

Challenges Learning Continual RL

- ▶ Transition function depends on *not observed* data
- ▶ for our LunarLander a single random variable $\in \{0, 1\}$ is enough
- ▶ Need to learn this from last transitions

Challenges Learning Continual RL

- ▶ Transition function depends on *not observed* data
- ▶ for our LunarLander a single random variable $\in \{0, 1\}$ is enough
- ▶ Need to learn this from last transitions
- another network (RNN?) with multiple transitions as input and a *task embedding* as output
- ▶ hard to learn → agent could do this itself otherwise
- ▶ even harder if endlessly changing
- ▶ May already be covered by “standard” env-learning algorithms

Opportunities Learning Continual RL

- ▶ sample task embeddings, no need to start from beginning
- agent can learn all of them equally often/priority weighted

Opportunities Learning Continual RL

- ▶ sample task embeddings, no need to start from beginning
- agent can learn all of them equally often/priority weighted
- ▶ out of distribution embeddings may have meaning to them
- better learning through more data