Generating open source chess puzzles - notes

Aatu Selkee

January 2, 2026

Contents

1 RL

Compute the log-probability of the models in the same way as with autoregressive models (sum the log probabilities of the chosen tokens). The model must be called K times where K is the amount of tokens (the latter tokens depend on the previous tokens and teacher forcing is not possible I think?). Hence the computational complexity is a lot higher with masked diffusion than with an autoregressive model.