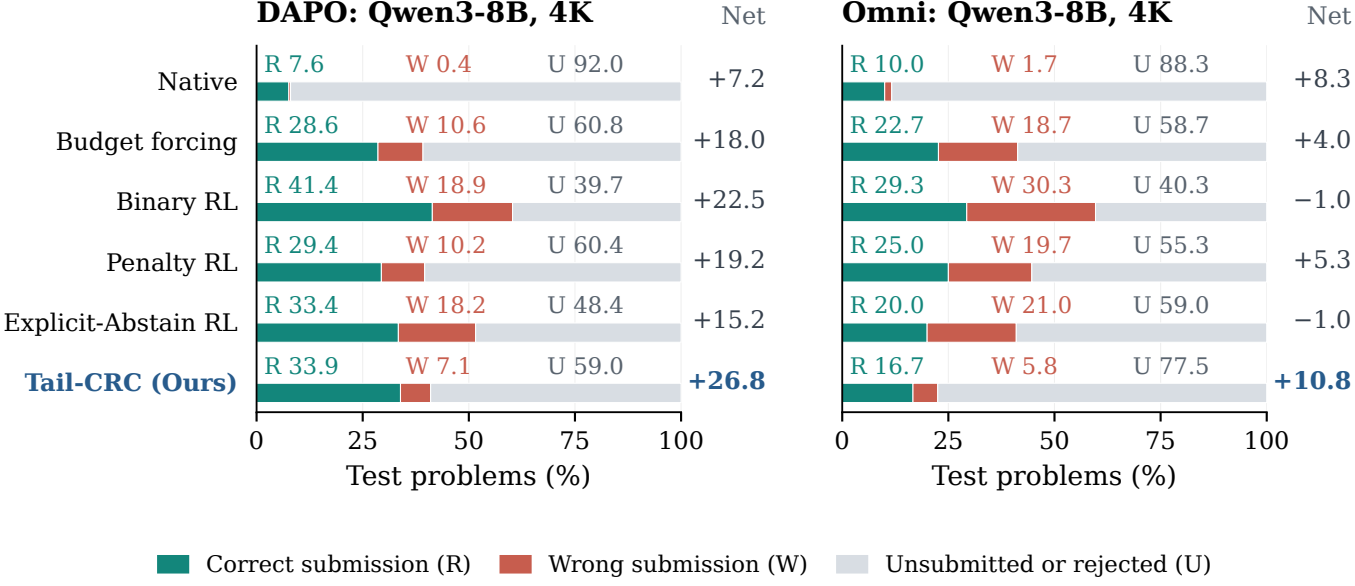

DAPO: Qwen3-8B, 4K
Net
Native
R 7.6
W 0.4
U 92.0
+7.2
Budget forcing
R 28.6
W 10.6
U 60.8
+18.0
Binary RL
R 41.4
W 18.9
U 39.7
+22.5
Penalty RL
R 29.4
W 10.2
U 60.4
+19.2
Explicit-Abstain RL
R 33.4
W 18.2
U 48.4
+15.2
Tail-CRC (Ours)
R 33.9
W 7.1
U 59.0
+26.8
0
25
50
75
100
Test problems (%)
Omni: Qwen3-8B, 4K
Net
R 10.0
W 1.7
U 88.3
+8.3
R 22.7
W 18.7
U 58.7
+4.0
R 29.3
W 30.3
U 40.3
−1.0
R 25.0
W 19.7
U 55.3
+5.3
R 20.0
W 21.0
U 59.0
−1.0
R 16.7
W 5.8
U 77.5
+10.8
0
25
50
75
100
Test problems (%)
Correct submission (R)
Wrong submission (W)
Unsubmitted or rejected (U)