Reasoning over input text GPT3.5 GPT4 Accuracy Gemini Pro 0.8 Mistral Medium Mixtral 8x7B 0.6 1500 2000 3000 250 500 1000 2500 Input length (# tokens) _____ ______ ______ ______ _______ _______ _____ ______ _____ --------_____ ____ _____ ______ _____

CoT

Normal