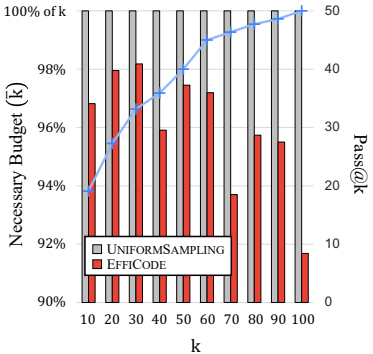
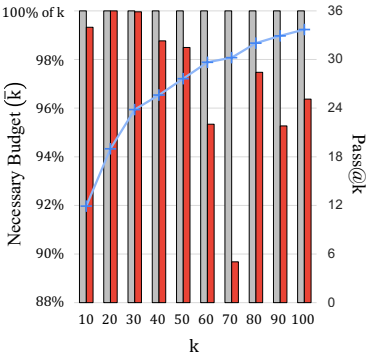
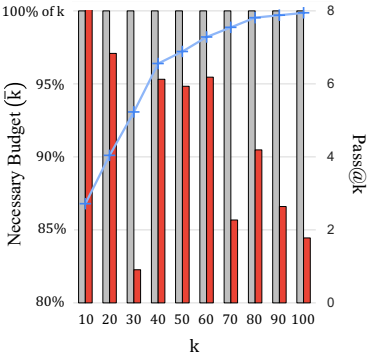


(a) HumanEval-Extreme



(b) MBPP-Extreme



(c) CodeContests-Extreme