



★ AgileThinker (Ours)

··· Max Performance of Single-Paradigm Methods

■ Reactive (V3)

+ Reactive (R1 + Budget Forcing)

▲ Planning (R1)

● Planning (R1 + Code-Policy)