NBDI: A Simple and Effective Termination Condition for Skill Extraction from Task-Agnostic Demonstrations

Myunsoo Kim*, Hayeong Lee*, Seong-Woong Shim, JunHo Seo, Byung-Jun Lee

Korea University, Gauss Lab Inc.

Motivation

➤ Recently, notable progress has been made in skill-based deep RL models, showing promising results in complex environments and robot manipula tions

➤ However, the use of fixed-length skills and the absence of appropriate termination conditions often restrict them from making decisions at critical decision points (e.g., crossroads), which can result in significant loss in performance
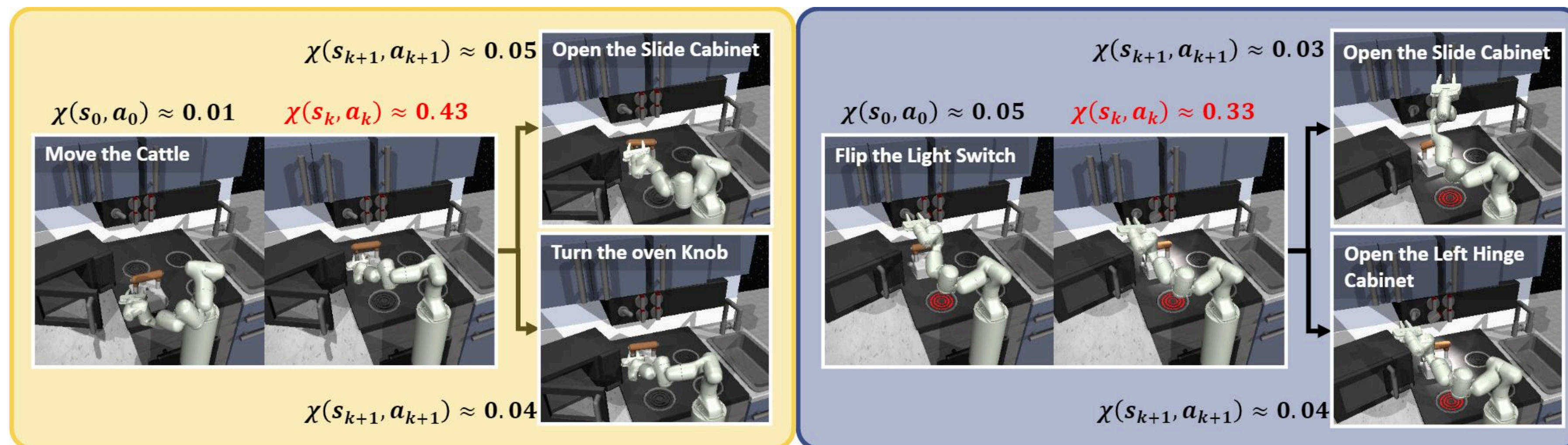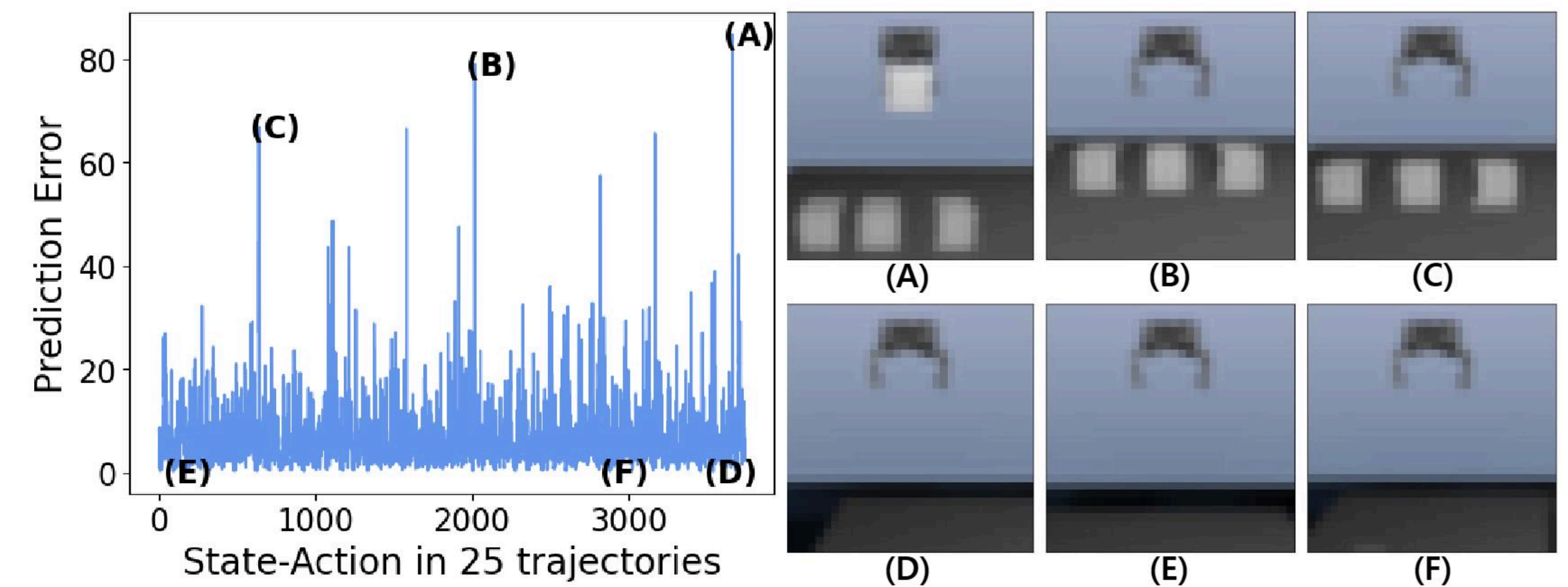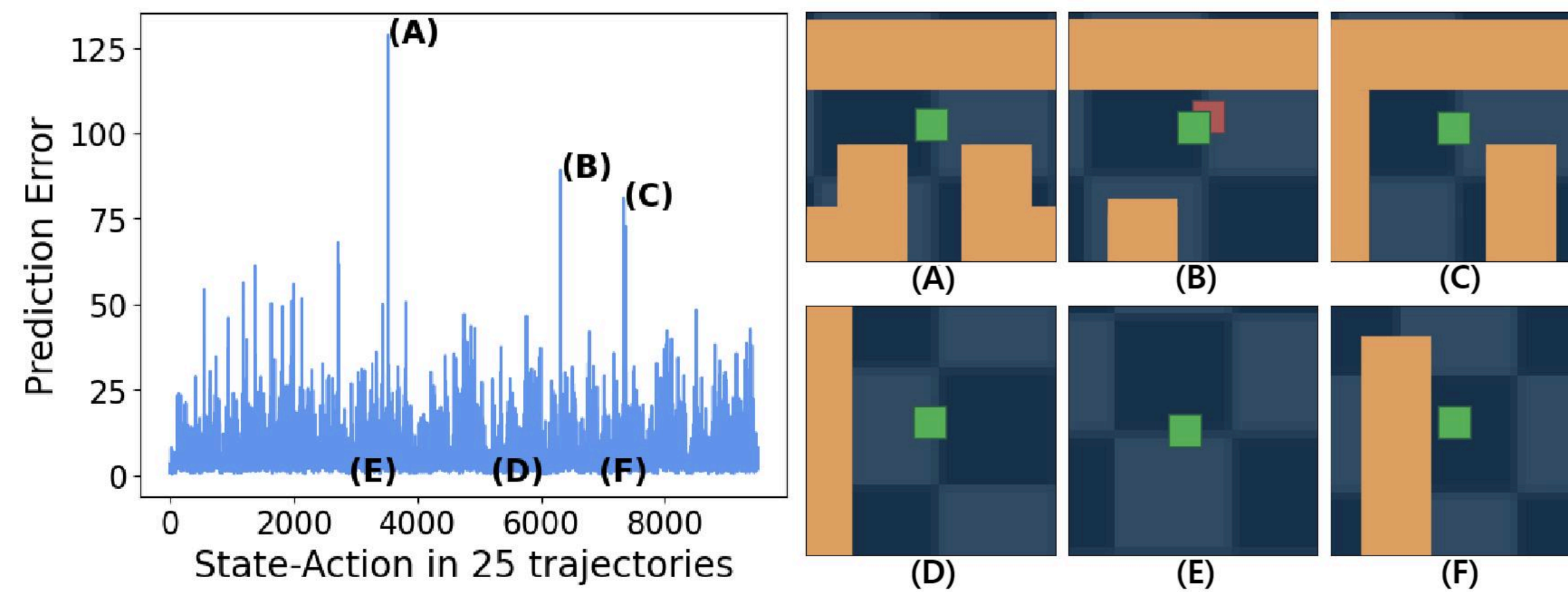
Introduction

➤ While maximizing skill length is advantageous in terms of temporal abstraction, extended skills can result in suboptimal behavior, especially when the skills are derived from task-agnostic trajectories.

➤ Such suboptimality of extended skills (or options) can be theoretically quantified using the termination improvement theorem (Sutton, 1998).

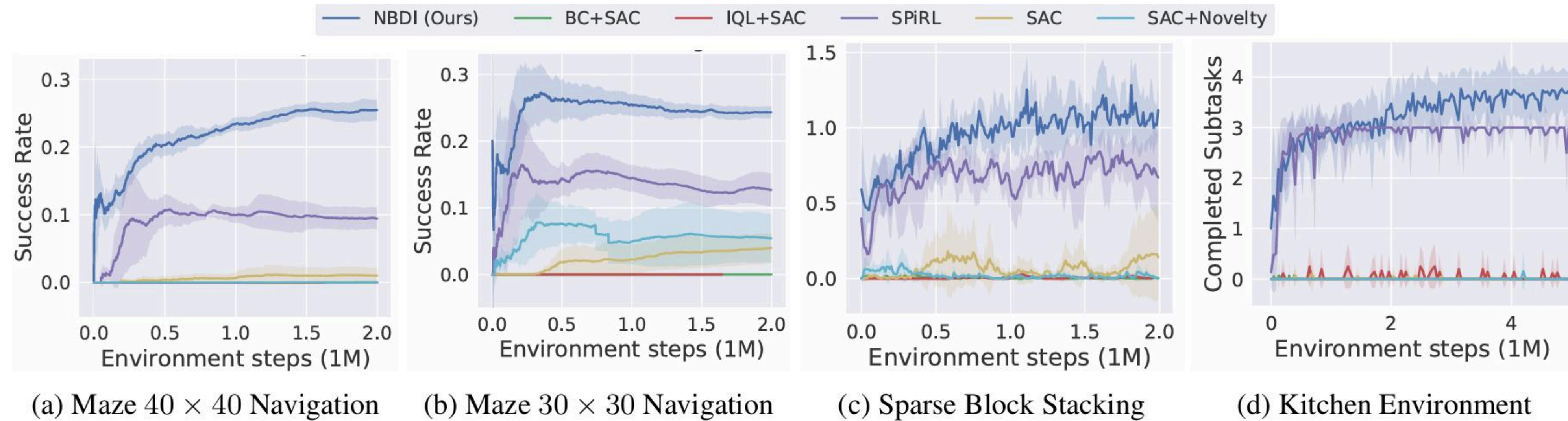**Termination Improvement**   **Conditional Action Novelty**   **State Novelty**
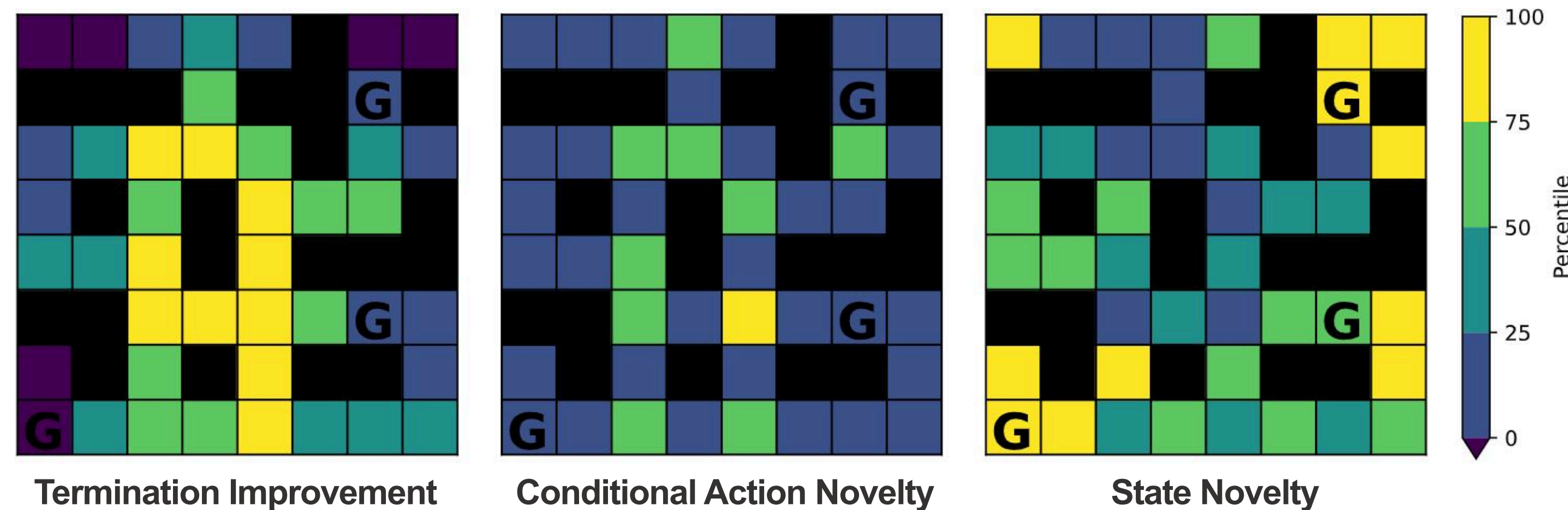
Method

➤ In this work, we propose to use state-action novelty to identify critical decision points for skill termination, which leads to the execution of variable-length skills.

➤ In particular, we use **intrinsic curiosity module (ICM)** (Pathak et al., 2017) as our state-action novelty estimator
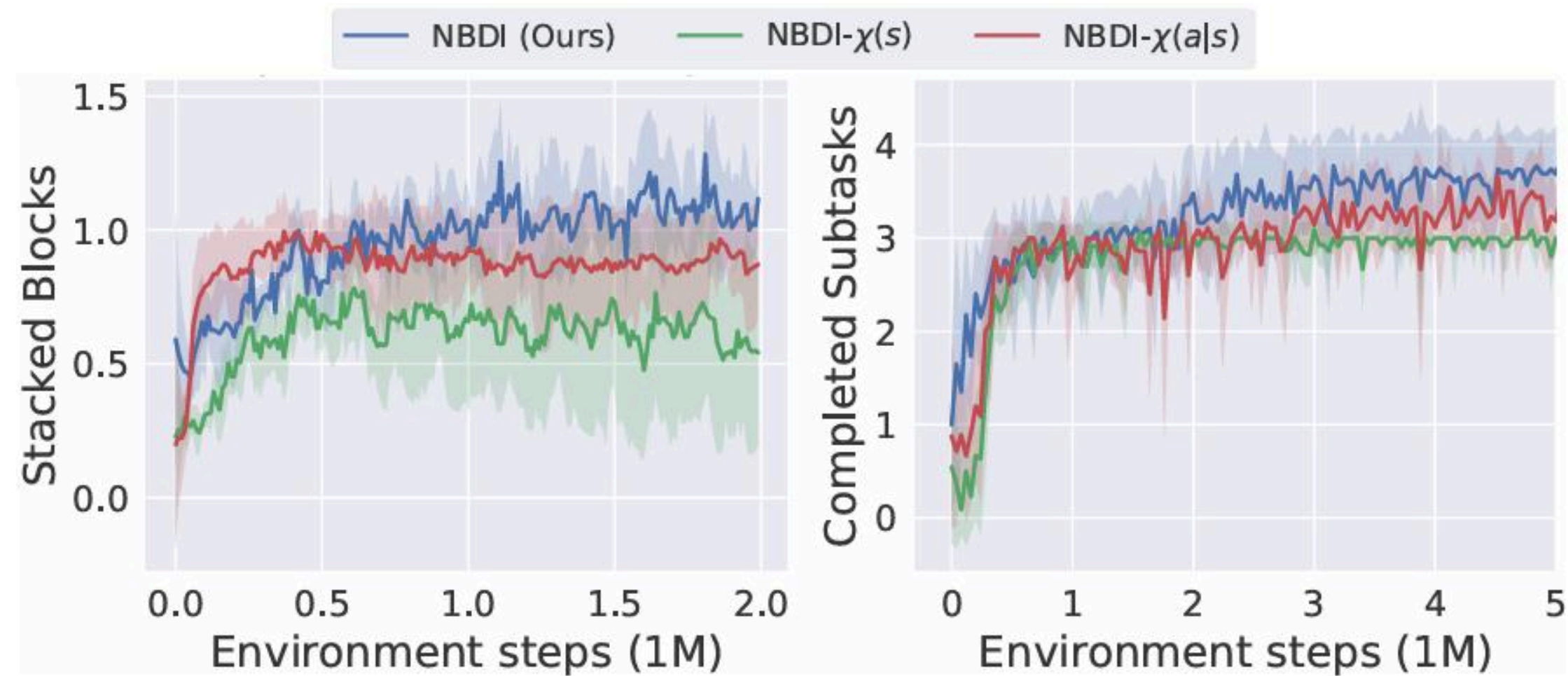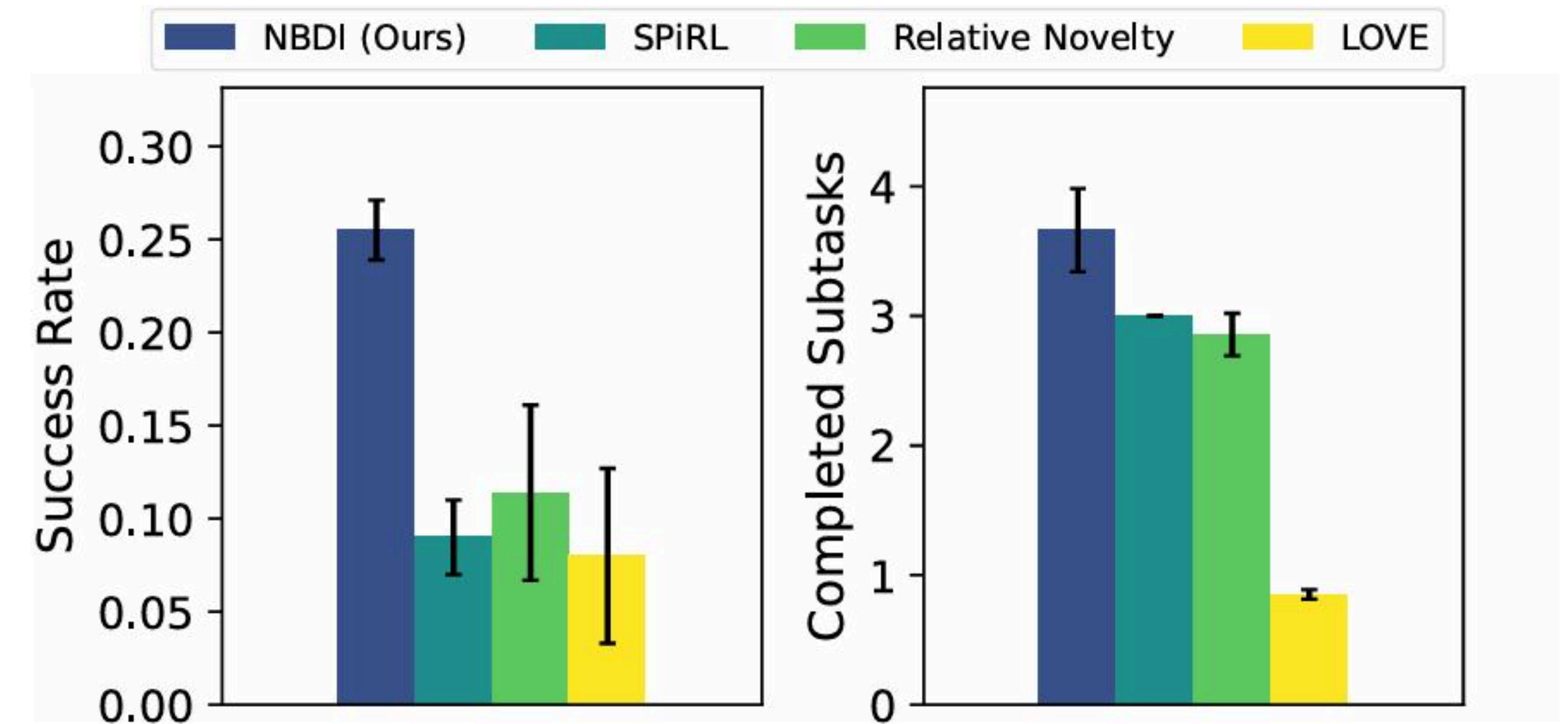
Experiments

| Legend: NBDI (Ours) — BC+SAC — IQL+SAC — SPiRL — SAC — SAC+Novelty |

(a) Maze 40 × 40 Navigation (b) Maze 30 × 30 Navigation (c) Sparse Block Stacking (d) Kitchen Environment

| ENVIRONMENT	SAC	SPiRL	NBDI	IMPROVEMENT OVER SPiRL(%)
MAZE 30x30 (*Success rate*)	0.04±0.03	0.13±0.03	**0.24**±0.01	84.62
MAZE 40x40 (*Success rate*)	0.01±0.01	0.09±0.02	**0.25**±0.02	177.78
SPARSE BLOCK STACKING (*Stacked Blocks*)	0.14±0.28	0.67±0.29	**1.12**±0.16	67.16
KITCHEN (*Completed Subtasks*)	0.0±0.0	3.0±0.0	**3.67**±0.43	22.33

Experiments

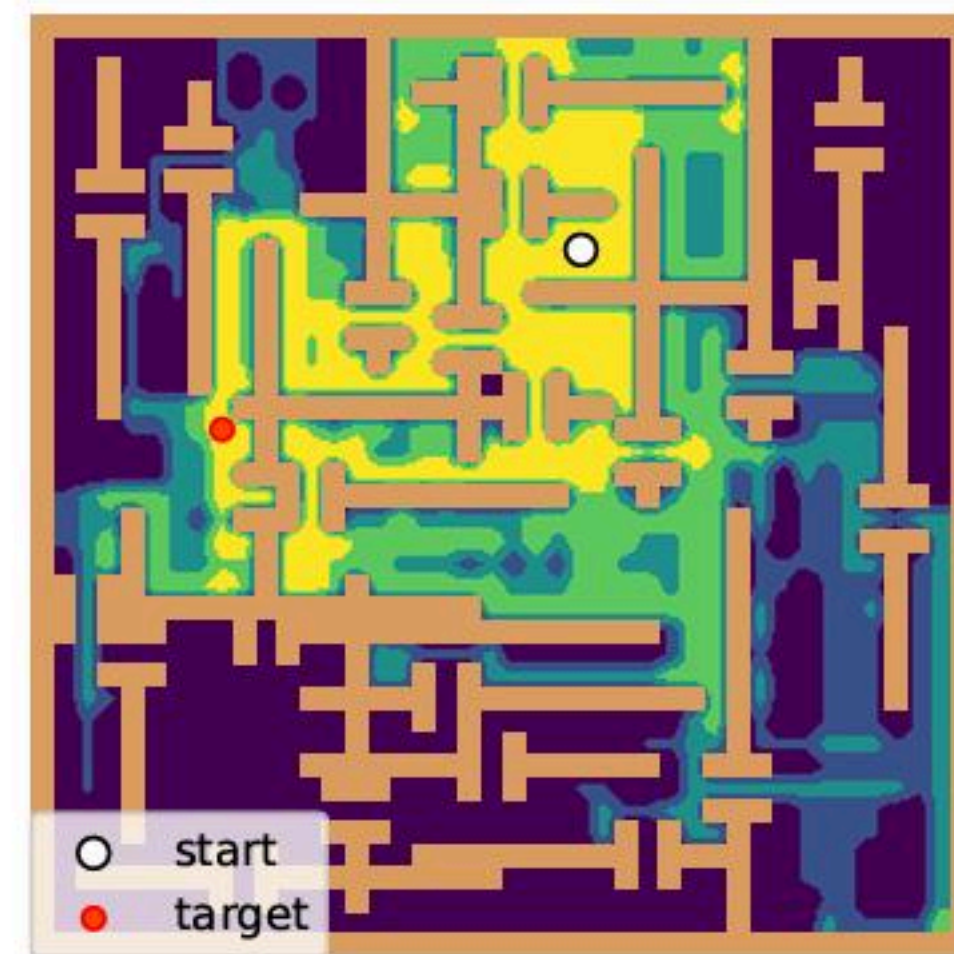(a) Comparison with other terminating criteria in sparse block stacking (left) and kitchen (right) environments

(b) Comparison with other variable skill length methods in 40×40 maze (left) and kitchen (right) environments
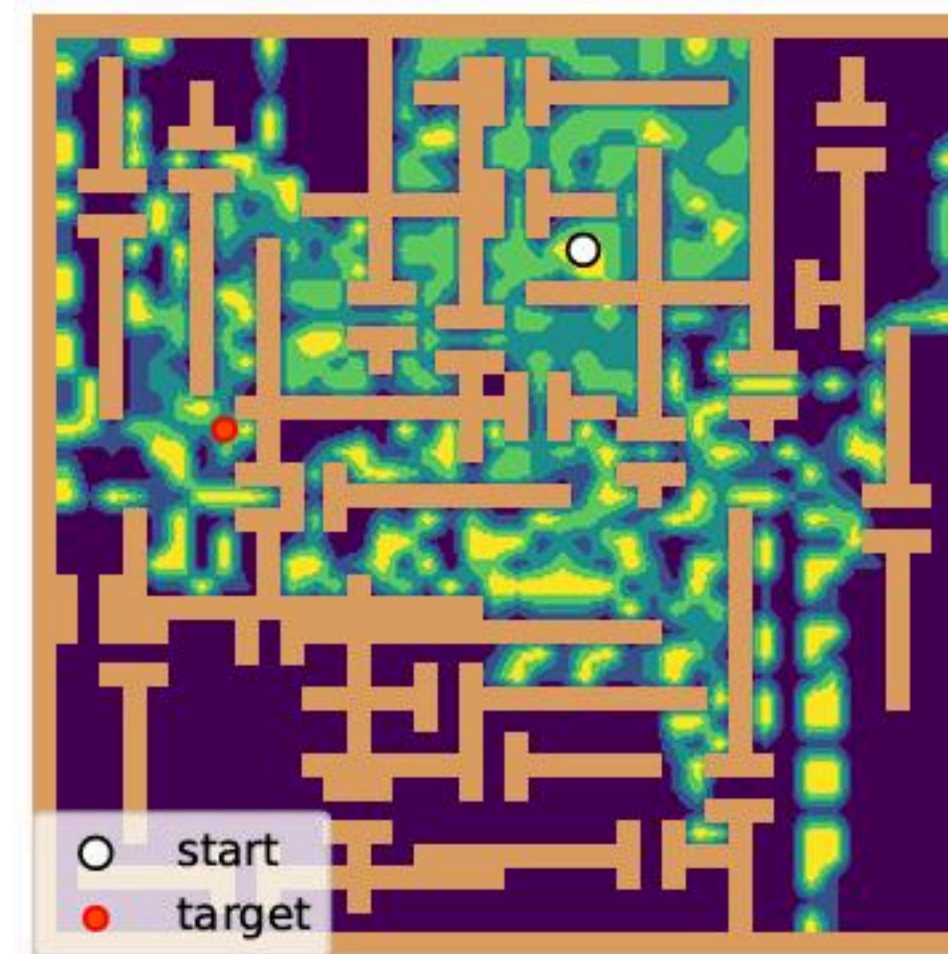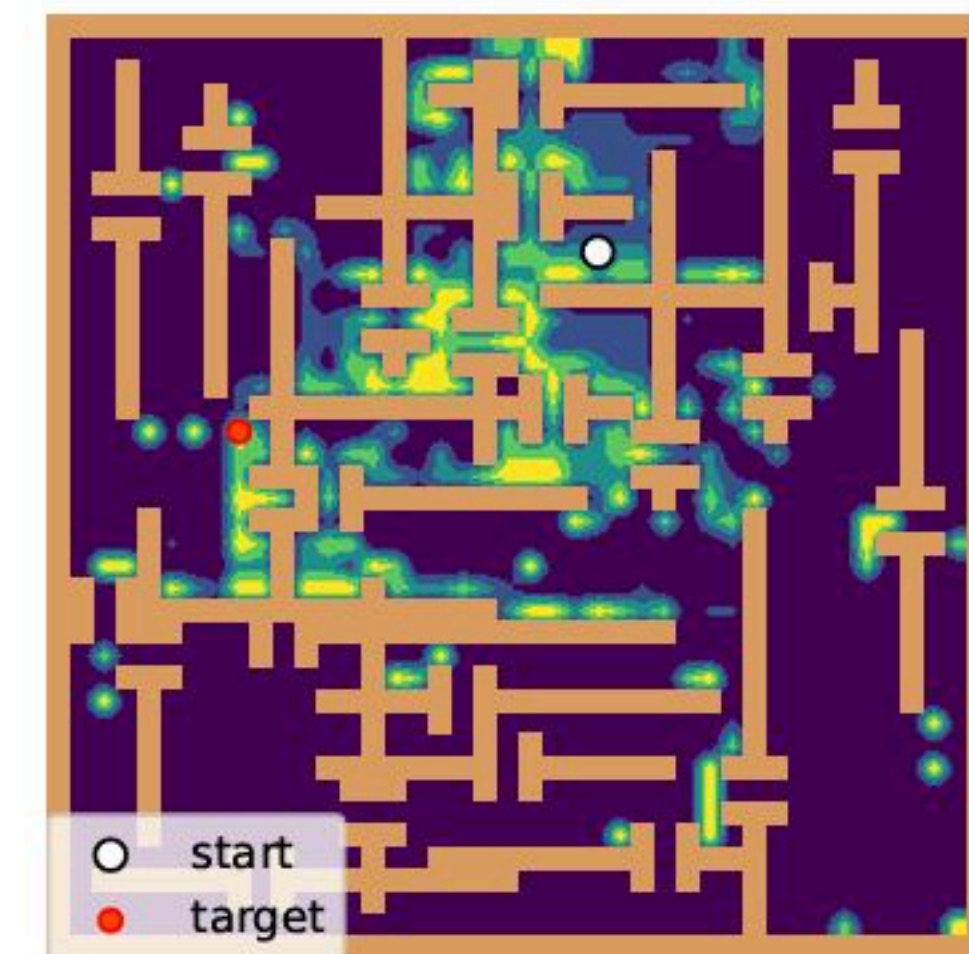
Experiments
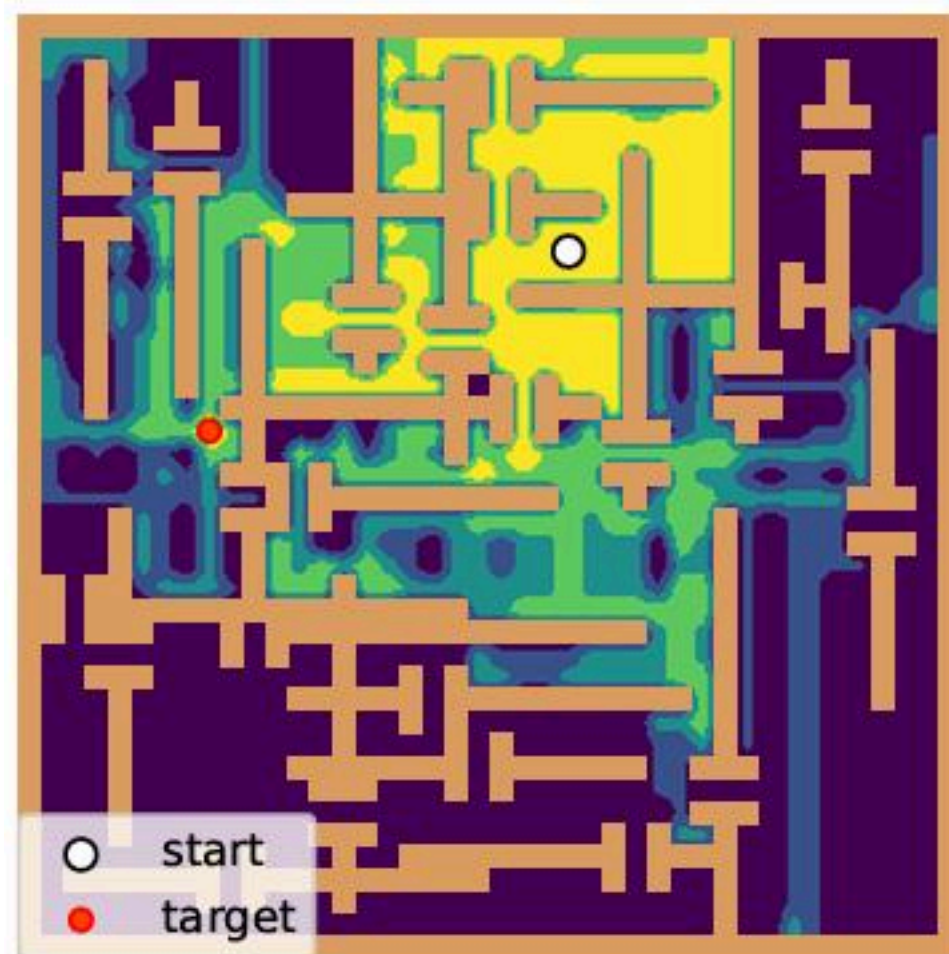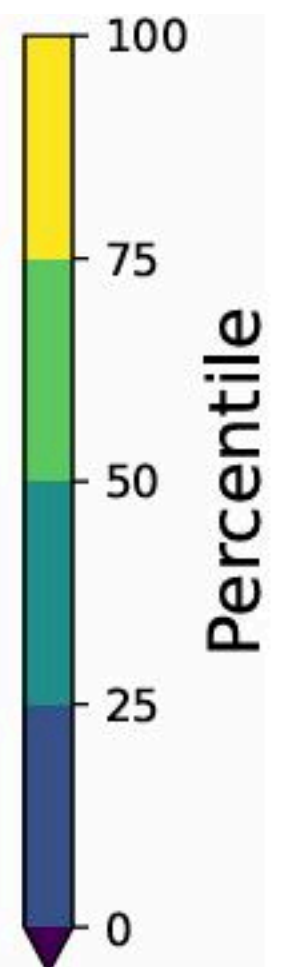

(a) Visitation(SPiRL)   (b) Visitation(NBDI)   (c) Skill Terminat.(SPiRL)   (d) Skill Terminat.(NBDI)

Thank you