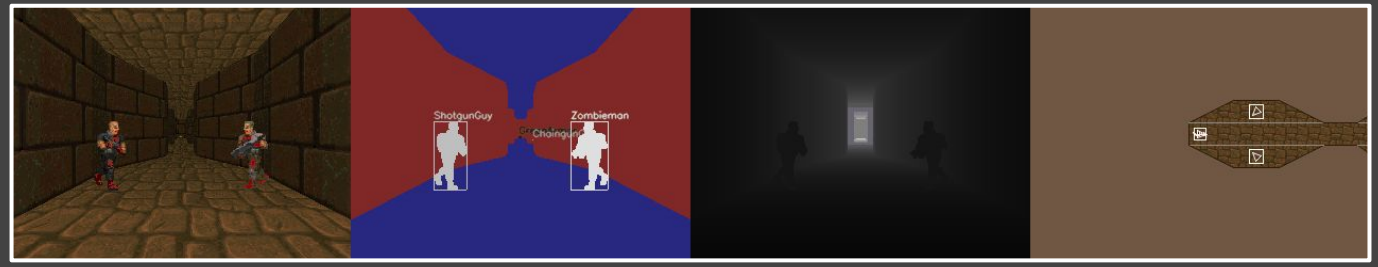
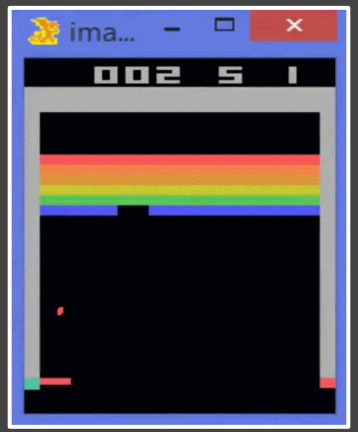
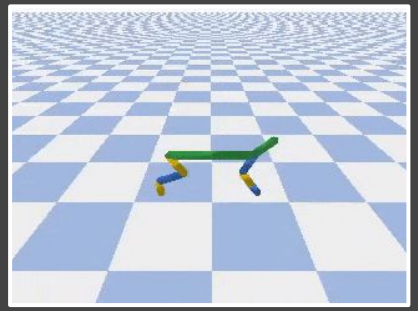
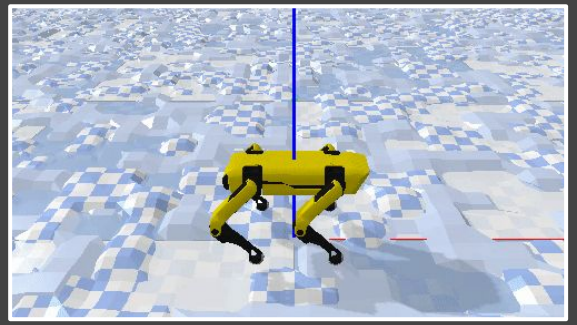
Advanced Reinforcement Learning Brainstorm Session









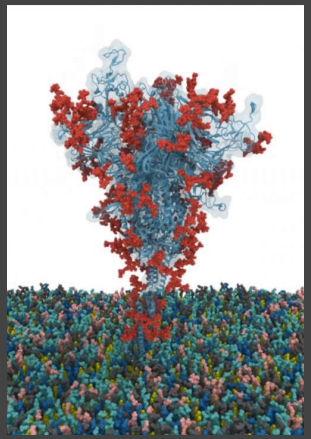
Poll: Check all the RL Algorithms you've heard of before

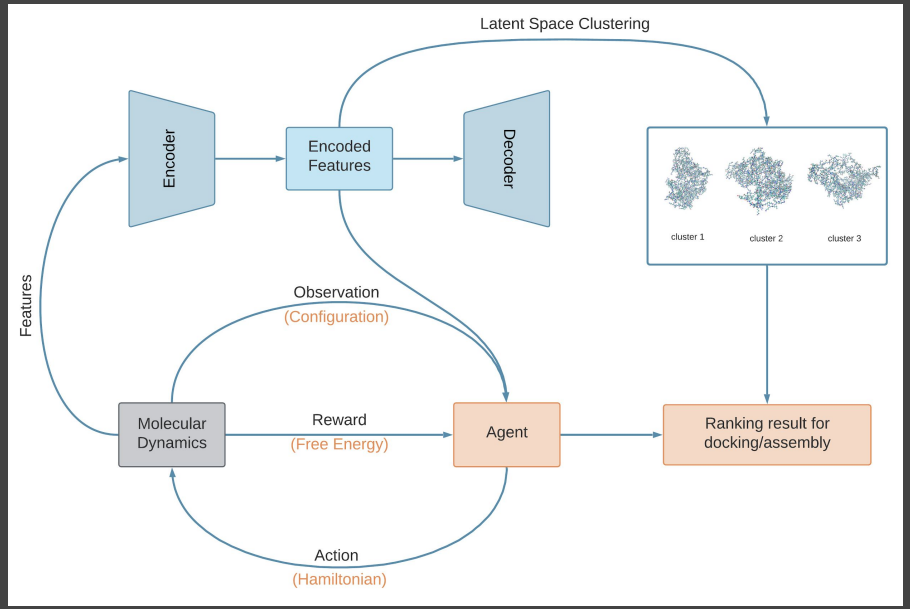
- TD(λ) (Temporal Difference)
- SARSA (State-Action-Reward-State-Action)
- DQN (Deep Q-Learning, Dual…, Double…, Actor-Critic…)
- GAE (Generalized Advantage Estimation)
- ACKTR (Actor Critic + Kronecker-Factored Trust Region)
- ACER (Actor-Critic with Experience Replay)
- TRPO (Trust Region Policy Optimization)
- PPO (Proximal Policy Optimization)
- SAC (Soft Actor Critic)
- SLAC (Stochastic Latent Actor Critic)

About Me

- I’m pretty much David Silver/Richard Sutton/Andrew Barto (jk)

- RL Research
  - Protein Folding
  - Multi-Policy Swapping



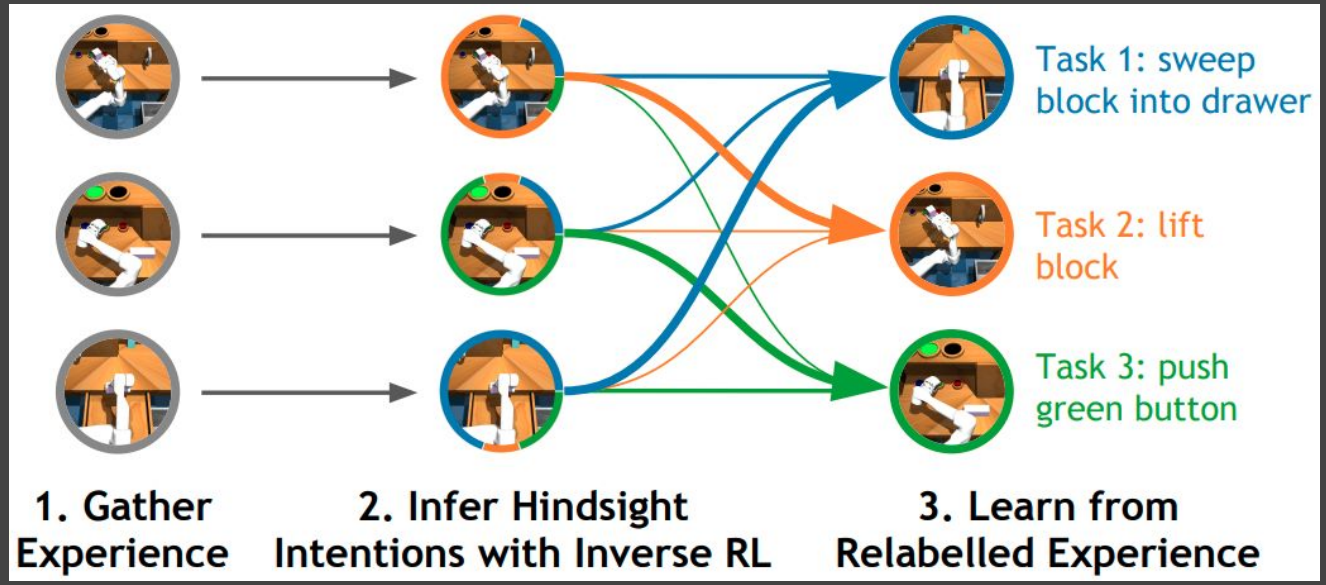


Ideas for this group

- Dive deeper into the aforementioned advanced algorithms
- Apply advanced algorithm on a cool project
- DooM AI Competition? Atari Env Competition?
- ...?

Keep Up with RL

- Inverse-RL:
  - Carnegie Mellon/UC Berkeley/Google brain
  - https://arxiv.org/pdf/2002.11089.pdf



Keep Up with RL

- SLAC:
  - DeepMind/UC Berekely
  - https://arxiv.org/abs/1907.00953

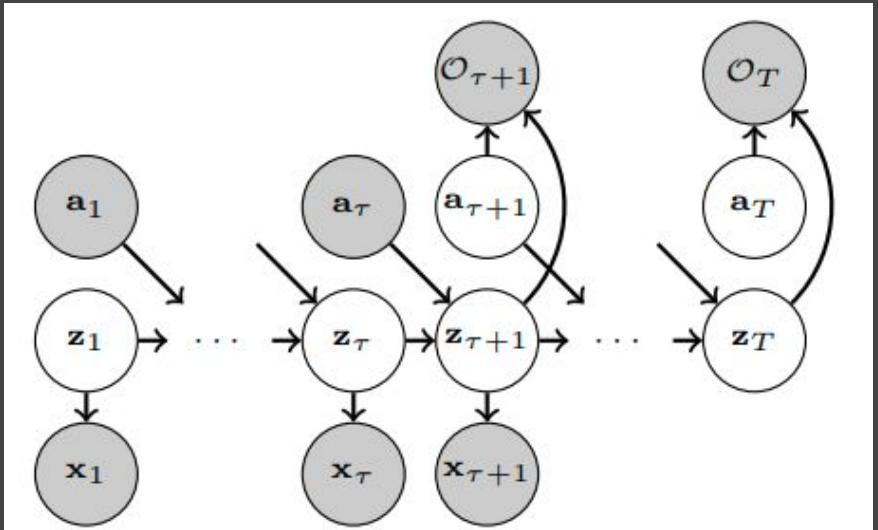
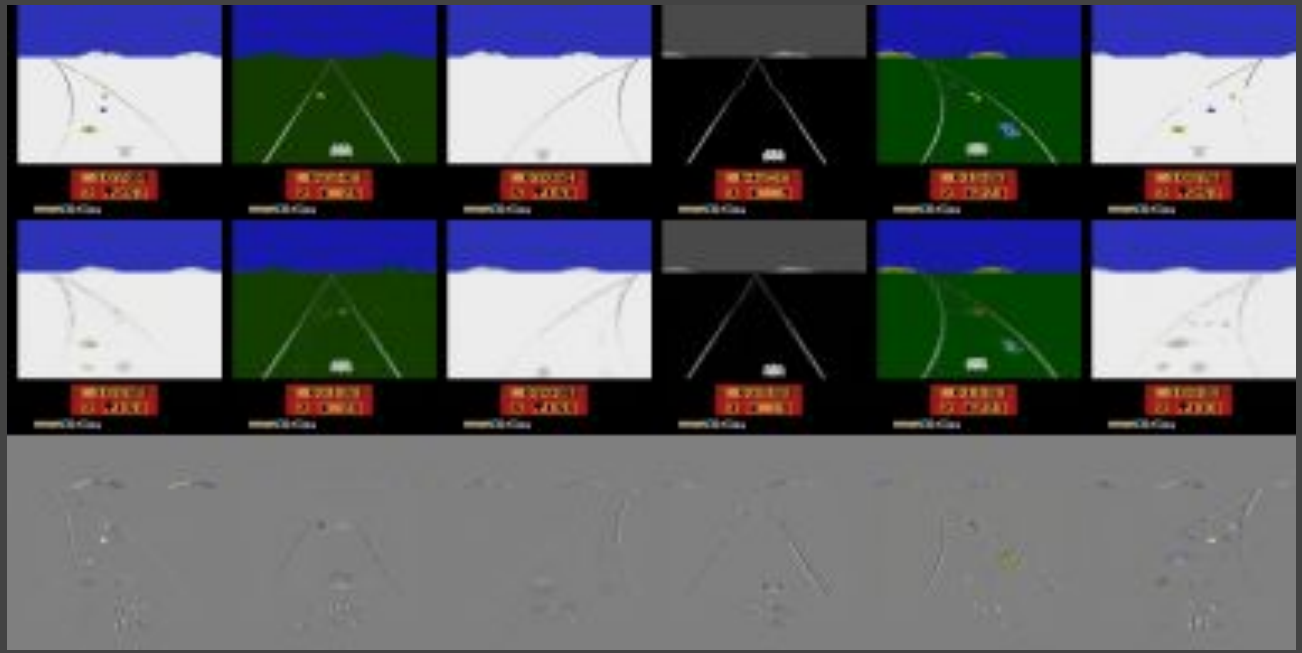


Figure 1: Graphical model of POMDP with optimality variables for $t \geq \tau + 1$.

Keep Up with RL

- DreamerV2:
  - GoogleAI
  - https://ai.googleblog.com/2021/02/mastering-atari-with-discrete-world.html



Keep Up with RL

- Go-Explore:
  - OpenAI/UberAI
  - https://www.nature.com/articles/s41586-020-03157-9
  - https://arxiv.org/abs/2004.12919