

Maze Shaping Report: Manhattan PBRS with Revisit-Termination (v1)

February 13, 2026

1 Setup

This experiment changes the episode termination rule: an episode ends when either (i) the agent reaches goal, or (ii) the agent revisits any previously visited state in the same episode.

Core settings:

- Potential: Manhattan PBRS
- Discount: $\gamma = 0.99$
- Rewards: step = 0.0, goal = 1.0
- Episodes: 500, Runs: 12
- Validation: every 25 episodes (plus episode 1), 30 rollouts

2 Curves

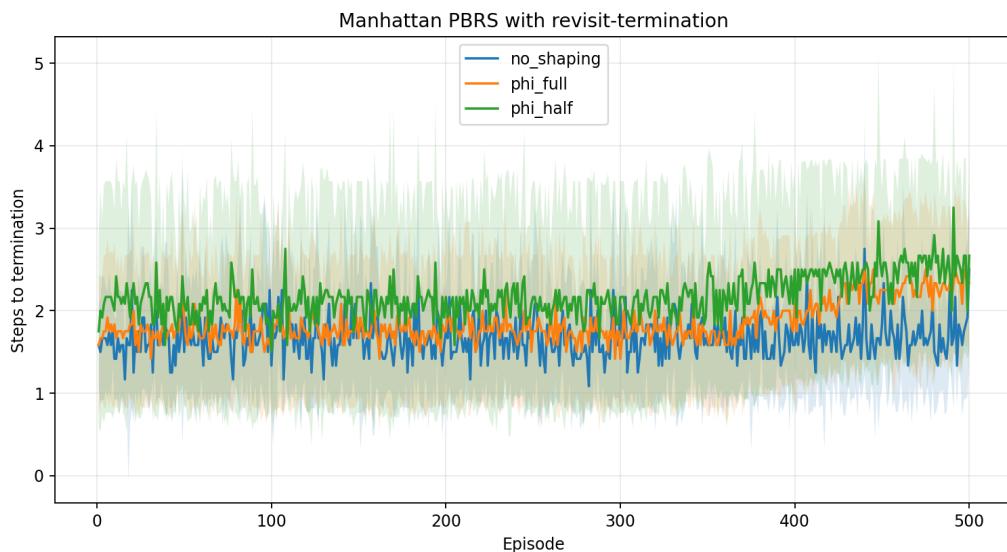


Figure 1: Training curve (steps to termination).

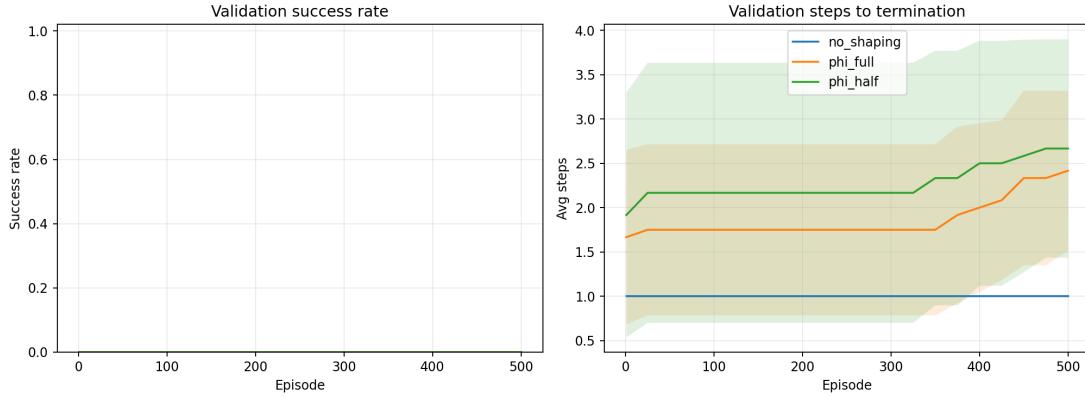


Figure 2: Validation success rate and validation average steps to termination.

Condition	Train mean steps	Validation success rate	Validation mean steps
no_shaping	2.500000	0.0	1.000000
phi_half	2.666667	0.0	2.666667
phi_full	2.333333	0.0	2.416667

Table 1: Episode-500 summary under revisit-termination rule.

3 Final Metrics (Episode 500)

4 Interpretation

Under revisit-termination, episodes collapse to very short trajectories (roughly 1–3 steps), and no condition achieves goal-reaching validation success.

5 Artifacts

- Output root: `../../outputs/maze_shaping_revisit_termination_manhattan_v1`
- Files: `learning_curve.csv/png`, `validation_progress.csv/png`, `run_summary.json`