Maze Shaping Report: Manhattan PBRS with Revisit-Termination (v2 Long-Train)

February 13, 2026

1 Setup

This run keeps the same revisit-termination rule but increases the training budget from 500 episodes to 5000 episodes.

Core settings:

- Termination: goal reached or revisit of any previously visited state in same episode

- Potential: Manhattan PBRS

- Discount: $\gamma = 0.99$

- Rewards: step $= 0.0$, goal $= 1.0$

- Episodes: 5000, Runs: 12

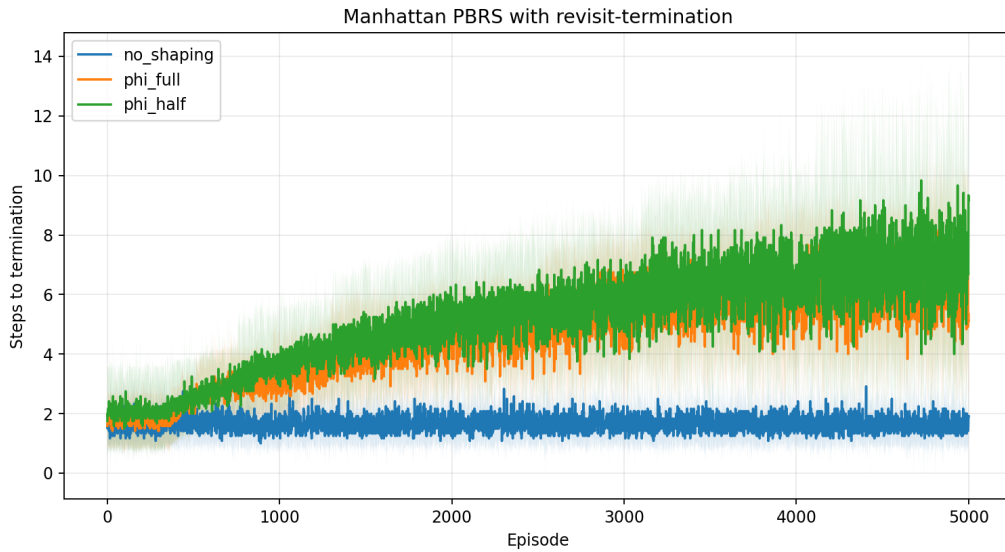- Validation: every 25 episodes (plus episode 1), 30 rollouts

2 Curves

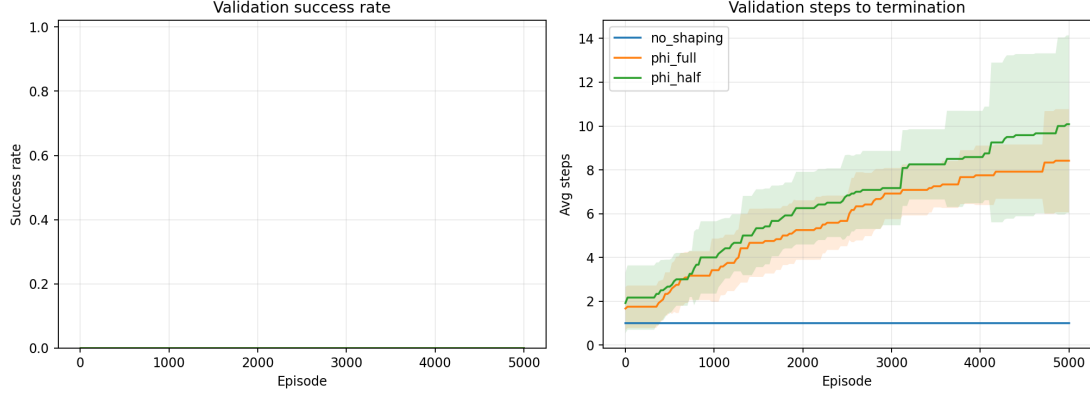Figure 1: Training curve (steps to termination).

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

Condition	Train mean steps	Validation success rate	Validation mean steps
no_shaping	1.916667	0.0	1.000000
phi_half	9.166667	0.0	10.083333
phi_full	5.333333	0.0	8.416667

Table 1: Episode-5000 summary under revisit-termination rule.

3 Final Metrics (Episode 5000)

4 Interpretation

Increasing training iterations alone did not recover success under this termination rule (all conditions remained at 0 validation success). This supports the view that the termination design, not only training budget, is the primary bottleneck.

5 Artifacts

- Output root: ../../outputs/maze_shaping_revisit_termination_manhattan_v2_longtrain

- Files: learning_curve.csv/png, validation_progress.csv/png, run_summary.json