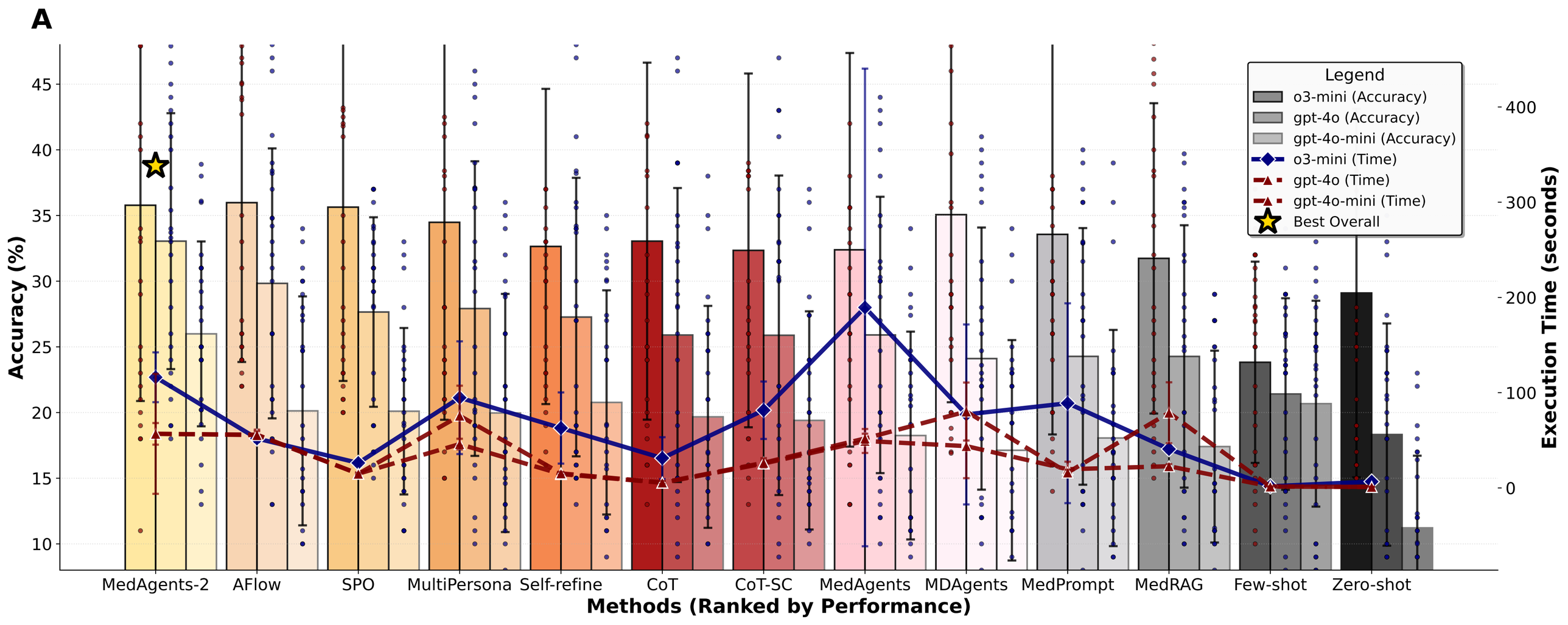
A
Legend
o3-mini (Accuracy)
gpt-4o (Accuracy)
gpt-4o-mini (Accuracy)
o3-mini (Time)
gpt-4o (Time)
gpt-4o-mini (Time)
Best Overall
Accuracy (%)
Execution Time (seconds)
Methods (Ranked by Performance)
MedAgents-2
AFlow
SPO
MultiPersona
Self-refine
CoT
CoT-SC
MedAgents
MDAgents
MedPrompt
MedRAG
Few-shot
Zero-shot