problem_id	iteration	dataset_type	status	passed_public	failed_public	timeout_public
0	iteration_0	plus	NO_TESTS	0	0	0
1	iteration_0	plus	PASSED	5	0	0
2	iteration_0	plus	PASSED	3	0	0
3	iteration_0	plus	FAILED	6	0	0
4	iteration_0	plus	FAILED	3	0	0
5	iteration_0	plus	PASSED	3	0	0
6	iteration_0	plus	PASSED	3	0	0
7	iteration_0	plus	PASSED	3	0	0
8	iteration_0	plus	PASSED	3	0	0
9	iteration_0	plus	PASSED	3	0	0
10	iteration_0	plus	PASSED	3	0	0
11	iteration_0	plus	PASSED	3	0	0
12	iteration_0	plus	PASSED	3	0	0
13	iteration_0	plus	PASSED	3	0	0
14	iteration_0	plus	PASSED	3	0	0
15	iteration_0	plus	PASSED	3	0	0
16	iteration_0	plus	PASSED	3	0	0
17	iteration_0	plus	PASSED	4	0	0
18	iteration_0	plus	PASSED	3	0	0
19	iteration_0	plus	PASSED	3	0	0
0	iteration_0	test	PASSED	1	0	0
1	iteration_0	test	PASSED	1	0	0
2	iteration_0	test	PASSED	1	0	0
3	iteration_0	test	PASSED	1	0	0
4	iteration_0	test	FAILED	0	1	0
5	iteration_0	test	FAILED	1	0	0
6	iteration_0	test	PASSED	1	0	0
7	iteration_0	test	PASSED	1	0	0
8	iteration_0	test	PASSED	1	0	0
9	iteration_0	test	PASSED	1	0	0
10	iteration_0	test	PASSED	1	0	0
11	iteration_0	test	FAILED	1	0	0
12	iteration_0	test	PASSED	1	0	0
13	iteration_0	test	PASSED	1	0	0
14	iteration_0	test	PASSED	1	0	0
15	iteration_0	test	FAILED	0	1	0
16	iteration_0	test	PASSED	1	0	0
17	iteration_0	test	PASSED	1	0	0
18	iteration_0	test	PASSED	1	0	0
19	iteration_0	test	PASSED	1	0	0
0	iteration_0	test	PASSED	1	0	0
1	iteration_0	test	PASSED	1	0	0
2	iteration_0	test	FAILED	0	1	0
3	iteration_0	test	PASSED	1	0	0
4	iteration_0	test	FAILED	0	1	0
5	iteration_0	test	FAILED	1	0	0
6	iteration_0	test	PASSED	1	0	0
7	iteration_0	test	PASSED	1	0	0
8	iteration_0	test	PASSED	1	0	0
9	iteration_0	test	PASSED	1	0	0
10	iteration_0	test	PASSED	1	0	0
11	iteration_0	test	FAILED	1	0	0

12	iteration_0	test	PASSED	1	0	0
13	iteration_0	test	PASSED	1	0	0
14	iteration_0	test	PASSED	1	0	0
15	iteration_0	test	FAILED	1	0	0
16	iteration_0	test	PASSED	1	0	0
17	iteration_0	test	PASSED	1	0	0
18	iteration_0	test	PASSED	1	0	0
19	iteration_0	test	PASSED	1	0	0
20	iteration_0	test	FAILED	0	1	0
21	iteration_0	test	PASSED	1	0	0
22	iteration_0	test	PASSED	1	0	0

passed_private	failed_private	timeout_private	passed_generate	failed_generate	timeout_generate
0	0	0	0	0	0
103	0	0	0	0	0
106	0	0	0	0	0
90	18	0	0	0	0
66	36	0	0	0	0
103	0	0	0	0	0
101	0	0	0	0	0
103	0	0	0	0	0
106	0	0	0	0	0
108	0	0	0	0	0
113	0	0	0	0	0
101	0	0	0	0	0
104	0	0	0	0	0
102	0	0	0	0	0
101	0	0	0	0	0
125	0	0	0	0	0
102	0	0	0	0	0
105	0	0	0	0	0
101	0	0	0	0	0
102	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
0	3	0	0	0	0
1	2	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
2	1	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
1	2	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
0	3	0	0	0	0
3	0	0	0	0	0
0	3	0	0	0	0
1	2	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
2	1	0	0	0	0

3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
2	1	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0
0	3	0	0	0	0
3	0	0	0	0	0
3	0	0	0	0	0

total_passed	total_failed	total_timeout	total_tests	dataset_name
0	0	0	0	mbpp
108	0	0	108	mbpp
109	0	0	109	mbpp
96	18	0	114	mbpp
69	36	0	105	mbpp
106	0	0	106	mbpp
104	0	0	104	mbpp
106	0	0	106	mbpp
109	0	0	109	mbpp
111	0	0	111	mbpp
116	0	0	116	mbpp
104	0	0	104	mbpp
107	0	0	107	mbpp
105	0	0	105	mbpp
104	0	0	104	mbpp
128	0	0	128	mbpp
105	0	0	105	mbpp
109	0	0	109	mbpp
104	0	0	104	mbpp
105	0	0	105	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
0	4	0	4	mbpp
2	2	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
3	1	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
1	3	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
0	4	0	4	mbpp
4	0	0	4	mbpp
0	4	0	4	mbpp
2	2	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
4	0	0	4	mbpp
3	1	0	4	mbpp

4	0	0	4 mbpp
4	0	0	4 mbpp
4	0	0	4 mbpp
3	1	0	4 mbpp
4	0	0	4 mbpp
4	0	0	4 mbpp
4	0	0	4 mbpp
4	0	0	4 mbpp
0	4	0	4 mbpp
4	0	0	4 mbpp
4	0	0	4 mbpp

results_directory
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\plus\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-3.5-turbo\resu
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results

C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results
C:\Users\tayla\OneDrive\Dokumente\Paircoder\A-Pair-Coder\Examples\PairCoder-main\logs\mbpp\test\gpt-4\results