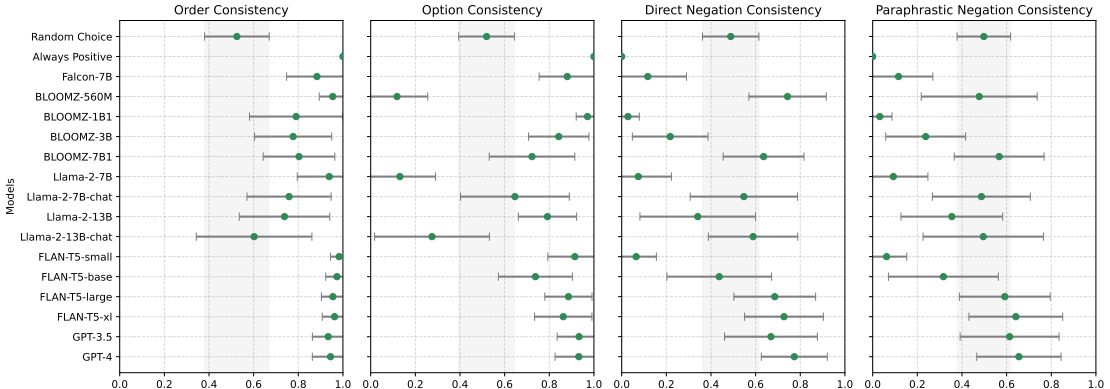

Order Consistency
Option Consistency
Direct Negation Consistency
Paraphrastic Negation Consistency
Models
Random Choice
Always Positive
Falcon-7B
BLOOMZ-560M
BLOOMZ-1B1
BLOOMZ-3B
BLOOMZ-7B1
Llama-2-7B
Llama-2-7B-chat
Llama-2-13B
Llama-2-13B-chat
FLAN-T5-small
FLAN-T5-base
FLAN-T5-large
FLAN-T5-xl
GPT-3.5
GPT-4
0.0
0.2
0.4
0.6
0.8
1.0
Percentage of Consistent Answers