

Transformer train loss on random tasks with different wildcard alphabet sizes

