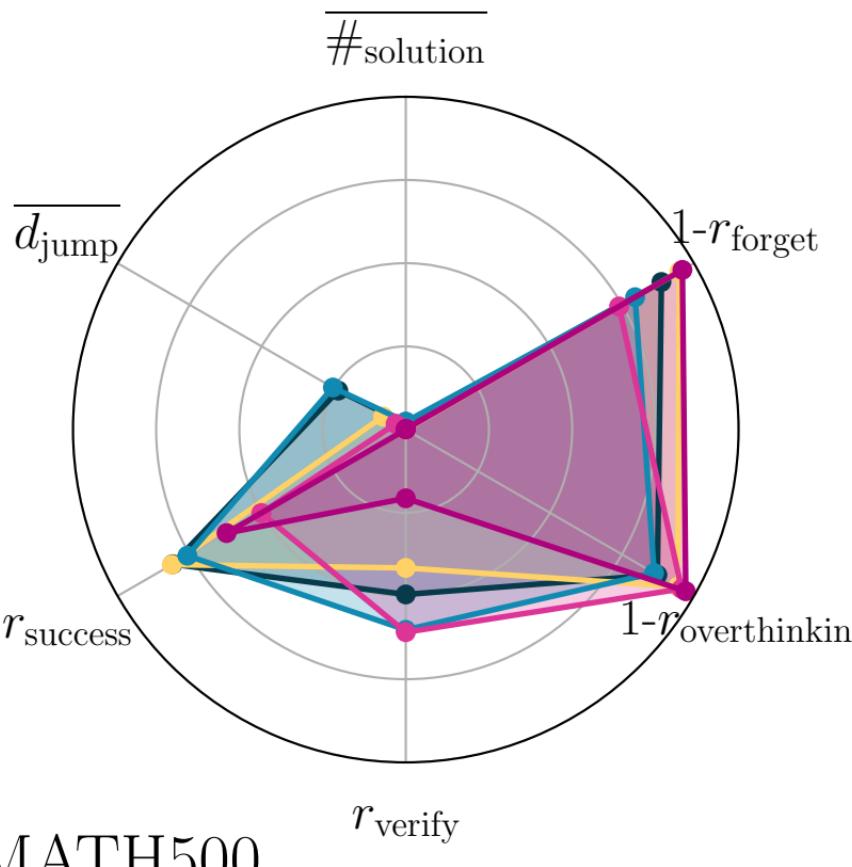
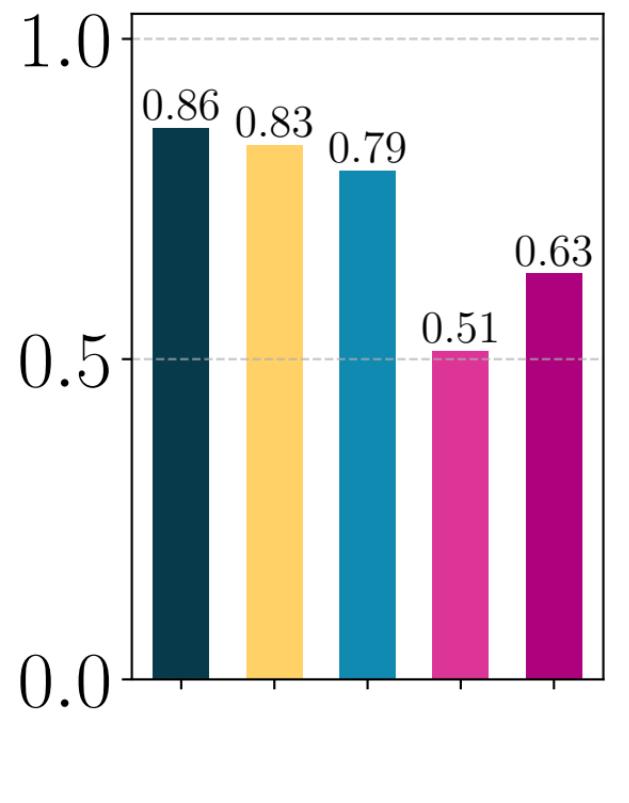


DeepSeek-R1 Grok 3 Mini Beta QwQ-32B Phi-4-reasoning-plus Claude 3.7 Sonnet

Pass@1



Pass@1

