

Summary of performance on some simple inferences

Temperature: 0, Condition: chain_of_thought

