

Attention-only transformer
test loss on 3-state ICL-MC

