



—×— Baseline —■— GRPO —◆— GRPO-16R
—●— SFT —▲— GRPO-8E —+— GRPO-L