GPT2 SV: Attn-6 1.0 0.9 **Ablated Accuracy** 8.0 0.7 0.6 Full Model Acc. 0.5 Abl. Circuit Abl. Random

Condition