

# Recent LLMs can do 2-hop and 3-hop no Chain-of-Thought reasoning on natural facts

## Multi-Hop No Chain-of-Thought Reasoning Task

### 2-hop example:

Who won Nobel Chemistry in  $[1900 + [\text{atomic \# of Tc}]]$ ? **1900 + 43**

→ **George de Hevesy**

### 3-hop example:

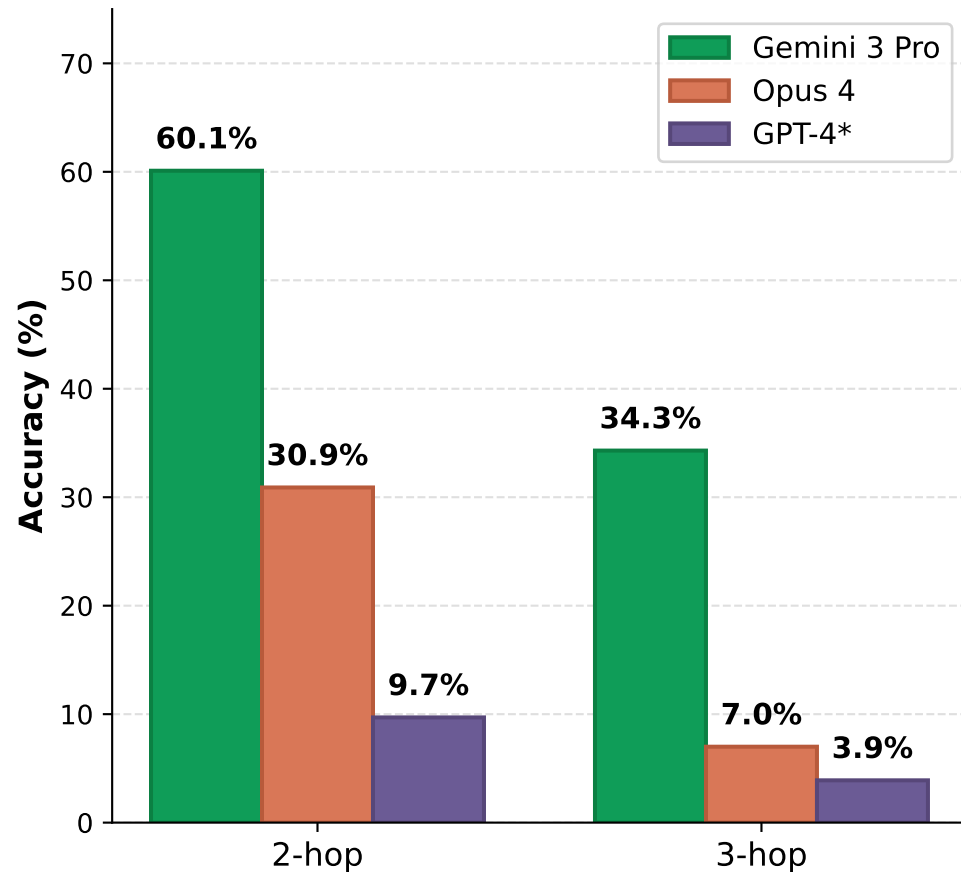
Counties in state that joined the US  $[\text{age Chekhov died}]$ th? **Wyoming** **44**

→ **23**

Intermediate values  
Final answer

*Models must solve in a single forward pass (no chain-of-thought)*

## Performance (with filler tokens, f=300)



*(4-hop accuracy is near chance for all models)*

*\*GPT-4 results don't use filler tokens*