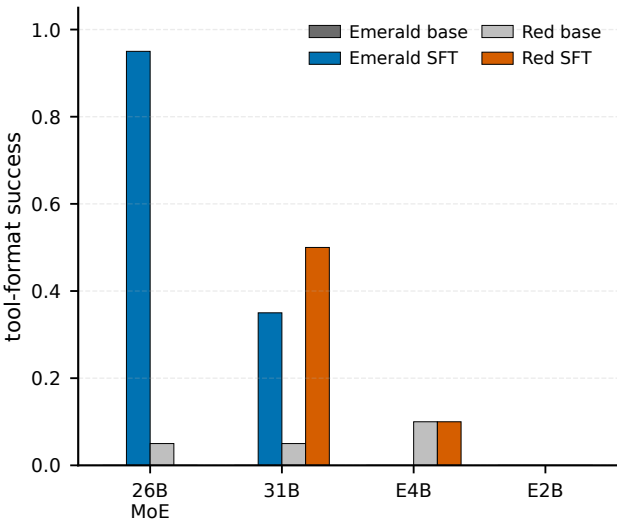


(A) SFT lift across Gemma-4 sizes



(B) Offline GRPO reward curves

