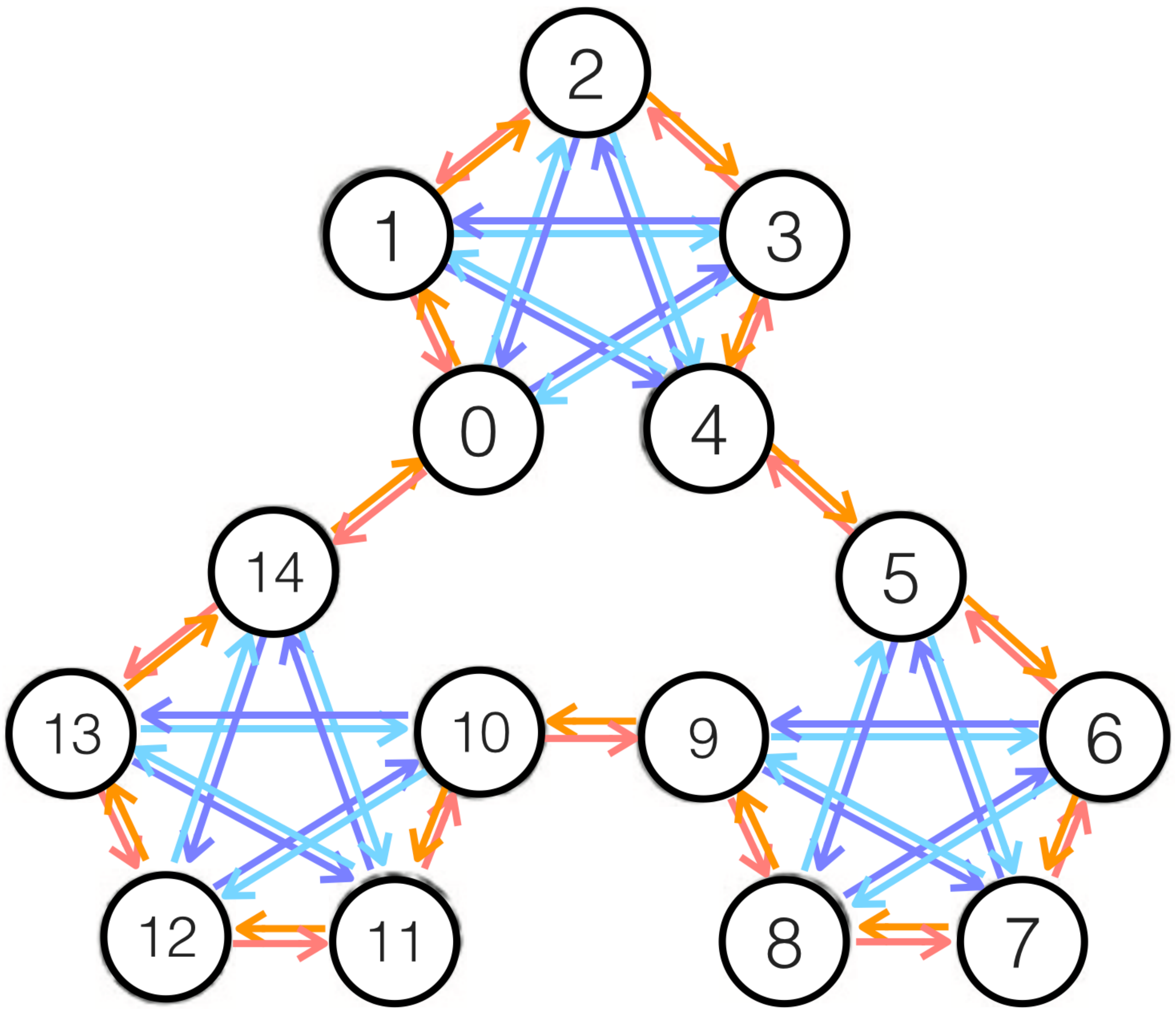
Curious transition learning facilitates subsequent planning

Effie Li | EDUC234 pre-report

World 1

15 states, connected by 4 actions



Question

Does learning step-level transitions facilitate subsequent multi-step planning?

Transition learning:

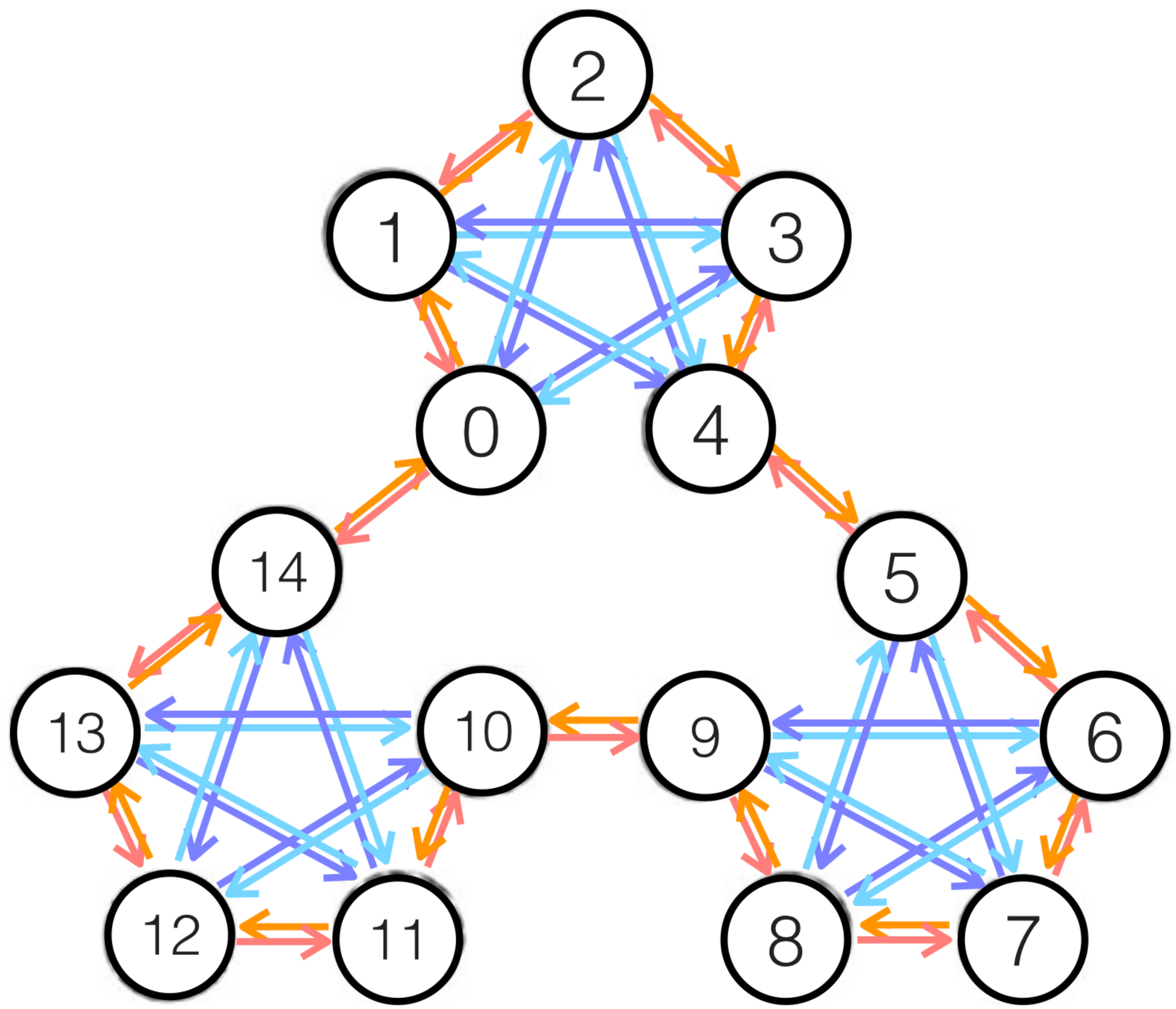
(s1, s2) —> a

where s1, s2 are adjacent

Multi-step planning:

(s, g) —> (a1, a2,…)

where s, g are non-adjacent



Model

Transition learning:

3-layer MLP

Multi-step planning:

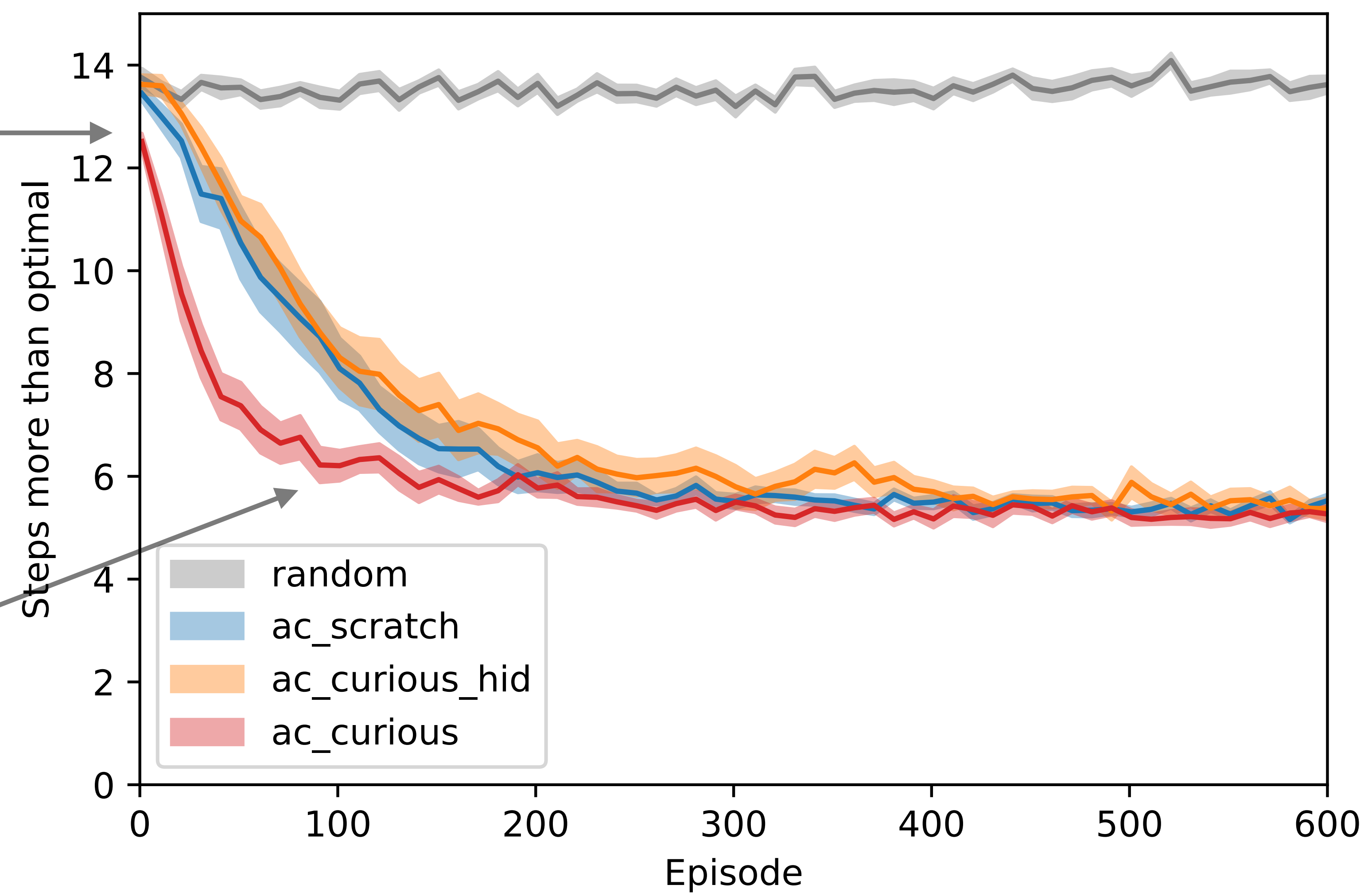
actor-critic agent + weights from the transition learning model

Performance on multi-step planning

Baseline 1 (random): random walk agent
Baseline 2 (ac_scratch): actor-critic trained from scratch
Baseline 3 (ac_curious_hid): actor_critic + transition learning hidden layer
Model (ac_curious): actor_critic + transition learning hidden layer & output layer

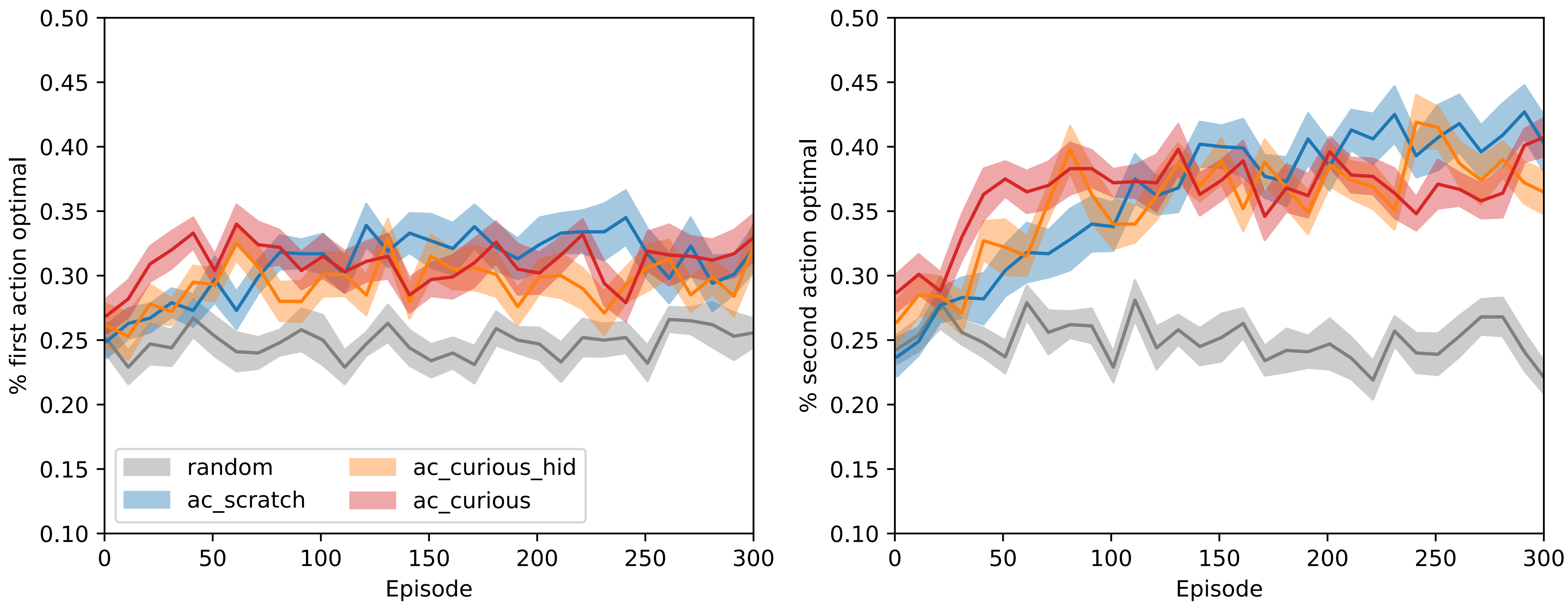
ac_curious has zero-shot adapting advantage

ac_curious learns faster



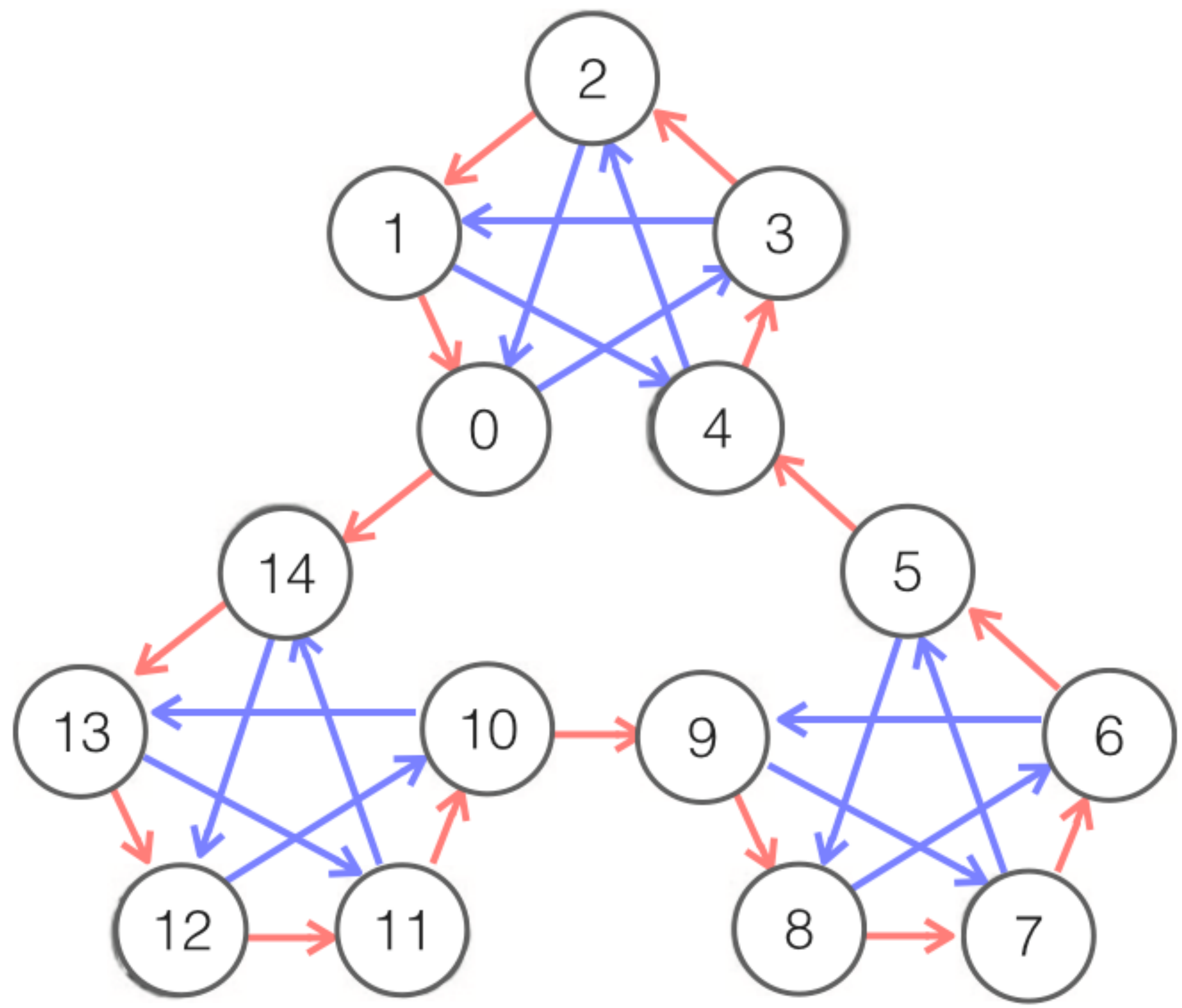
Performance on multi-step planning

Action optimal rate in 2-step problems



World 2

15 states, connected by 2 actions



Question

Does learning step-level transitions facilitate subsequent multi-step planning?

Transition learning:

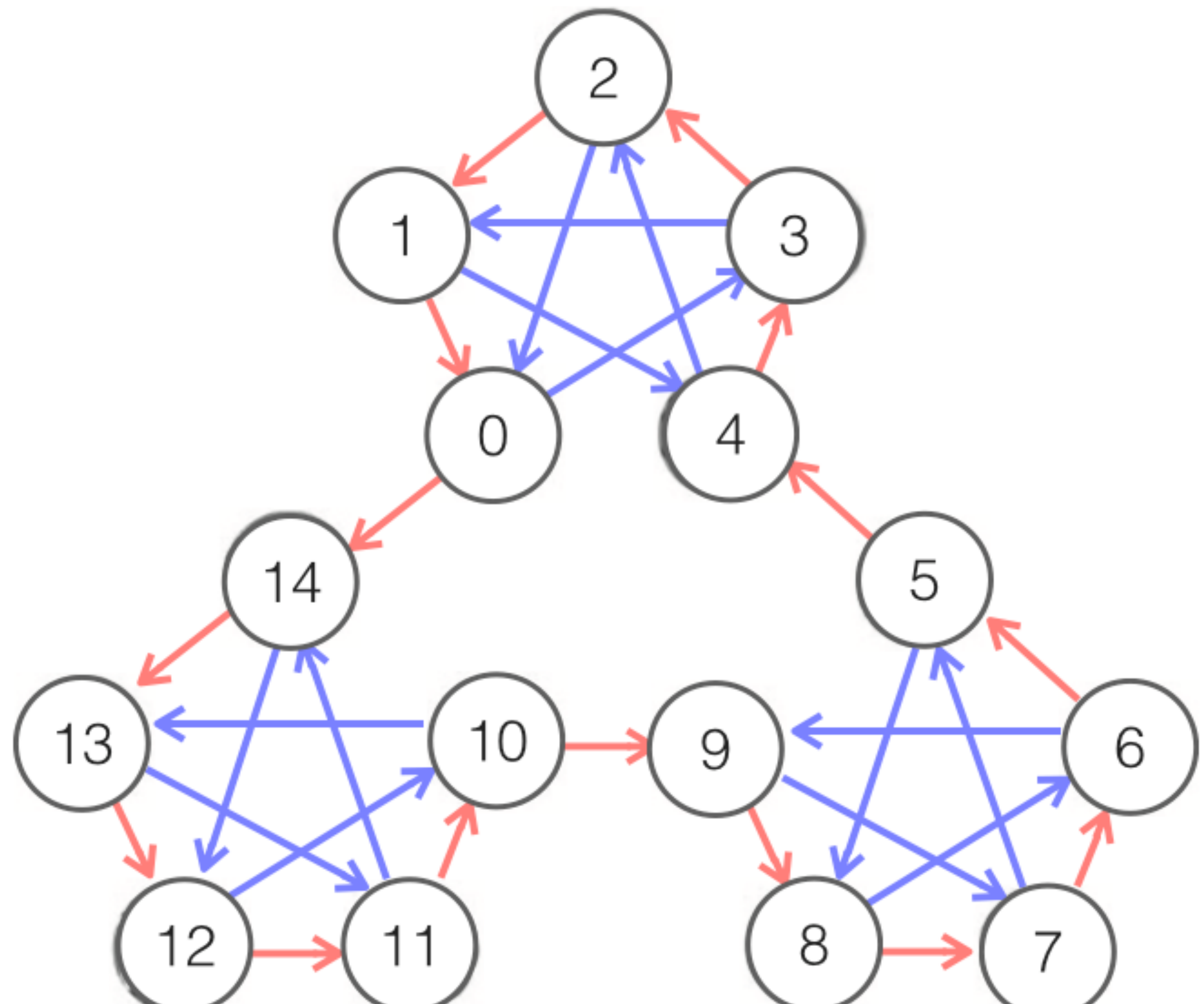
(s1, s2) —> a

where s1, s2 are adjacent

Multi-step planning:

(s, g) —> (a1, a2,…)

where s, g are non-adjacent

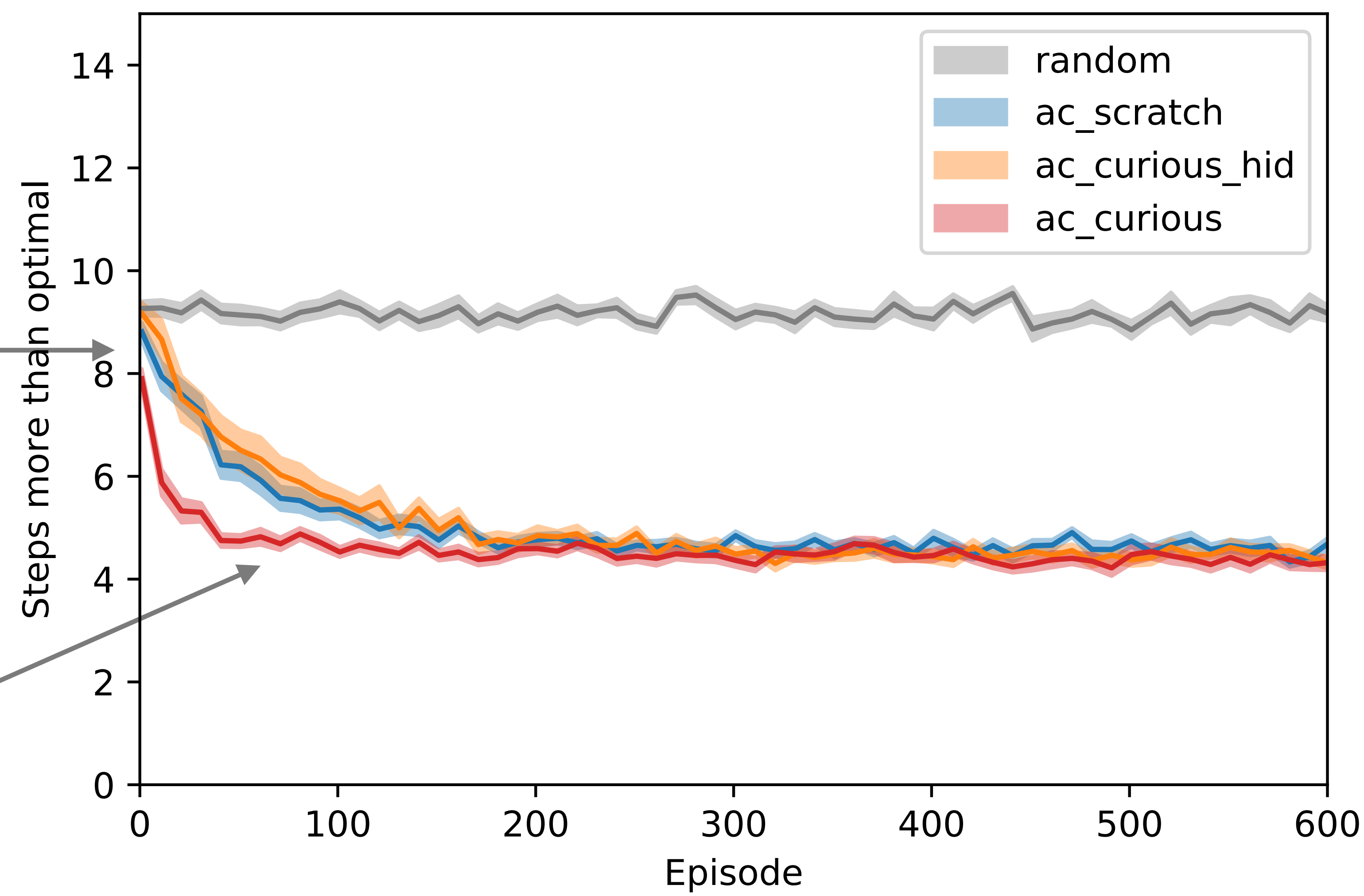


Performance on multi-step planning

Baseline 1 (random): random walk agent
Baseline 2 (ac_scratch): actor-critic trained from scratch
Baseline 3 (ac_curious_hid): actor_critic + transition learning hidden layer
Model (ac_curious): actor_critic + transition learning hidden layer & output layer

ac_curious has zero-shot adapting advantage

ac_curious learns faster



Next steps

1. Adjust model for better performance.
2. Systematically vary the transitions exposed to the transition learning model.
3. Test robustness to randomized state labels.