LLM Maze Navigation Dashboard

Tool Overreliance Research Framework

Maze Size

8

Difficulty

Medium

Episodes

5

Agent Strategy

Adaptive

LLM Model (for LLM Solver)

OpenAI (o1-mini)

Noise Type

None (Perfect)

Tool Accuracy

Broken (0%) **100%** Perfect (100%)

Run Experiment

Load Sample

Tool Usage & Accuracy

8

Tool Queries

58.0%

Usage Rate

92.8%
Tool Accuracy

38.3%
Follow Rate

SUCCESS RATE

100.0%

5/5 episodes solved

AVERAGE STEPS

14

Min: 13 | Max: 17

PATH OPTIMALITY

7.1%

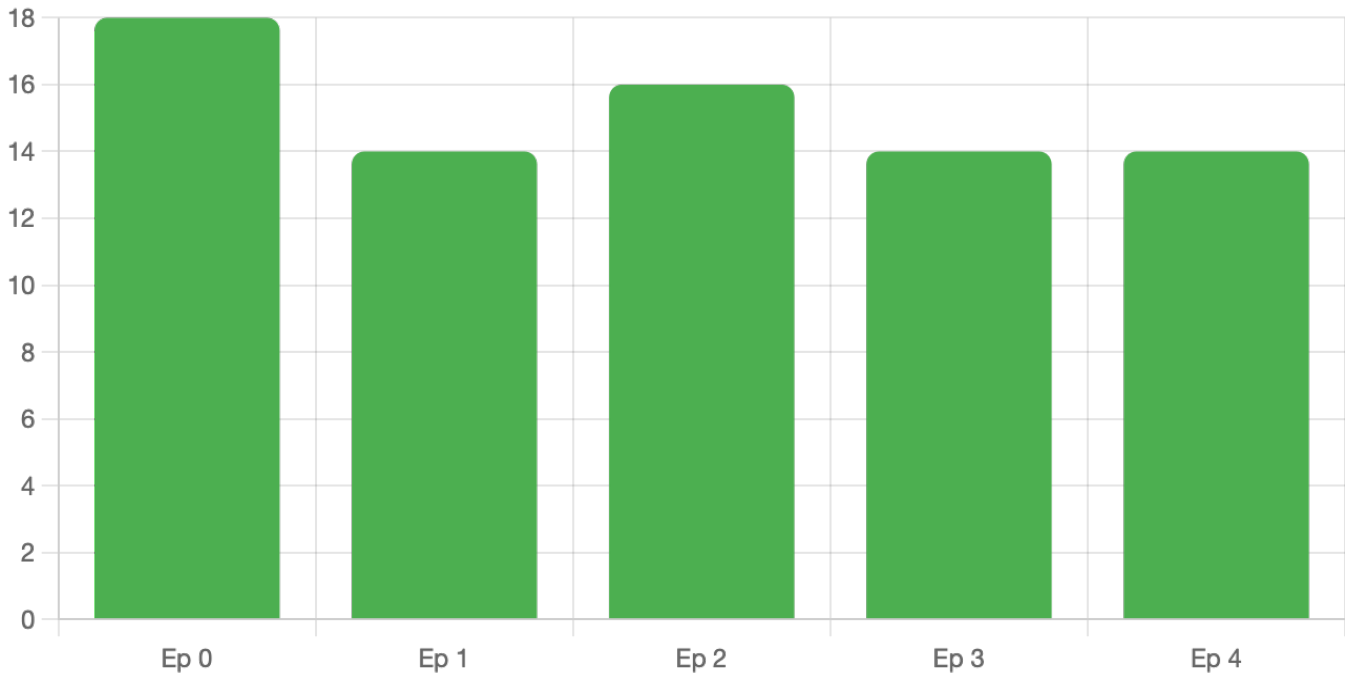
Efficiency vs optimal path

AVG FINAL DISTANCE

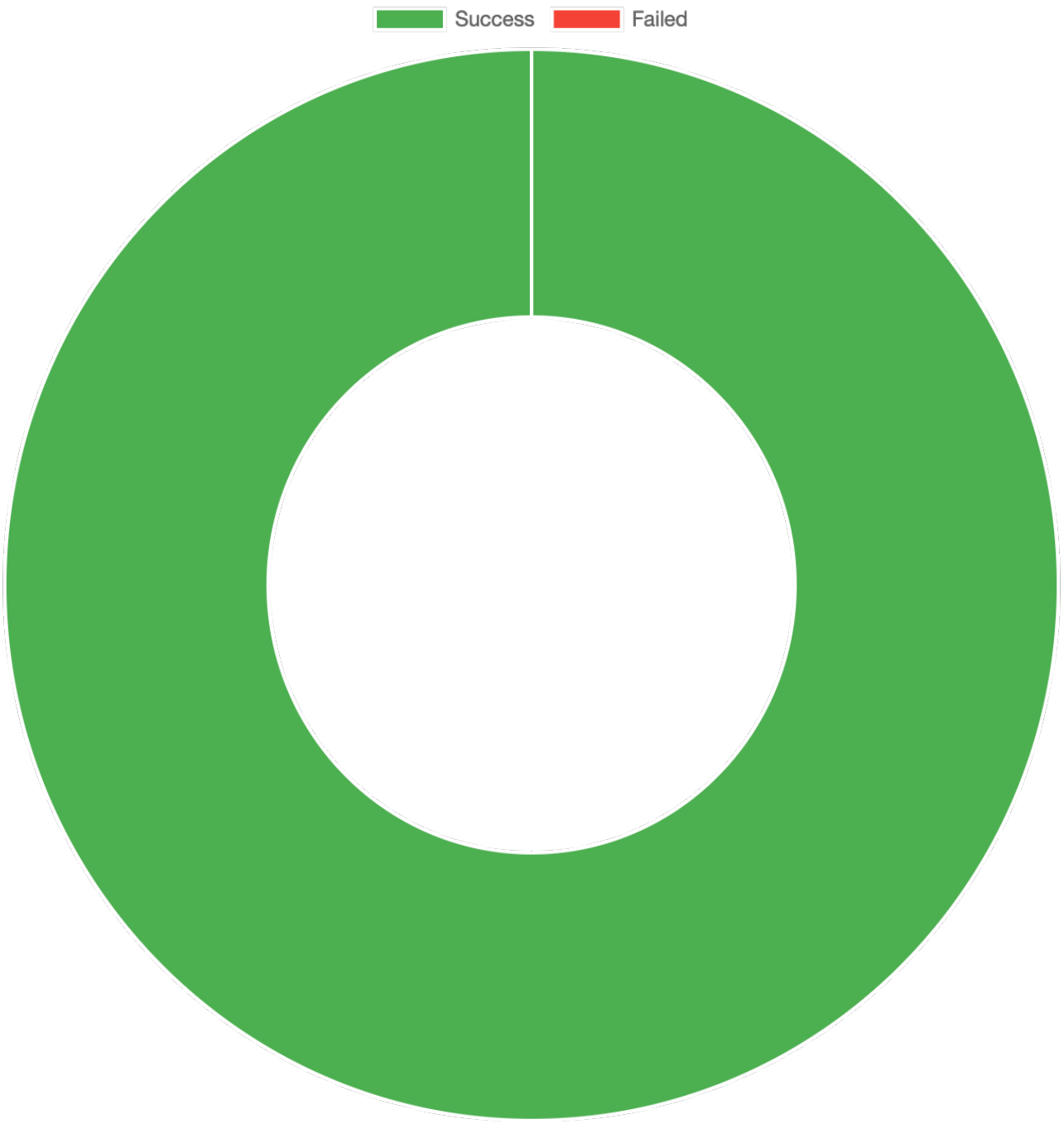
1

Steps from goal

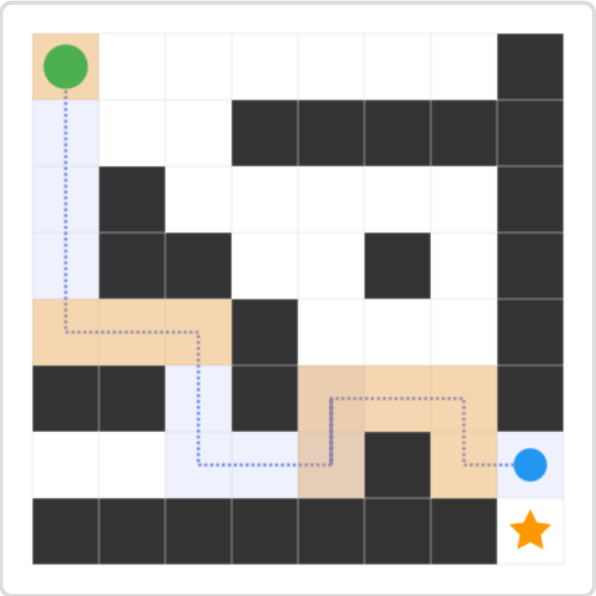
Episodes Performance



Metrics Overview



First Episode Path Visualization



Start Goal Agent Path

Wall (estimated) 🤖 LLM Used Tool

Episode Details

✓ OK	**Episode 0** Status: Success	Steps **18**	Tool Queries **9**	View LLM Decisions →
✓ OK	**Episode 1** Status: Success	Steps **14**	Tool Queries **8**	View LLM Decisions →
✓ OK	**Episode 2** Status: Success	Steps **16**	Tool Queries **9**	View LLM Decisions →
✓ OK	**Episode 3** Status: Success	Steps **14**	Tool Queries **8**	View LLM Decisions →

✓ OK

Episode 4

Status: Success

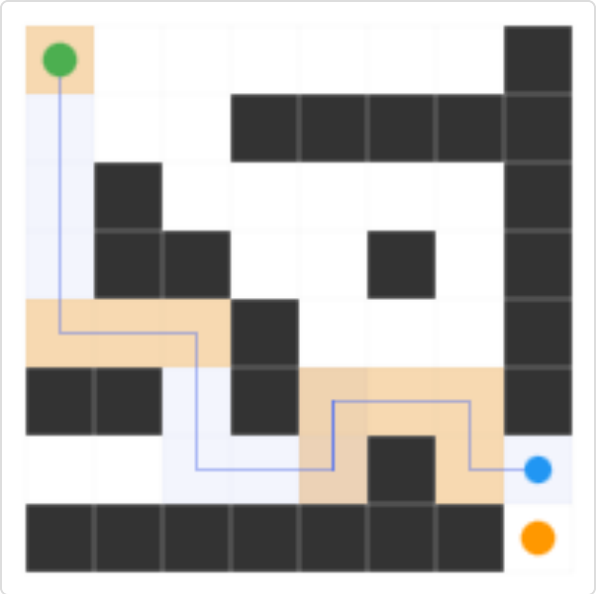
Steps **14**

Tool Queries **7**

View LLM Decisions →

📊 All Episodes Path Visualization

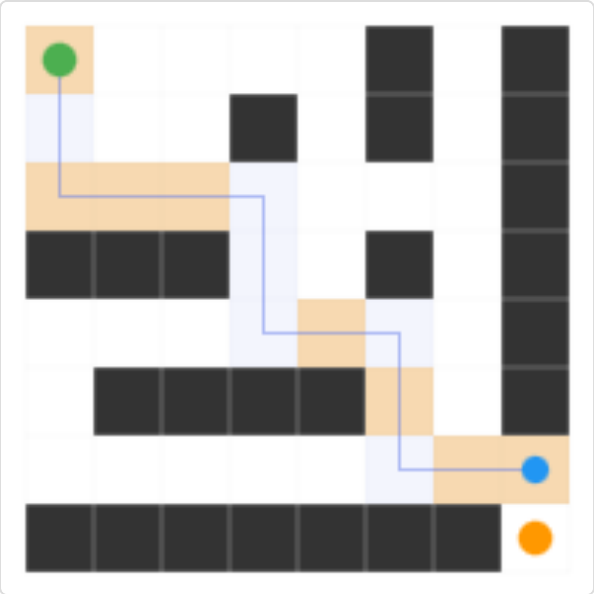
Episode 0 ✓



18 Steps

9 Tools

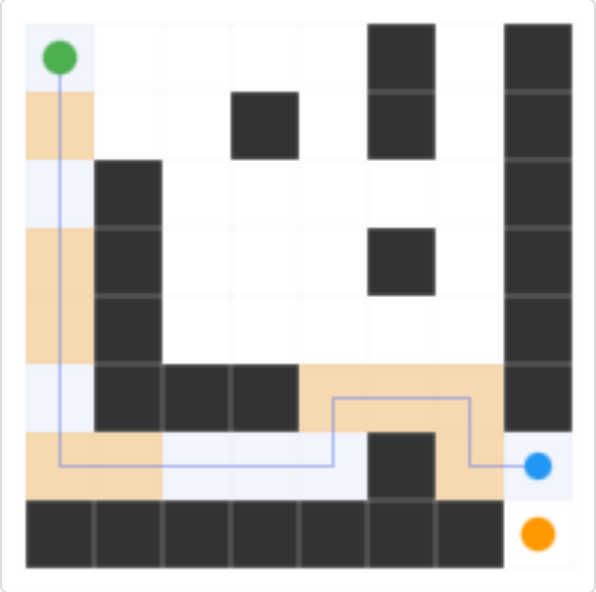
Episode 1 ✓



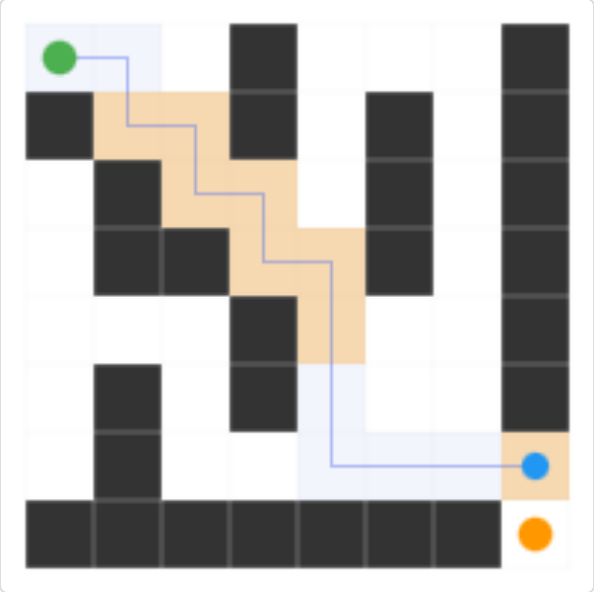
14 Steps

8 Tools

Episode 2 ✓



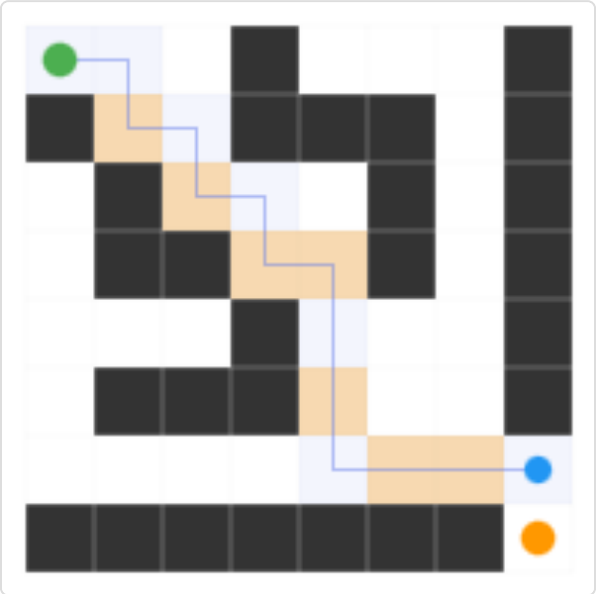
Episode 3 ✓



16	9
Steps	Tools

14	8
Steps	Tools

Episode 4 ✓



14	7
Steps	Tools