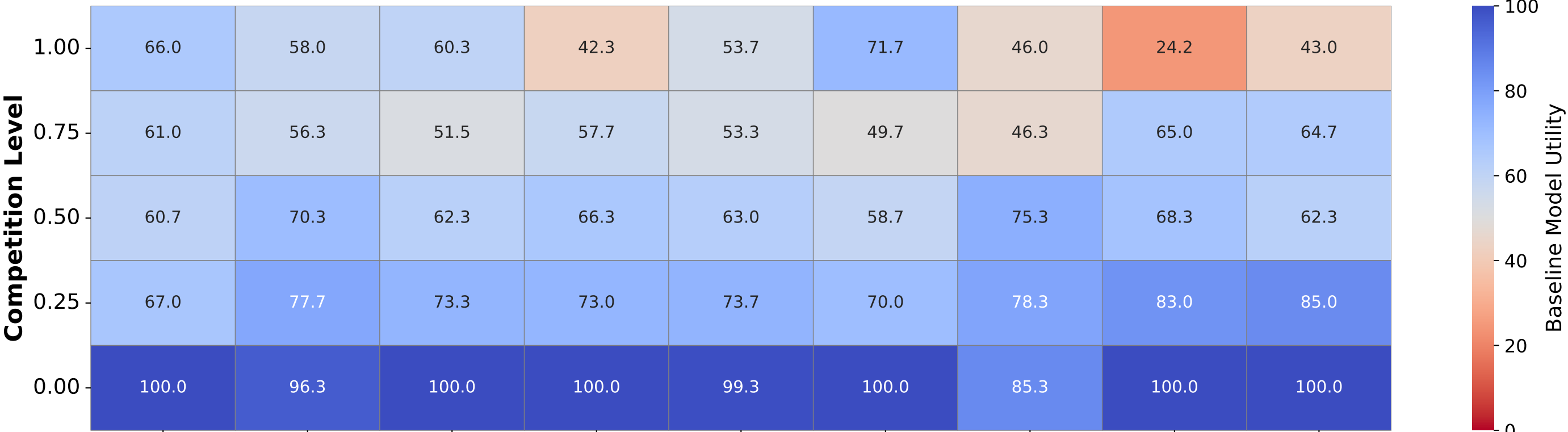
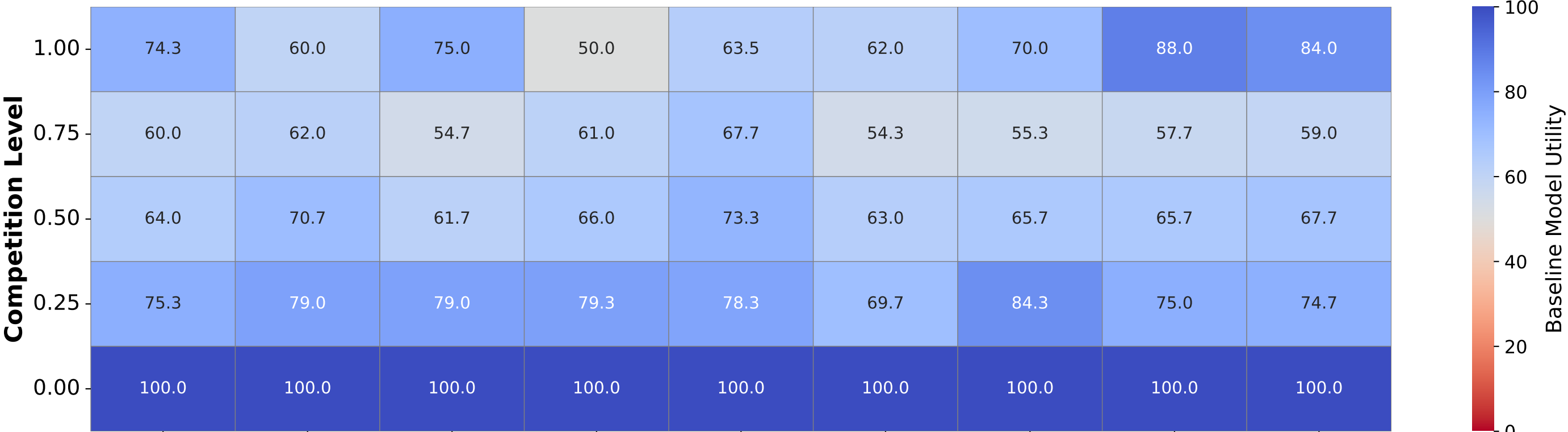


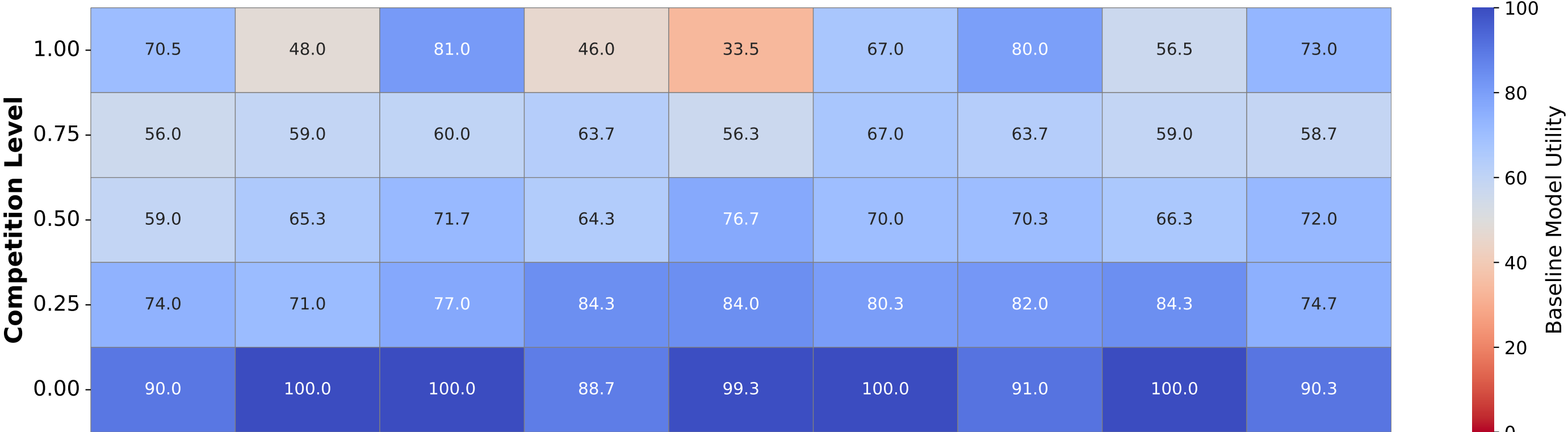
GPT-4o  
(May 2024): Baseline Utility



Gemini 1.5  
Pro: Baseline Utility



Claude 3  
Opus: Baseline Utility



Adversary Models Ordered by MMLU-Pro Performance →

