

qwen3-coder-plus — Flaky summary confusion

Gold

async wait	1	0	0	0	2	0	0	0	2	0	0	0	0	0
concurrency	4	0	0	0	0	0	0	0	1	0	0	0	0	0
floating point operations	0	2	0	1	0	0	0	0	1	0	1	0	0	0
io	0	1	0	4	0	0	0	0	0	0	0	0	0	0
network	1	0	0	0	3	0	0	0	1	0	0	0	0	0
platform dependency	1	0	0	0	0	2	0	0	0	0	0	0	2	0
randomness	1	0	0	1	0	0	2	0	0	0	0	1	0	0
resource leak	2	1	0	0	0	0	0	0	1	0	1	0	0	0
test case timeout	1	0	0	0	0	0	0	0	4	0	0	0	0	0
test order dependency	1	0	0	3	0	0	0	0	0	1	0	0	0	0
time	1	0	0	0	0	0	0	0	0	1	3	0	0	0
too restrictive range	1	0	0	0	0	0	0	0	2	0	0	2	0	0
unordered collection	2	0	1	0	0	0	0	0	0	0	0	0	2	0
unknown	0	0	0	0	0	0	0	0	0	0	0	0	0	0

async wait

concurrency

floating point operations

io

network

platform dependency

randomness

resource leak

test case timeout

test order dependency

time

too restrictive range

unordered collection

unknown

Predicted

4.0

3.5

3.0

2.5

2.0

1.5

1.0

0.5

0.0