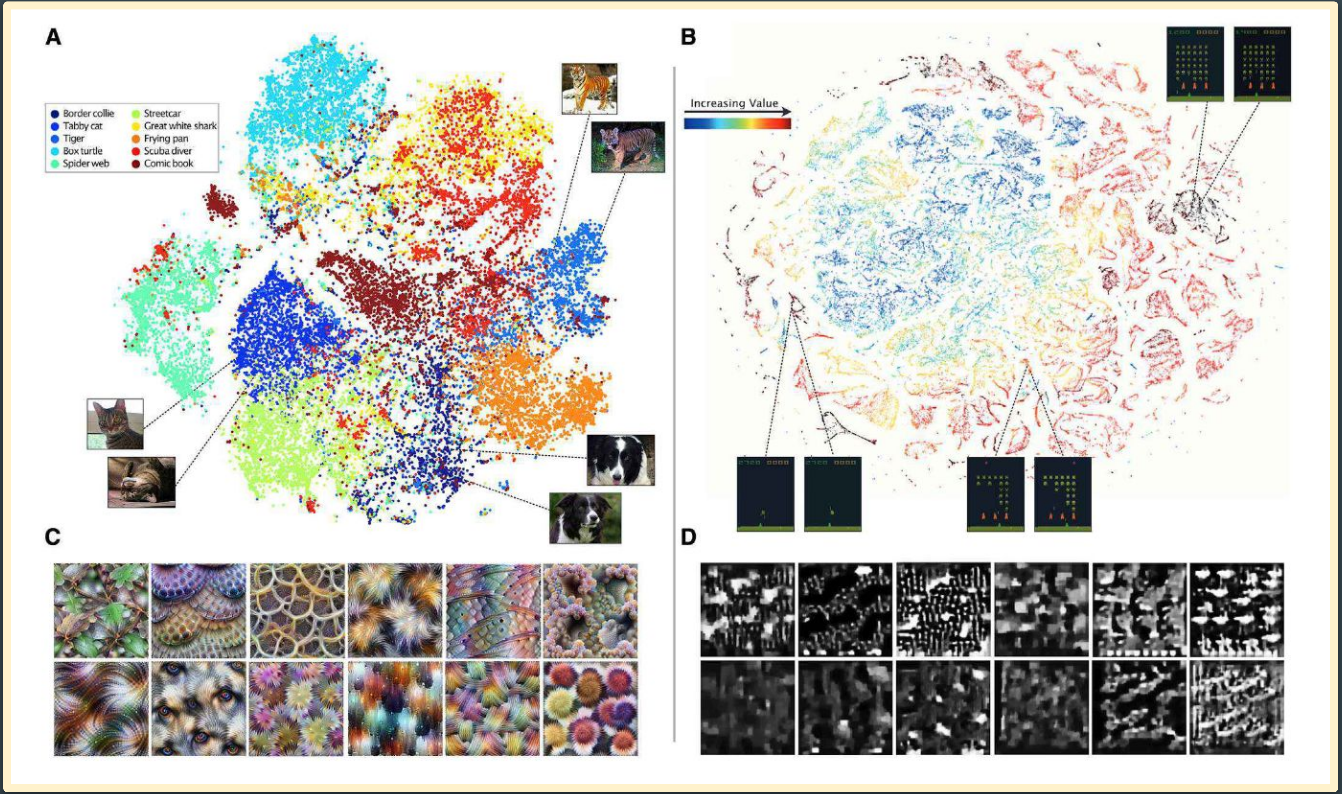
Comments?

Future Directions

Representation Learning

- Individual features associated with reward.
- Combination of features creates internal representation.
- Deep RL allows for the learning of complex features.
  - Easier transfer between related tasks.
- DQN Atari:
  - Good example of how one model can learn quickly through feature representation.
- Self-supervised learning.
  - Learned representations used to generate synthetic samples.
- Unsupervised learning.
  - Learns individual features of a task, rather than correct labels.

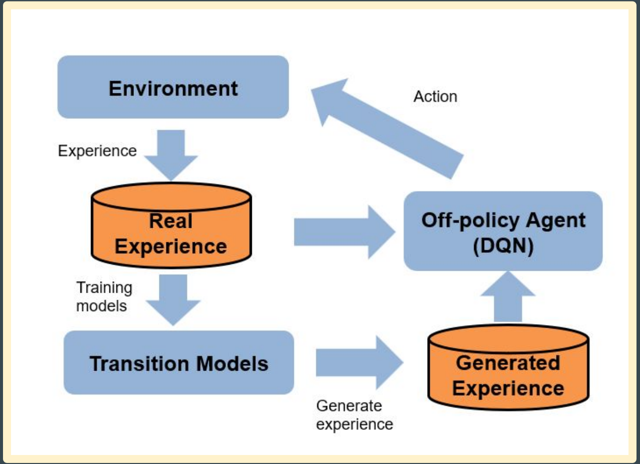
Representation and Feature Learning



(A: Deng et al., 2009, B: Olah et al., 2017, C: Bellemare et al., 2013, D: Suchet al., 2019)

Model-Based RL

- Agent develops internal model of rewards.
- Can forecast outcome of novel environment states and actions.
- AlphaGo uses model-based with help from model-free.
  - Model-free training used to improve planning.
- Deep RL can produce model-free planning.
  - Model-based behavior emerges from deep model-free agents.
  - Insight into how real brains decide between model-free and model-based approaches.



Memory

- Some approaches use *experience replay*.
  - Stored experiences are used along with new experiences during learning.
  - Use of recall helps to reinforce fruitful techniques.
  - Similar to recall and insertion from hippocampus.
  - Two types.
- Episodic memory.
  - Events written to long-term storage for later recall.
- Recurrent Memory
  - LSTM or GRU architectures.
  - Events stored in model weights.
  - Easy recall, easily forgotten or warped.
- Attention-based approaches.

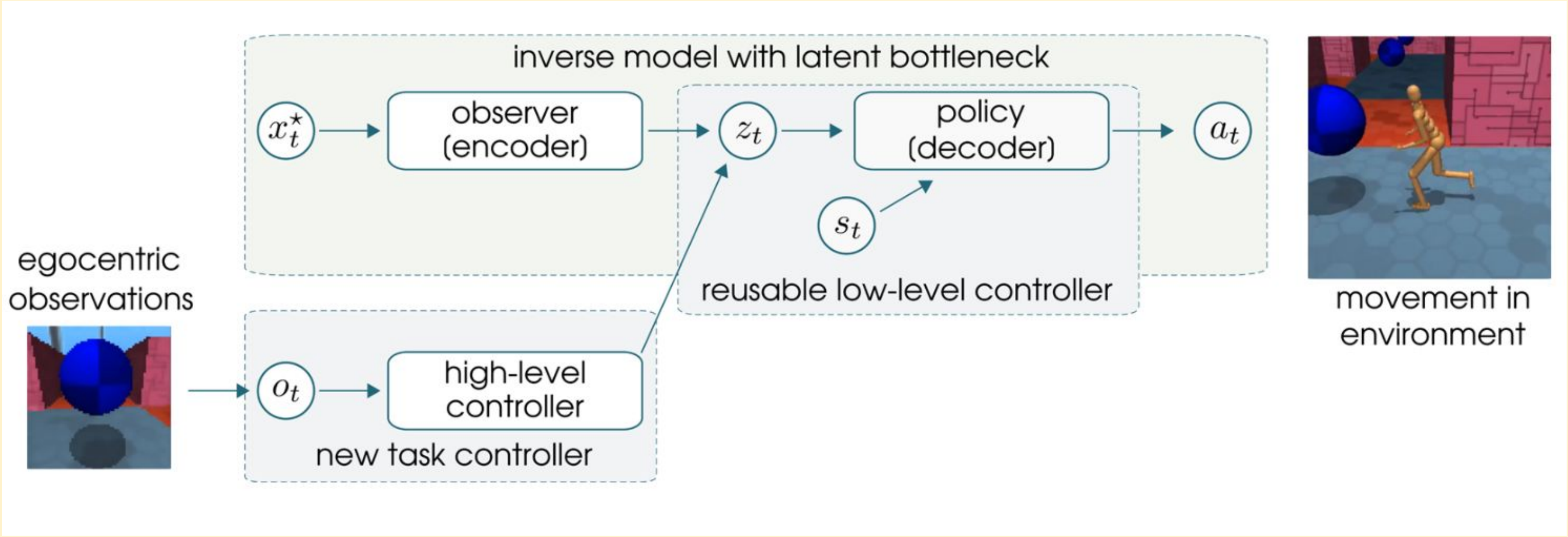
Exploration

- Both RL and deep RL need to explore to better understand their environment.
    - Novel situations harder for deep networks due to higher dimensionality.
    - Several distinct approaches.
- Based on uncertainty.
    - By exploring uncertain areas, the agent is exposed to more novel stimuli.
- Using meta-RL.
    - Similar to hypothesis-driven experimentation.
    - Deliberate choice to test unknowns.
- Foraging models.

Cognition Hierarchies

- Complex minds often involve hierarchies of control.
  - Simpler, low-level systems perform most tasks.
  - Governed by higher-level systems with more general goals.
- Can be applied to deep-RL.
  - Higher-level agents control simpler tasks.
  - Abstracts complexity away from agents.
  - They can focus on the more broad goals.
  - Mimics structures found in the brain.
  - Merel et al. 2018

Hierarchical motor control



Hierarchical motor control in mammals and machines (Merel et al., 2018)

Social Cognition

- Deep-RL agents have capacity to behave socially.
  - Demonstrated in multi-agent environments.
  - Deep networks allow for complex strategies.
  - Can cooperate, be anti-social, etc.
- Agents can develop basic *theory of mind.*
  - Predict the behavior of other agents.
  - Modify their behavior based on this.

Emergent tool use from multi-agent autocurricula (Baker et al., 2019)

Challenges

- Many places where artificial agents not comparable to biology.
  - Shortcomings in deep RL make good future research.
- Deep RL much slower than biological learning.
  - Requires orders of magnitudes more data.
  - Can be improved with meta-RL / episodic learning.
  - Humans already have many “built-in” relevant knowledge.
- Deep RL still not as flexible.
  - Brains still much better at generalizing tasks.
  - Learning new tasks from scratch.
- Brains much better at credit assignment.
  - Knowing the exact cause of reward and how to reliably replicate.

More Caveats

- Backpropagation.
  - Very little biological connection.
  - Likely separate from biological means of learning.
  - May be underlying reason for “slowness” of learning.
  - Used because it works, not because of neurological inspiration.
- Most research conducted strictly for better performance.
  - May align with biology, may not.
  - Could converge on an approach that works well.
    - Completely disconnected from biology.
  - Usually not a generic or evolutionary approach.

Conclusions

- Deep RL rapidly gaining popularity.
- Many future approaches mirror biological ways of learning.
  - May give further insight into how real brains learn.
  - "Virtuous circle" between AI and neuroscience.
- Many different areas of research.
  - Open questions with room for formalization.
- Still many issues to face.
  - Deep RL still exhibits sample inefficiency.
  - Founded on non-biological algorithms.
  - Motivation to improve, regardless of the methods.

Questions

- What would happen if both neuroscientific and non-neuroscientific methods are both successful at solving a complex RL task? (Paraphrased from Jenny Shen)
  - Which would draw the most research attention?
  - If it better to go with something easy to improve or something inspired by the brain?
- “How can multi-agent deep RL be used to develop new models of social behavior?” (Omar Cobas)
  - Can this discover novel solutions to complex social problems?

Questions

- “What are the potential implications of applying Deep RL to our current understanding of the roles of emotions and affective states in decision-making processes?” (Brianna Lopez)
  - Can emotions be modeled as different policies for choosing actions?
    - An “angry” model may have a greedy policy, a “depressed” model may be pessimistic when estimating rewards.
  - Would this be useful or would it only hurt performance?

- “Are there insights from the way that biological brains work that could help in making deep RL more sample efficient?” (AJ Jansen)
  - A more generalized or hierarchical set of representations?

Thank You