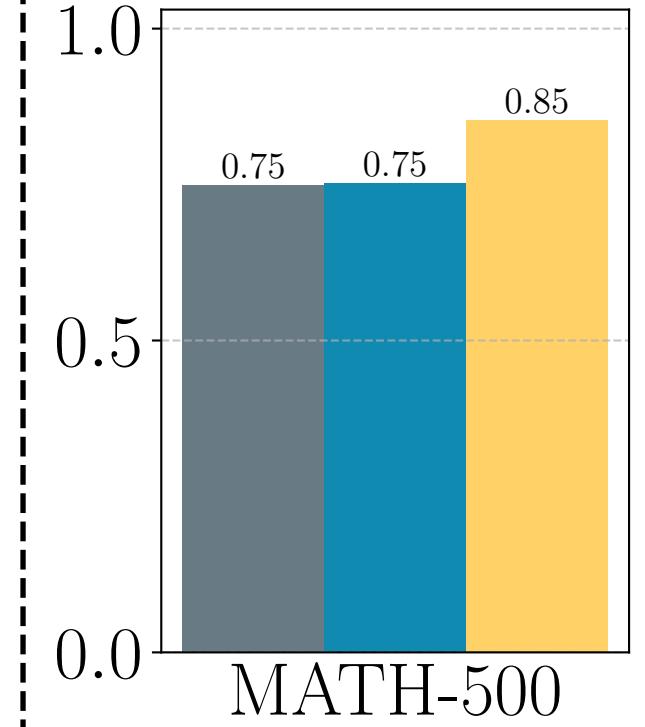


openrouter-deepseek/deepseek-r1-distill-qwen-32b

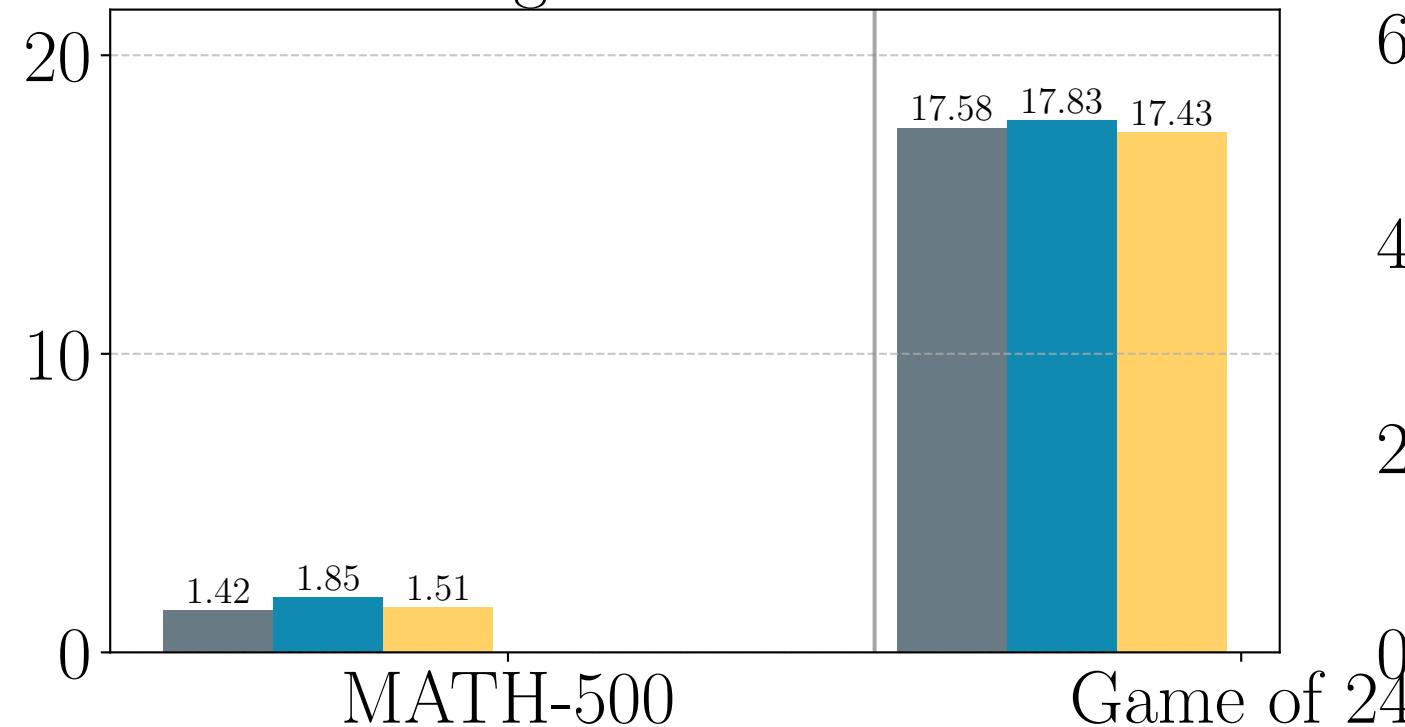
openrouter-qwen/qwq-32b

deepseek-ai/deepseek-reasoner

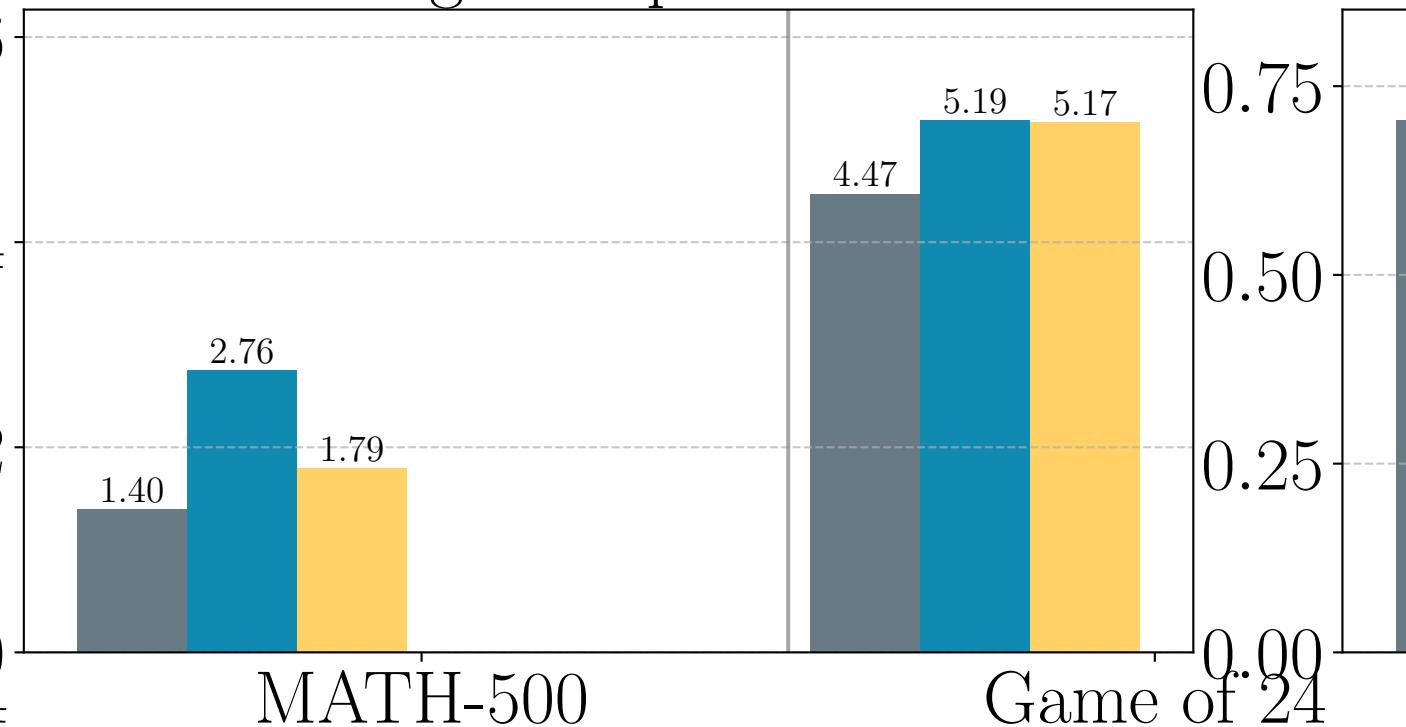
Pass@1



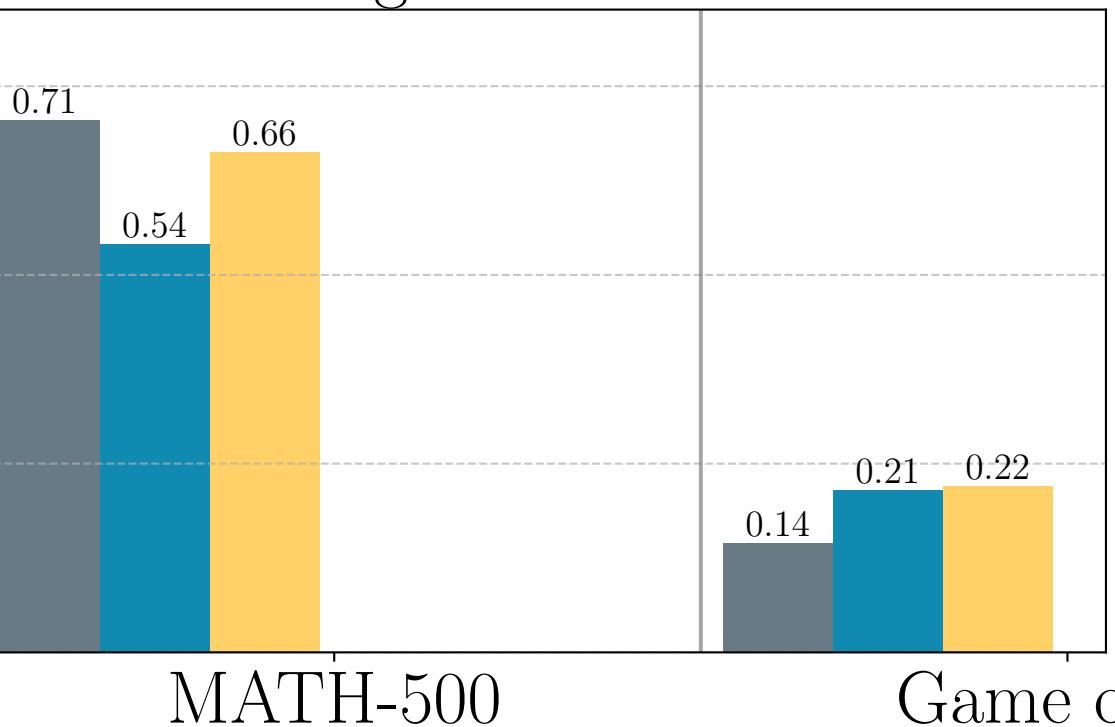
Average Solution Count



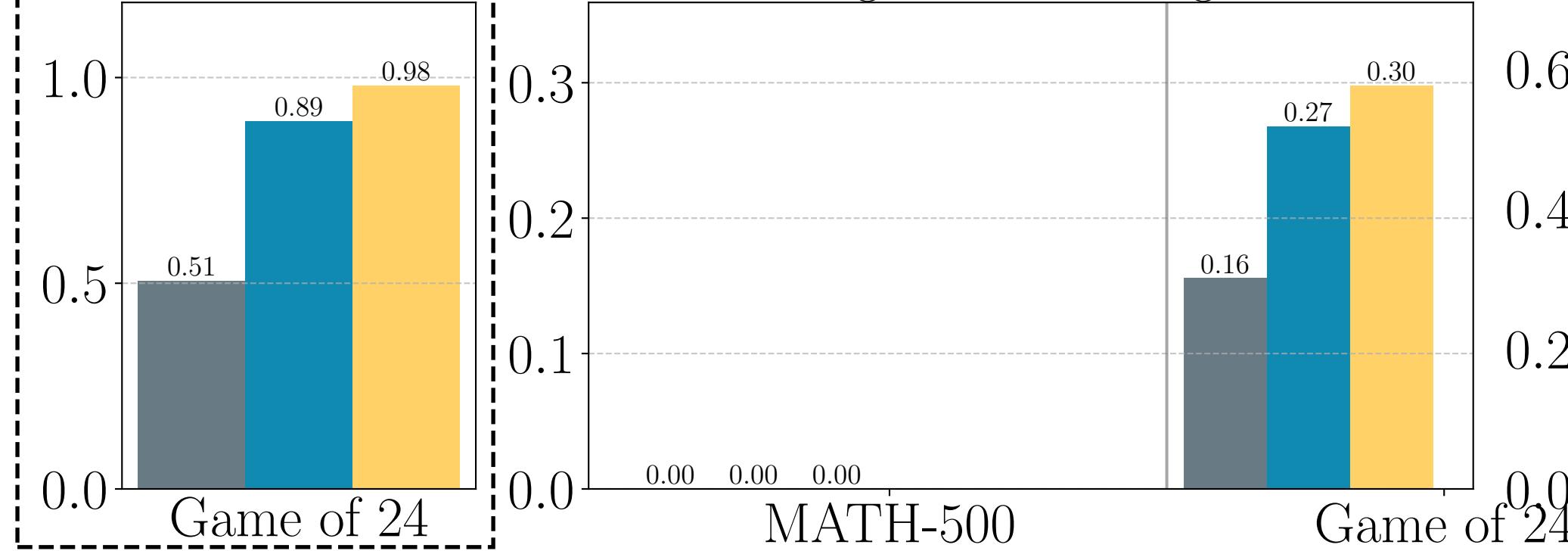
Average Jump Distance



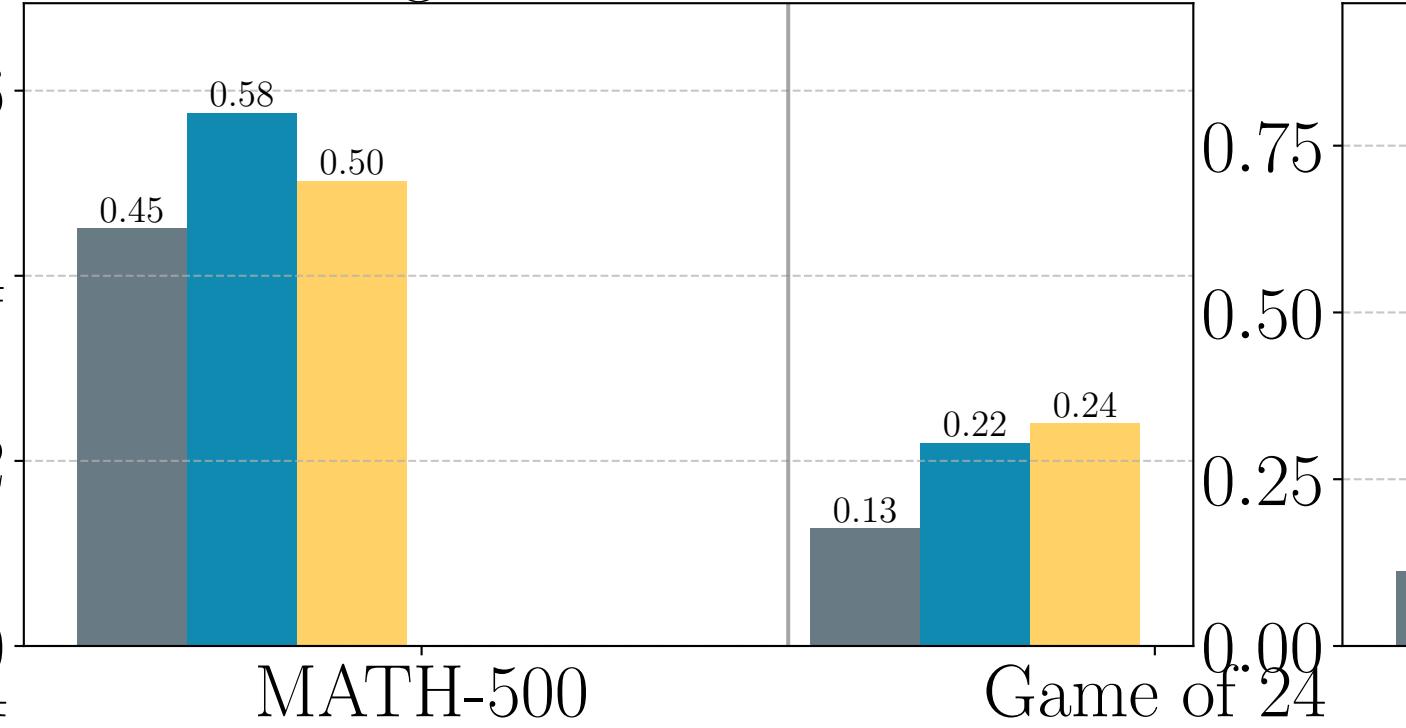
Average Success Rate



Average Overthinking Rate



Average Verification Rate



Forgetting Rate

