## Sequential log probabilities gpt2-large distilgpt2 gpt2 Gemma-2-2B Gemma-2-9B -3 -5 -6 gpt2-medium OLMo-2-1124-7B opt-1.3b opt-125m gpt2-xl -3 Average sequential log probabilities opt-2.7b opt-350m opt-6.7b pythia-1.4b pythia-12b -5 -6 pythia-160m pythia-2.8b pythia-410m pythia-6.9b pythia-70m -3 -5 -6 e continuation target content main main main appositive appositive appositive main Continuation Type $\stackrel{\square}{=}$ event association $\stackrel{\square}{=}$ lexical overlap $\stackrel{\square}{=}$ situational inference