

Maze Shaping Report: Manhattan PBRS with Revisit-Termination (v3 Very-Long)

February 13, 2026

1 Setup

This run keeps the revisit-termination rule and further increases the training budget to 20000 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: 20000, Runs: 12
- Validation: every 25 episodes (plus episode 1), 30 rollouts

2 Curves

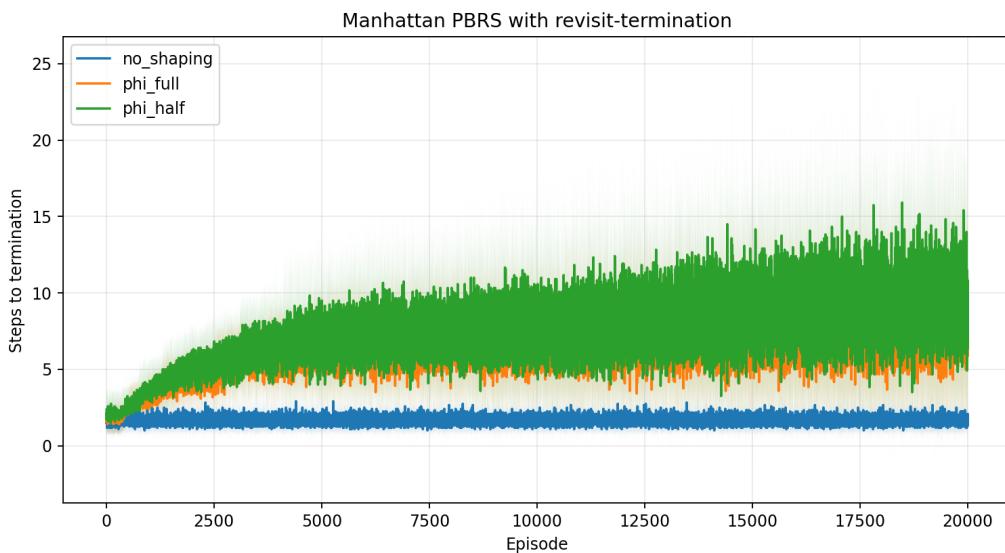


Figure 1: Training curve (steps to termination).

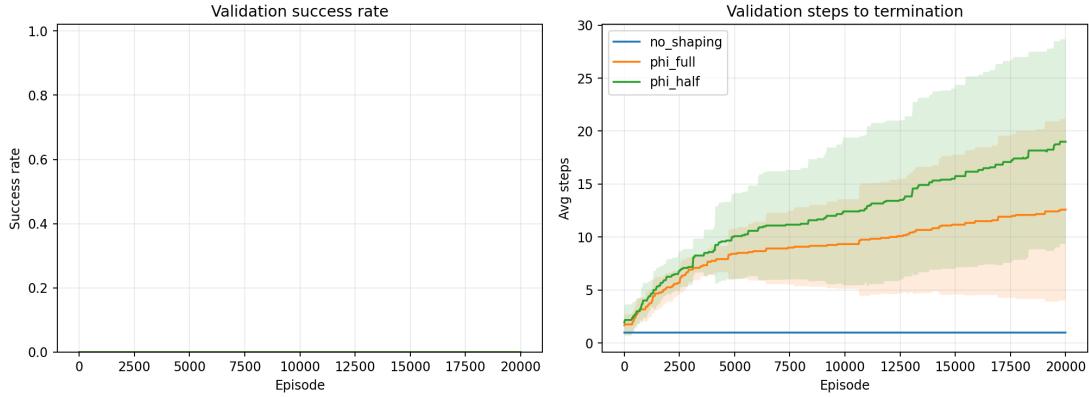


Figure 2: Validation success rate and validation average steps to termination.

Condition	Train mean steps	Validation success rate	Validation mean steps
no_shaping	1.666667	0.0	1.000000
phi_half	8.666667	0.0	19.000000
phi_full	6.083333	0.0	12.583333

Table 1: Episode-20000 summary under revisit-termination rule.

3 Final Metrics (Episode 20000)

4 Interpretation

Even with 20000 episodes, all conditions remain at zero validation success. PBRs conditions increase steps-to-termination but do not convert that into goal-reaching behavior.

5 Artifacts

- Output root: `../../../../outputs/maze_shaping_revisit_termination_manhattan_v3_verylong`
- Files: `learning_curve.csv/png`, `validation_progress.csv/png`, `run_summary.json`