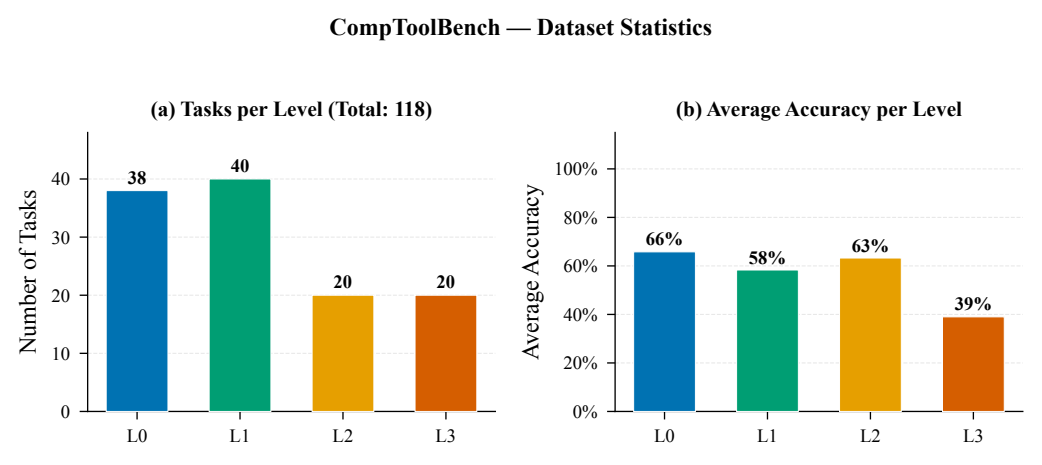
CompToolBench — Dataset Statistics
(a) Tasks per Level (Total: 118)
Number of Tasks
0
10
20
30
40
38
40
20
20
L0
L1
L2
L3
(b) Average Accuracy per Level
Average Accuracy
0%
20%
40%
60%
80%
100%
66%
58%
63%
39%
L0
L1
L2
L3