

On the Complexity of Exploration in Goal-Driven Navigation

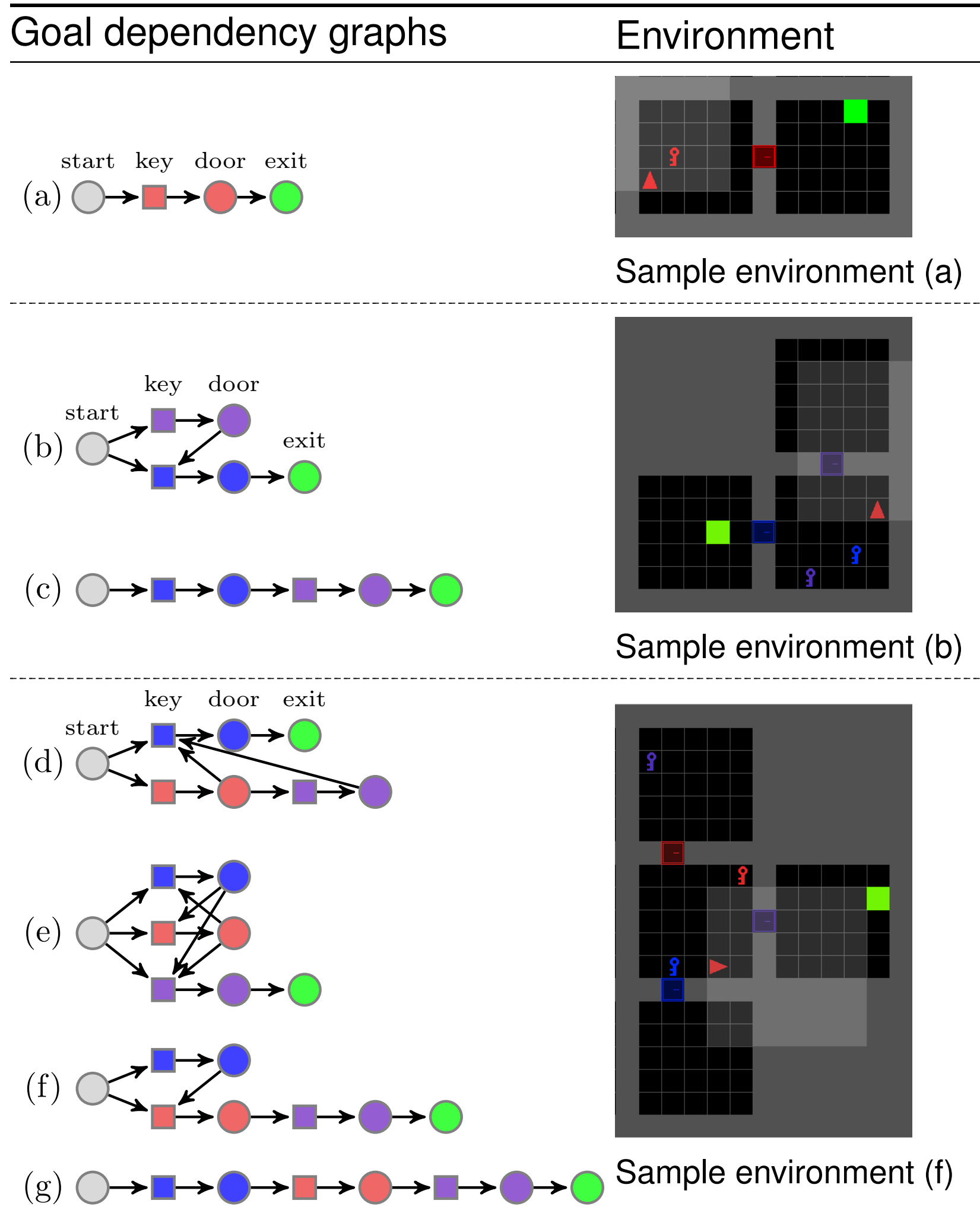
Lisa Lee ^{*1} Maruan Al-Shedivat ^{*1} Ruslan Salahkutdinov ^{1,2} Eric P. Xing ^{1,3}

1. Motivation

- **Measure the complexity of RL environments** by computing *hitting times* in goal dependency graphs.
- **Understand** how a hierarchical design of the agent’s policy can affect its exploration capabilities.

2. Environments

- Procedurally generate EscapeRoom environments.
- Environment complexity is controlled by the number of subgoals and relationships between them.



4. Results

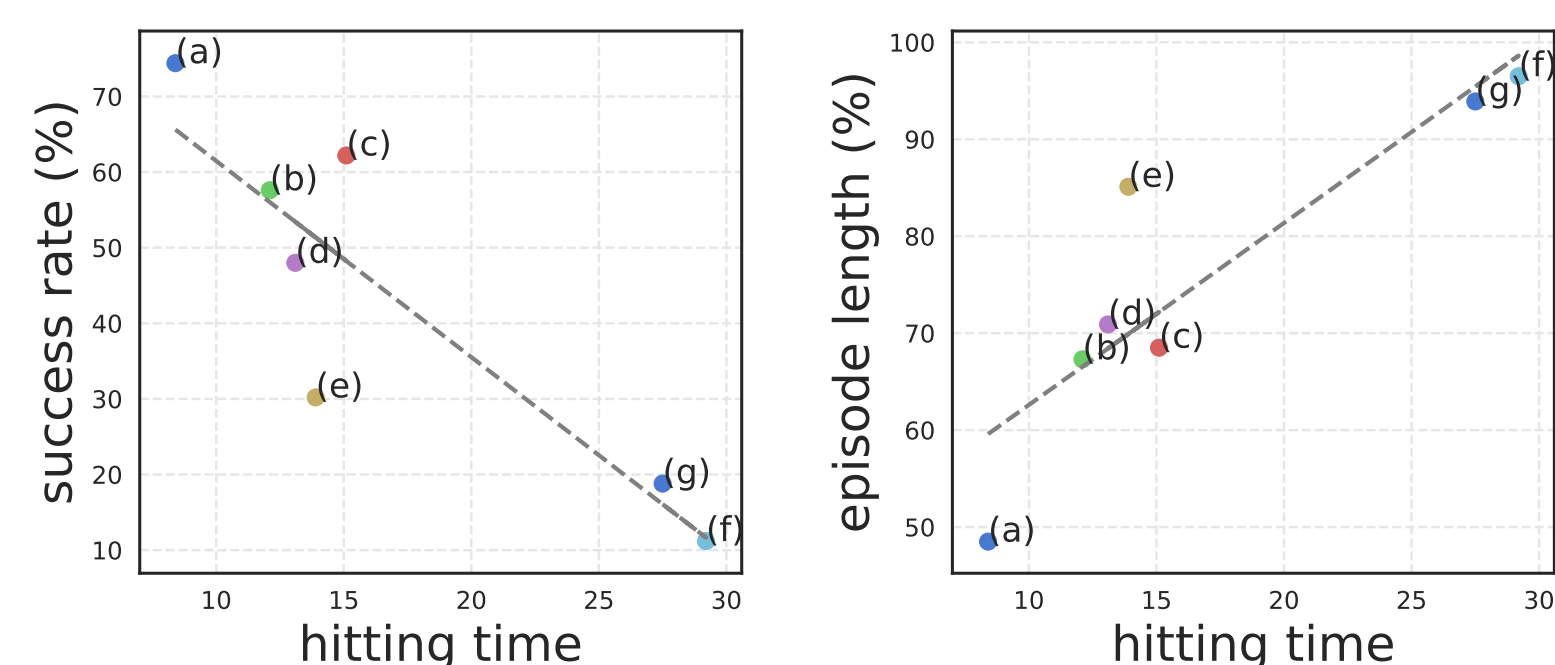
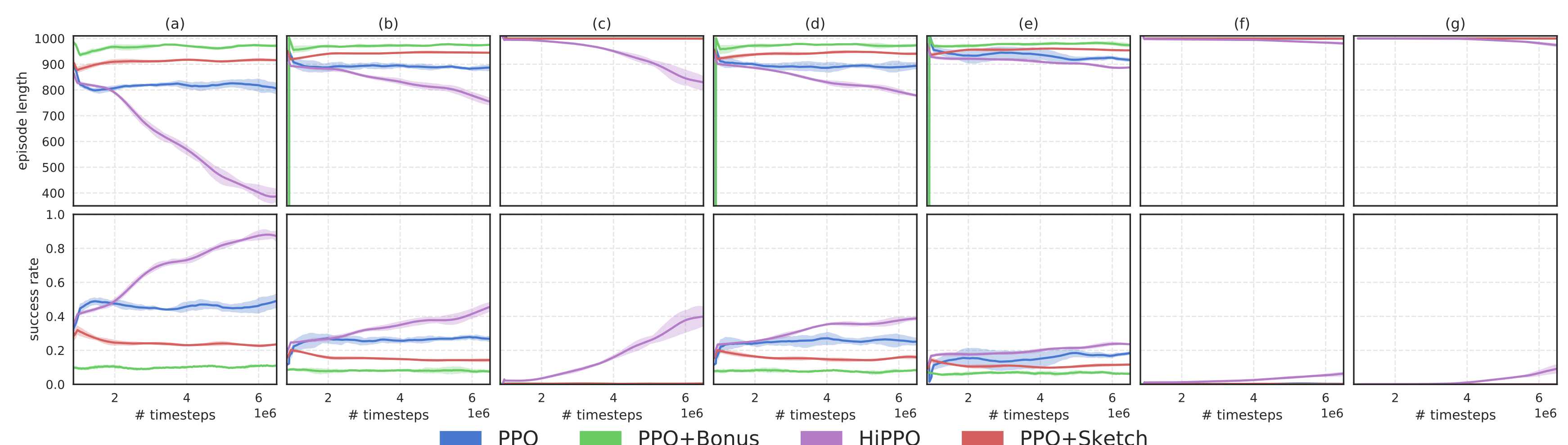


Table 2: Left: Average success rate (%) to reach the final goal over the last 10 training episodes. **Right:** Average episode length (% of the max length, smaller is more efficient) over the last 10 training episodes. “—” indicates failure to reach the final goal within 1000 steps.

	Average Success Rate							Average Episode Length						
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(a)	(b)	(c)	(d)	(e)	(f)	(g)
PPO	56.1	28.2	0.2	22.7	19.1	0.0	0.0	78.5	88.0	—	90.7	91.4	—	—
PPO+Bonus	9.0	6.0	0.0	11.0	4.5	0.5	0.0	97.8	96.7	99.9	97.3	98.0	99.9	—
PPO+Sketch	23.2	14.5	0.4	12.6	10.7	0.1	0.0	91.9	94.4	99.9	95.0	95.7	—	—
HiPPO	74.9	57.0	60.8	48.0	29.9	11.2	19.0	48.2	67.4	69.1	71.0	85.3	96.4	93.8



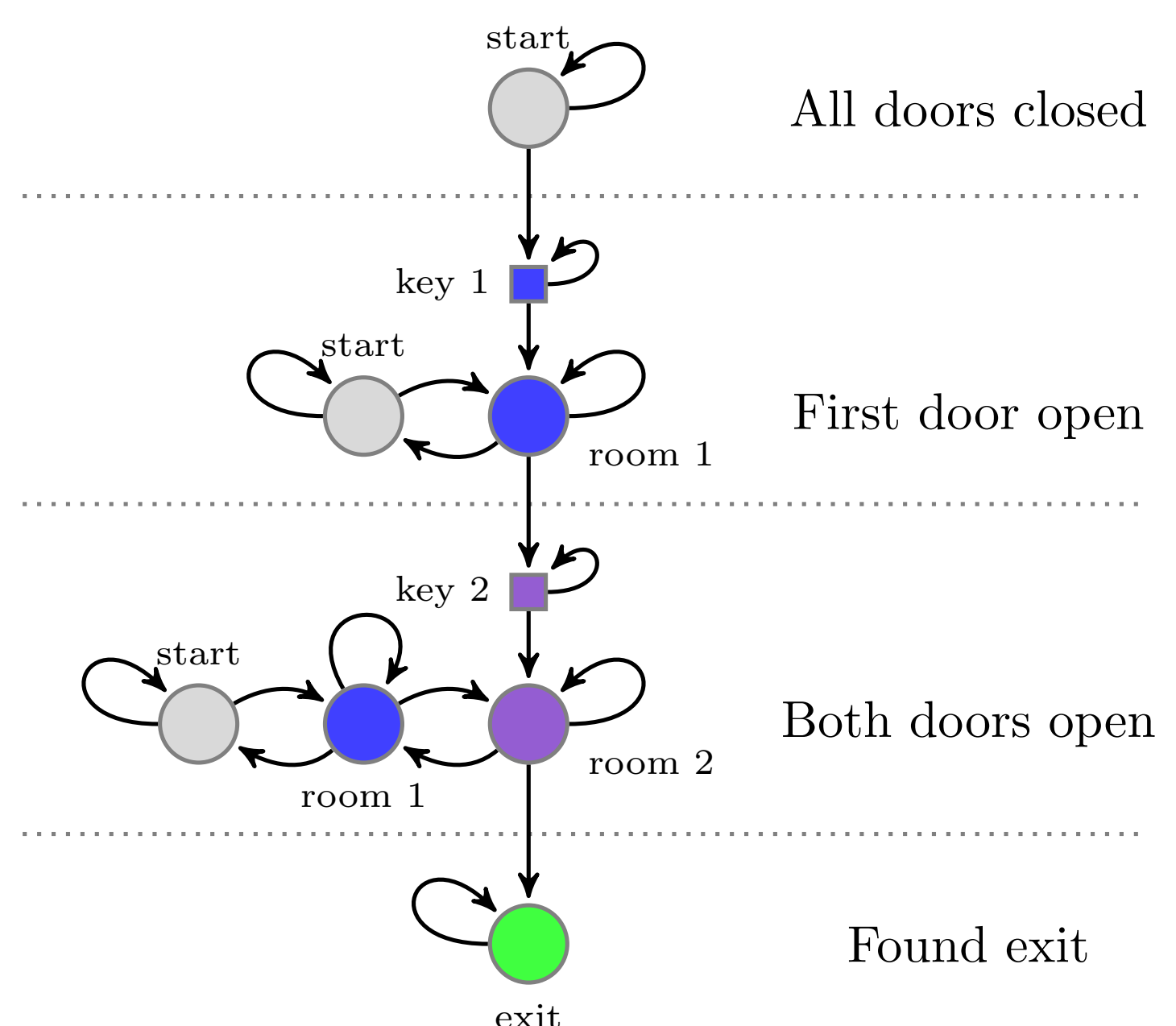
3. Method

- **Complexity measure:** Hitting time in the dependency graph between the subgoals in the environment.
- **Definition:** Let $L \in \mathbb{R}^{n \times n}$ be the Laplacian matrix of the *augmented* graph, then hitting time is the solution

$$Lx = b \quad \text{s.t. } x_t = 0, \quad x, b \in \mathbb{R}$$

where $b_s = 1, b_t = -1, b_k = 0 \quad \forall k \notin \{s, t\}$

Augmented subgoal graph



Hierarchical PPO (HiPPO)

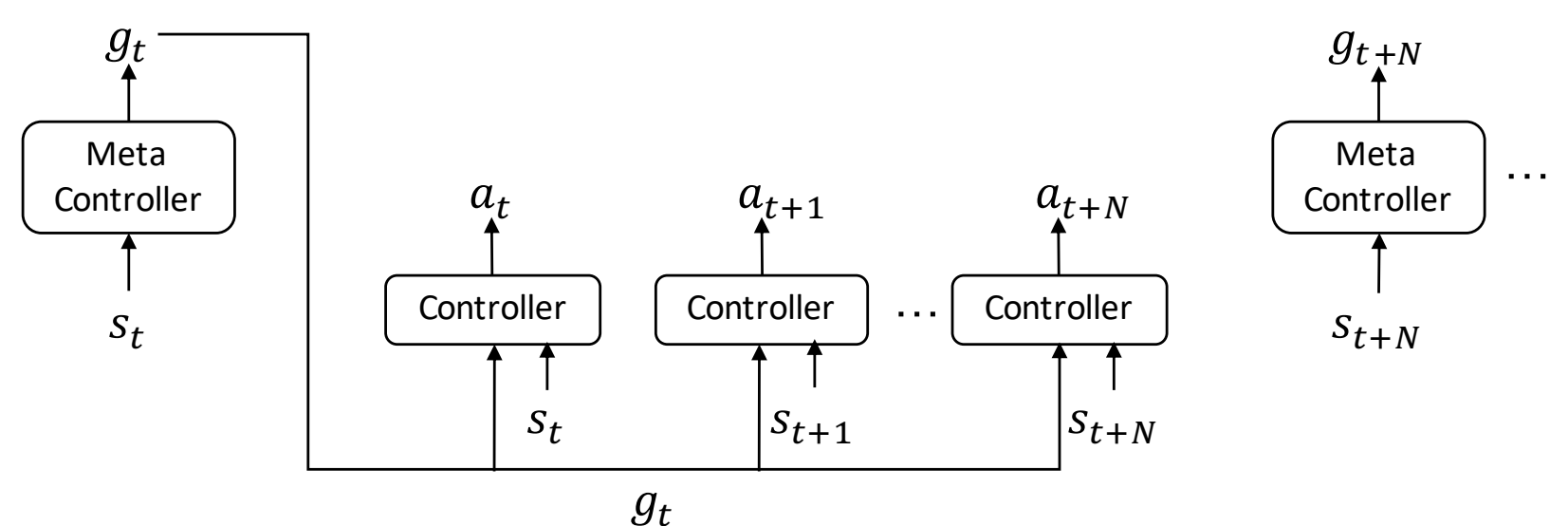


Table 1: Depth, width, and hitting time (HT) statistics computed for EscapeRoom environments (a)-(g).

	(a)	(b)	(c)	(d)	(e)	(f)	(g)
exit depth	2	2	4	2	2	4	6
graph width	1	2	1	2	3	2	1
HT (w/o drop-key)	8.4	12.1	15.1	13.1	13.9	29.2	27.5
HT (w/ drop-key)	16.5	25.2	39.5	27.5	26.7	86.1	82.5