KL Divergence Control

Purpose of KL Divergence

KL divergence measures how much the optimized policy deviates from the original base model, preventing harmful divergence from foundational knowledge.

Visual Representation

π_θ

Current Policy
(Being Optimized)

$KL(\pi_\theta \,||\, \pi_{ref})$

π_{ref}

Reference Policy
(Original Base Model)

⬍ KL measures distribution difference