Exploration via Posterior Sampling: A Bayesian Approach to Meta-RL

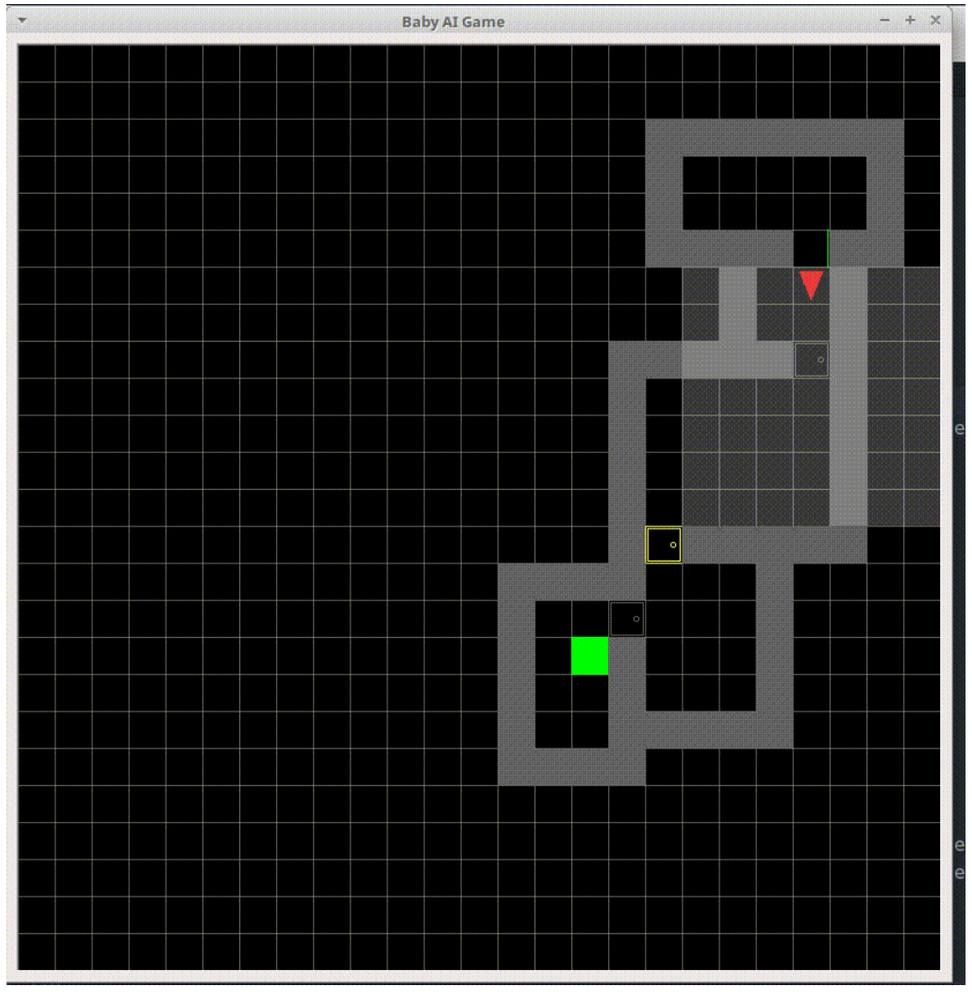
Team: MetaBayes

Member: Alexa Bosworth

Experience: Summer research on the topic; implemented a benchmark for the problem

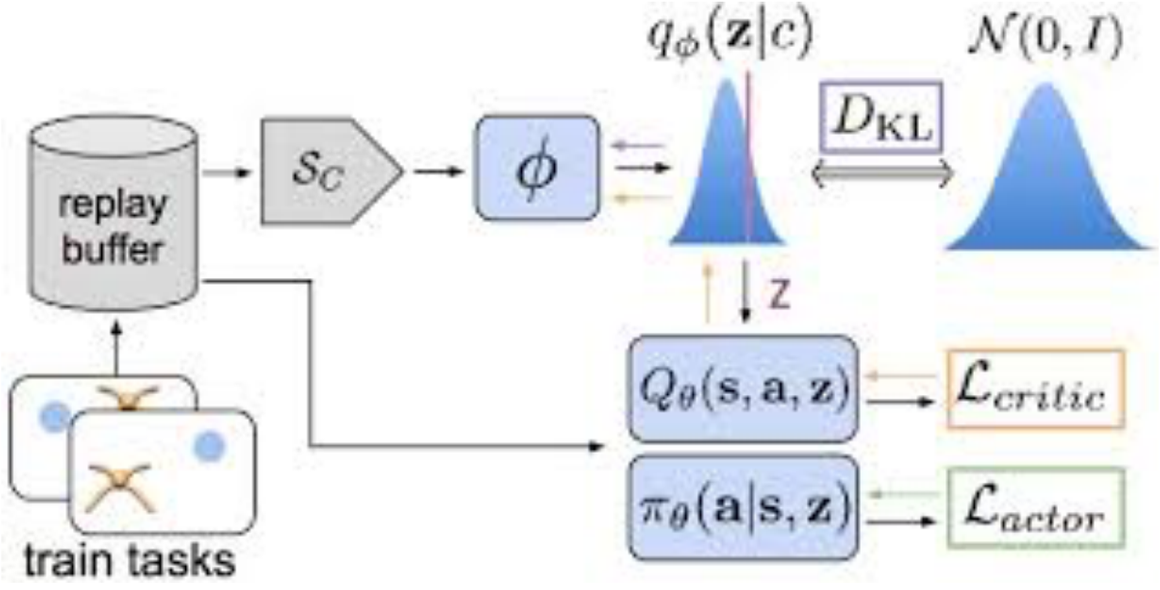
Problem Overview

- Reinforcement Learning problem
- Dataset: Gym-MiniGrid; multi-room environment
- Benchmark: DQN
- Performance metrics: steps, compare performance in environments of varying complexity



BDL relevant research question

- Hypothesis: When introduced with an entirely new environment PEARL will outperform a DQN.



- Models: PEARL and DQN

- Algorithms: Thompson sampling, Q function, SAC

Timeline*

- Major milestones:
    - Environment setup, implement the sampler, posterior network (2 weeks)
    - Integrate the SAC network (2 weeks)
    - Design Training curriculum (2 weeks)
        - Implement Benchmark (1 week)
        - Collect and interpret data (1 week)
- Challenges:
    - Training curriculum
    - Not a widely reproduced algorithm; limited to authors literature and documentation
    - Training time will be extensive