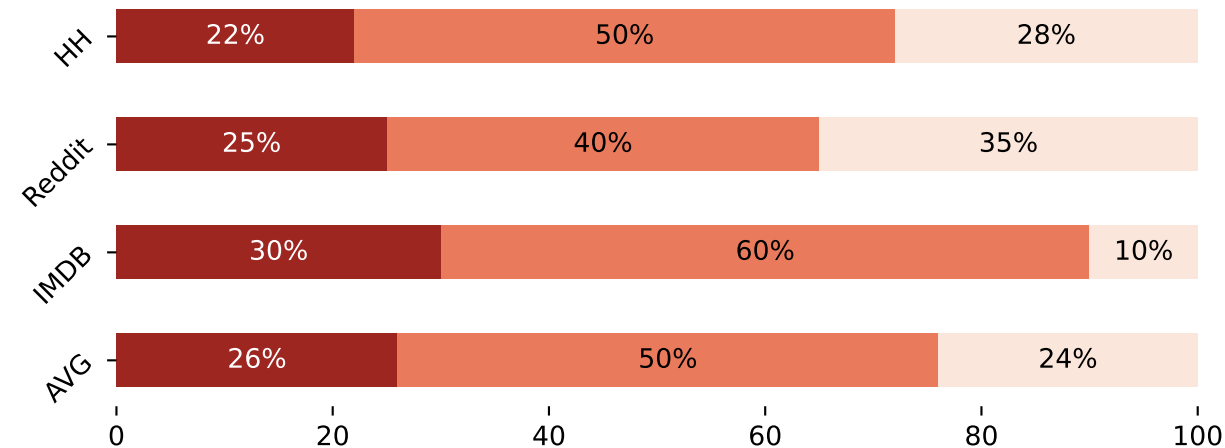
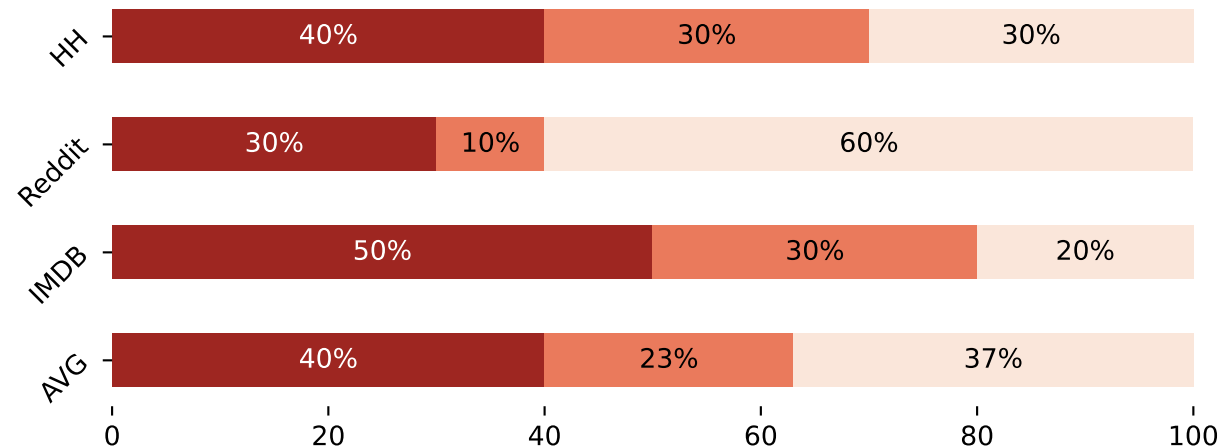


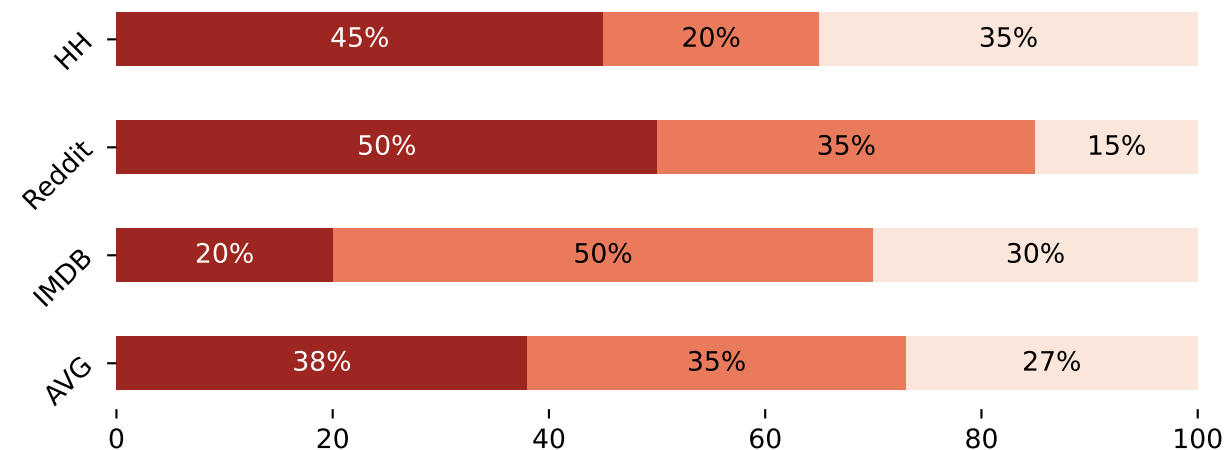
COPR v.s. Golden (Human Eval)



COPR v.s. Golden (GPT-4 Eval)



COPR v.s. DPO-ER (Human Eval)



COPR v.s. DPO-ER (GPT-4 Eval)

