Dimension

Instuction Comprehension

Perception Reasoning

Generalization Planning

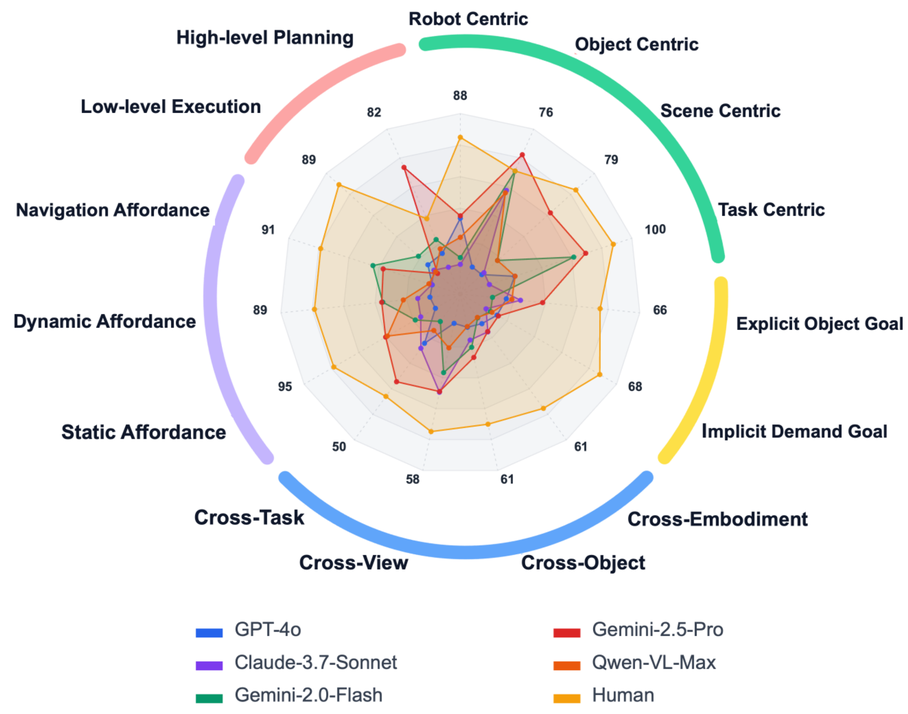
Affordance Prediction

Failure Ananlysis

Subdimension

Explicit Object Goal	Implicit Demand Goal		
Robotic-Centric	Object-Centric	Scene-Centric	Task-Centric
Cross-Embodiment	Cross-Object	Cross-View	Cross-Task
Static Affordance	Dynamic Affordance	Navigation Affordance	
Low-Level Execution Error	High-Level Planning Error		

Performance



1. Instuction Comprehension

Implicit Demand Goal:
Please help me tidy up the apple

Put the apple into the drawer

interpret human intention

gather task-relevant information

2. Perception Reasoning

Task-Related:
robotic type?
object attribute?
operation note?
spatial relation?

single arm
apple: rigid
drawer: closed
apple lower than drawer

3-1. Long-Horizon Planning

Question:
How to complete the goal?

1 open(drawer) 2 pick_up(apple)
3 place(apple, drawer) 4 close(drawer)

decompose long-horizen goal into substeps

3-2. Next-Step Planning

Question:
What's the next subgoal?

open(drawer)

adapt subgoal in closed-loop

3-3. Task State Estimation

Question:
Has current subgoal been completed?

monitor state in closed-loop

Yes

4-1. Static Affordance

Question:
What's the contact point?

[x1, y1]

refine subgoal into spatial aligned cues

4-2. Dynamic Affordance

Question:
What's the movement trajectory?

[x1, y1], [x2,y2], [x3, y3],[x4, y4]

refine subgoal into spatial aligned cues

monitor state in closed-loop

No

5. Failure Analysis

Question:
Is there any error? What type of error?

Yes, high-level error---missing steps

diagnose, analyze, adjust error