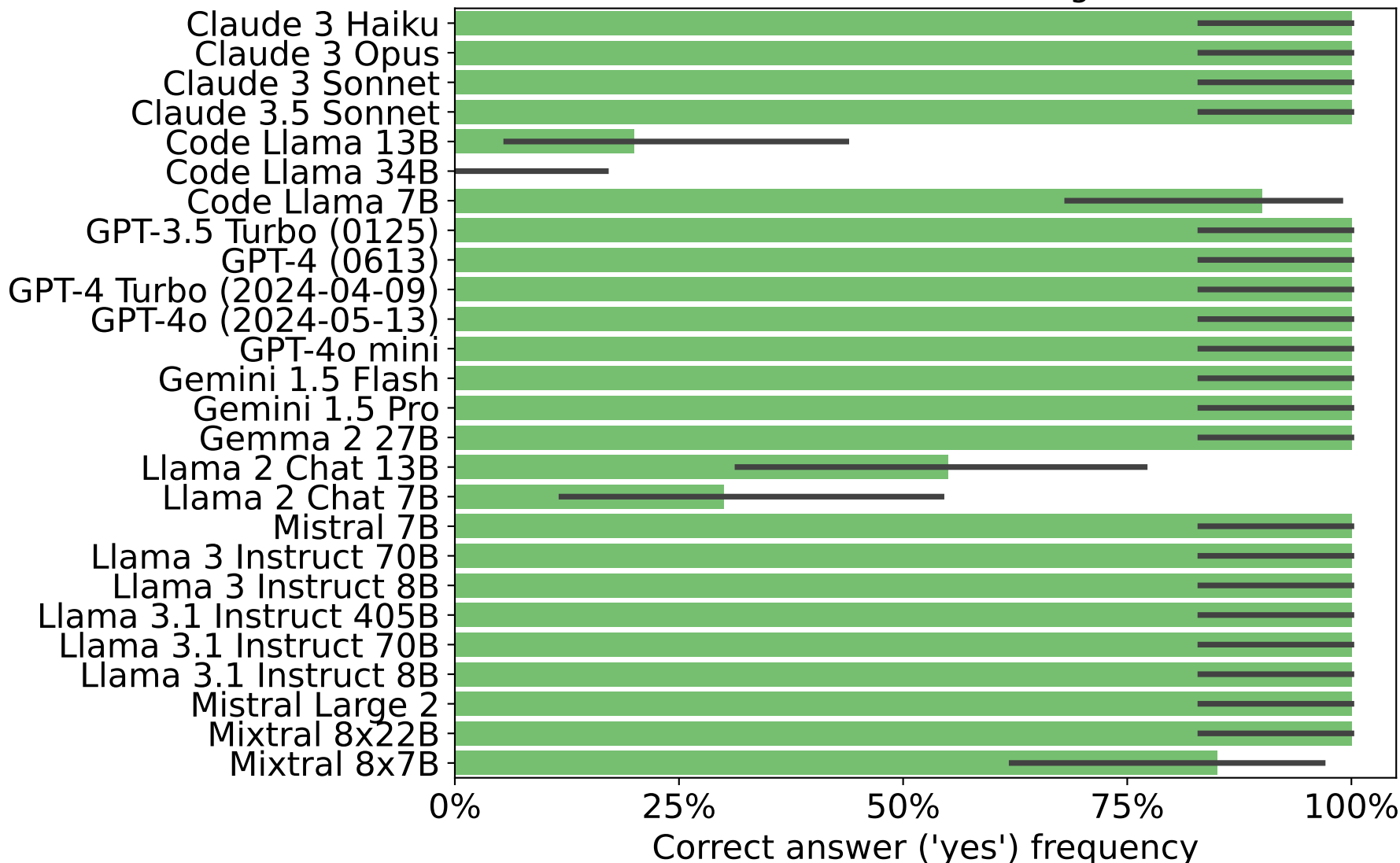uMiN example: Mary might not have been at the wedding. Can we infer that it's not the case that Mary must have been at the wedding?

Model	
Claude 3 Haiku	
Claude 3 Opus	
Claude 3 Sonnet	
Claude 3.5 Sonnet	
Code Llama 13B	
Code Llama 34B	
Code Llama 7B	
GPT-3.5 Turbo (0125)	
GPT-4 (0613)	
GPT-4 Turbo (2024-04-09)	
GPT-4o (2024-05-13)	
GPT-4o mini	
Gemini 1.5 Flash	
Gemini 1.5 Pro	
Gemma 2 27B	
Llama 2 Chat 13B	
Llama 2 Chat 7B	
Mistral 7B	
Llama 3 Instruct 70B	
Llama 3 Instruct 8B	
Llama 3.1 Instruct 405B	
Llama 3.1 Instruct 70B	
Llama 3.1 Instruct 8B	
Mistral Large 2	
Mixtral 8x22B	
Mixtral 8x7B	

Correct answer ('yes') frequency

Temperature: 0, Condition: zero-shot