









-  with CoT
-  with SB
-  GPT-4
-  text-bison
-  Claude-2
-  Claude-1
-  LLaMA-2-70
-  Aggregated scores