How Should an Agent Practice?
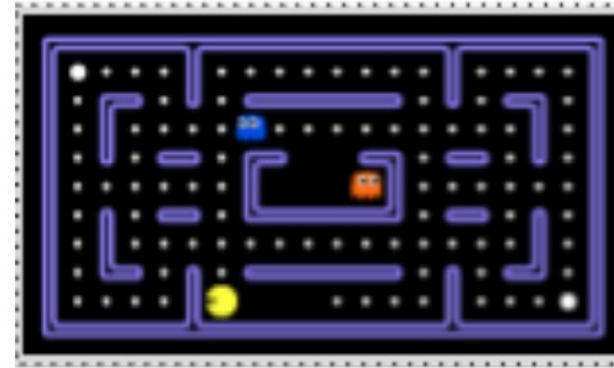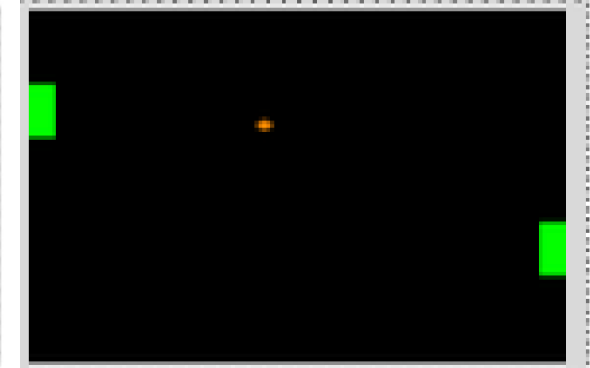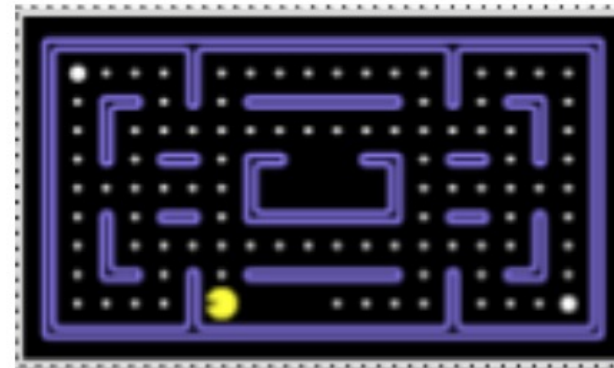
Rajendran et al. (2019)

- Practice environment with intrinsic rewards

- Match environment with extrinsic rewards

- Analogy to sports or other real world applications where "matches" are infrequent

- Learn on both environments alternately

- Intrinsic reward model is updated with meta-gradients based on extrinsic rewards from match environment

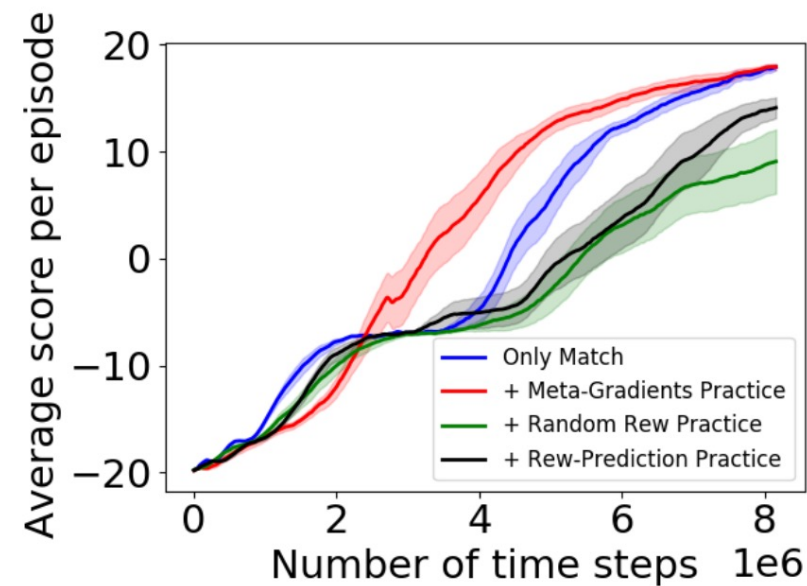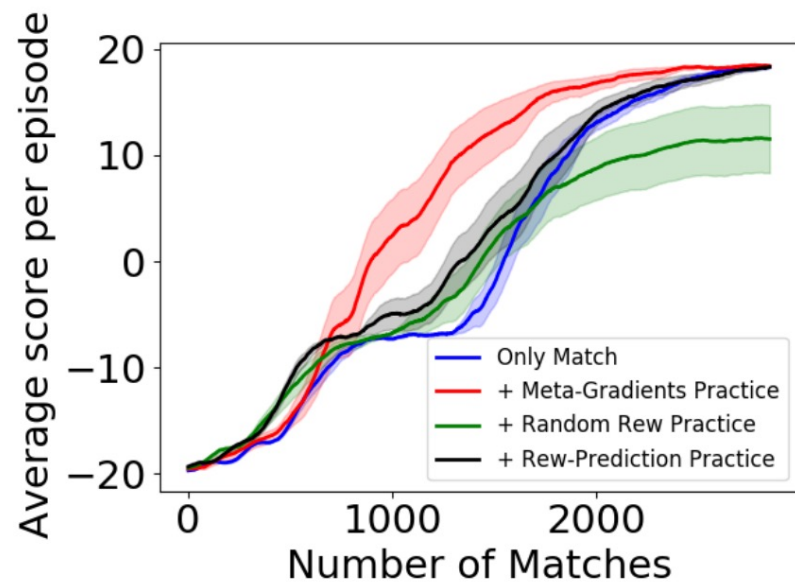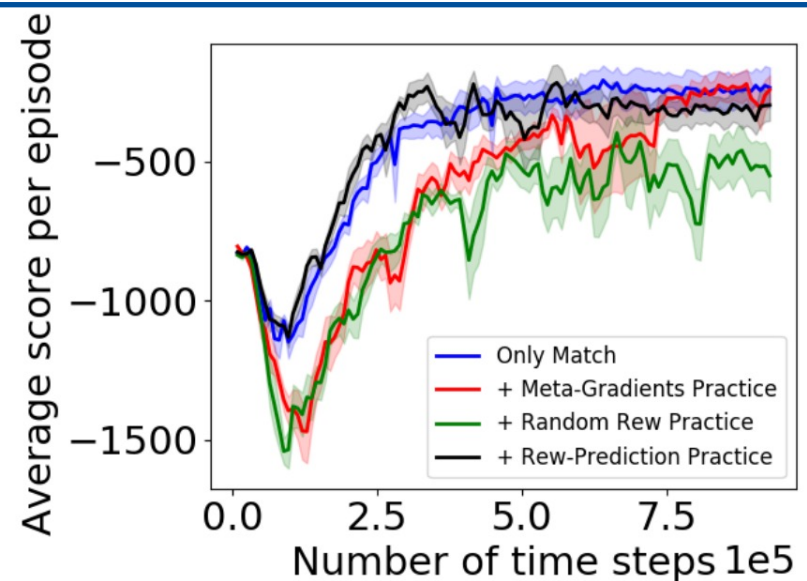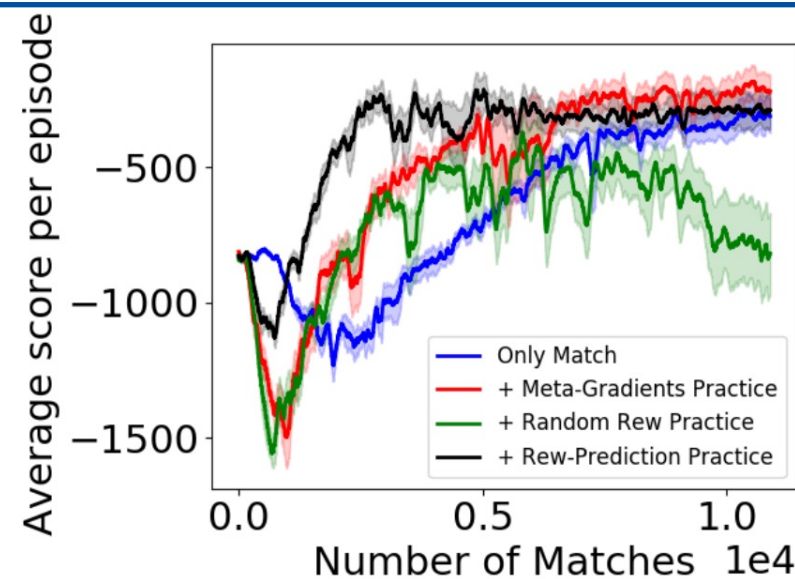(a) Match env

(a) Match env

(b) Practice env

(b) Practice env

- Pacman

- Pong

Results

- Pong

- Pacman

Rajendran et al. (2019)