## Iterative Tasks: Success Rate (Bars) and Hard Efficiency (Lines) Zero\_shot Tasks: Success Rate (Bars) and Hard Efficiency (Lines) ab\_openai\_init\_refined (Success Rate) **-**O- ab\_openai\_init\_refined (Hard Efficiency) 1.0 o4-mini (Success Rate) -O- o4-mini (Hard Efficiency) 0.2

Tasks and Precision Levels

Tasks and Precision Levels

