Total Plausible@1 vs "Greedy" Plausible@10 0.60 claude-3-5-sonnet-20241022 (2024-10-22) 0.55 Total "Greedy" Plausible@10 gpt-4o-2024-08-06 (2024-08-06) grok-2-1212 (2024-12-12) 0.50 claude-3-5-sonnet-20240620 (2024-06-20) gpt-4o-2024-11-20 (2024-11-20) mistral-large-24blek-v2.5 mistral-(2024-11) 18)24-09-05) (2024-07-24) llama-3.1-405b-instruct 0.45 qwen-2.5-coder-32b-instruct 7-23) gemini-1.5-pro-002 (2024-09-24) (2024-11-12) llama-3.3-70b-instruct (2024-12-06) qwen-2.5-72b-instruct codestral-2405 (2024-09-19)(2024-05-23) gemini-1.5-pro-001 (2024-05-24) 0.40 llama-3.1-nemotron-70b-instruct (2024-10-15)

0.30

Total Plausible@1

0.35

0.40

0.20

0.25