

STRATEGO

LLM FOR TEXT BASED GAMES



Stratego

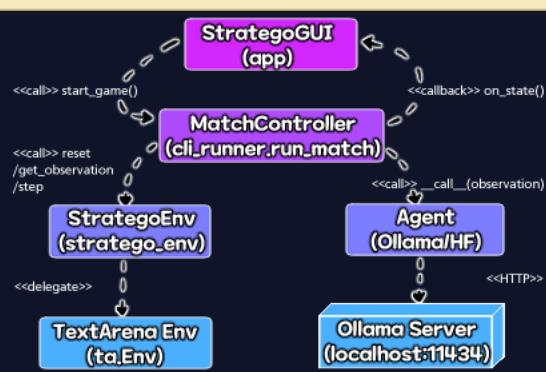


GAME MECHANICS

GAME BOARD VISUALIZATION



SYSTEM COMPONENT ARCHITECTURE



HOW TO PLAY

Hidden Information:

Enemy pieces are hidden unless revealed in combat.

Movement:

One piece per turn, up/down/left/right. Most move 1 tile; Scouts move any distance in a straight line if clear.

Lakes:

Impassable. No piece can enter or cross them.

Combat:

Attack an adjacent piece; higher rank wins (winner stays and advances). If the attacker loses, only the attacker is removed. The defender is briefly revealed, then hidden again.

Exceptions:

Spy beats Marshal only when attacking first. Only Miners can defuse Bombs; others lose to Bombs.

Win / Draw:

Capture the Flag or leave the opponent with no movable pieces to win. If only immovable pieces remain, it's a draw.

SPECIAL PIECES



MINER

These units can defuse enemy's Bombs



BOMB

Immovable Trap, which can be defused only by enemy's Miners



SPY

The weakest unit, but can defeat the enemy Marshal



SCOUT

Can move any number of squares in a straight line



FLAG

An immovable unit. If it's captured, you immediately lose the game

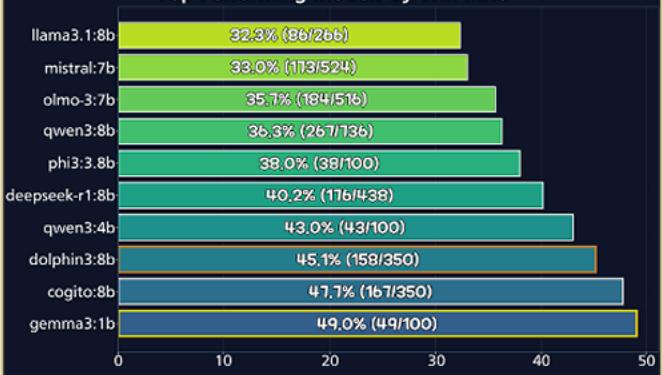
PERFORMANCE DASHBOARD

PERFORMANCE RESULTS - LLM BENCHMARKS

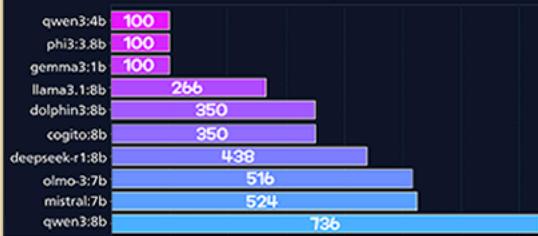
Win-Lose-Draw Comparison by Model



Top Performing Models by Win Rate



Total Games Played



Total models played: 10 Large Language Models

Total games played: 1740 Games and 3480 Data

Remarks: More games played means, the model played against other models more to get more precise data. Less games played means, the model played against other models less, so the data could not be precise.

