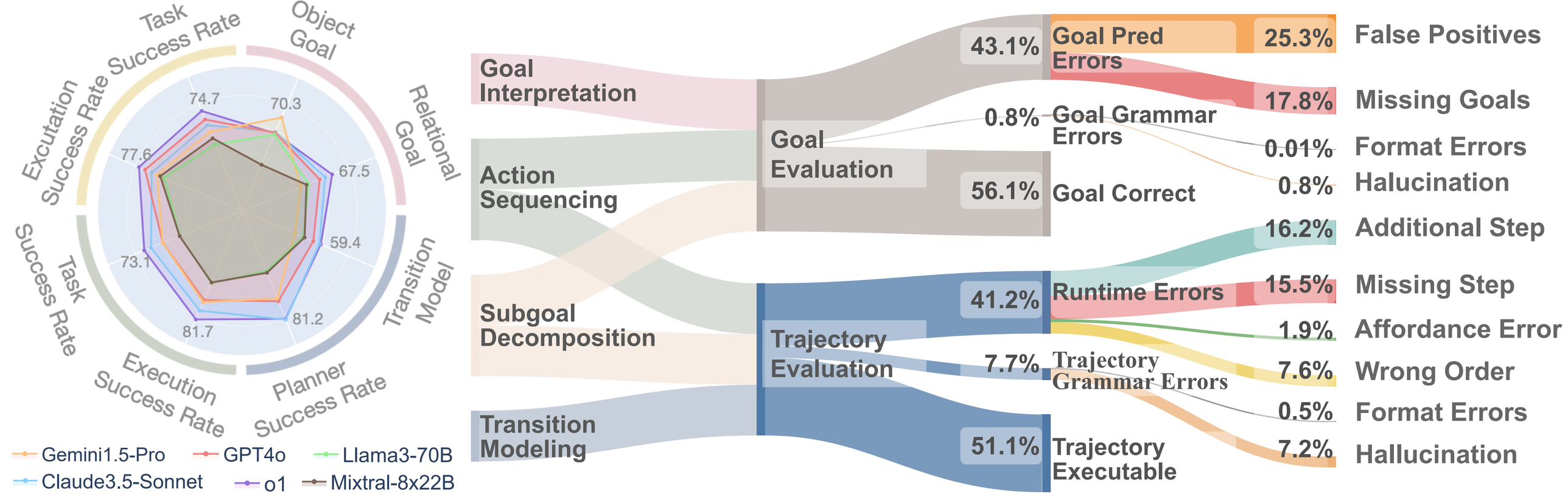
Task Success Rate
Object Goal
Relational Goal
Transition Model
Planner Success Rate
Execution Success Rate
Task Success Rate
Excutation Success Rate
74.7
70.3
67.5
59.4
81.2
81.7
73.1
77.6
Gemini1.5-Pro
GPT4o
Llama3-70B
Claude3.5-Sonnet
o1
Mixtral-8x22B
Goal Interpretation
Action Sequencing
Subgoal Decomposition
Transition Modeling
Goal Evaluation
Trajectory Evaluation
43.1%
0.8%
56.1%
41.2%
7.7%
51.1%
Goal Pred Errors
Goal Grammar Errors
Goal Correct
Runtime Errors
Trajectory Grammar Errors
Trajectory Executable
25.3%
17.8%
0.01%
0.8%
16.2%
15.5%
1.9%
7.6%
0.5%
7.2%
False Positives
Missing Goals
Format Errors
Halucination
Additional Step
Missing Step
Affordance Error
Wrong Order
Format Errors
Hallucination