📊 Coverage Reports Index

Total Solutions: 40   **LLM Implementations:** 4
Generated: 2025-11-10 12:11:15

gemma_Self_Planning

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/1	4/4	95.8%	87.5%	Good line coverage but some branches untested	View Report →
HumanEval/17	5/5	95.0%	87.5%	Good line coverage but some branches untested	View Report →
HumanEval/26	3/3	100.0%	100.0%	Excellent coverage - well-tested code	View Report →
HumanEval/77	7/8	100.0%	100.0%	1 test(s) failed - fix failing tests first	View Report →
HumanEval/84	0/5	100.0%	N/A	5 test(s) failed - fix failing tests first	View Report →
HumanEval/94	9/9	94.4%	93.8%	Excellent coverage - well-tested code	View Report →
HumanEval/101	6/6	100.0%	100.0%	Excellent coverage - well-tested code	View Report →
HumanEval/107	8/8	100.0%	100.0%	Excellent coverage - well-tested code	View Report →
HumanEval/114	12/12	85.7%	75.0%	Moderate branch coverage - some	View Report →

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
				conditional branches untested	
HumanEval/156	14/14	100.0%	100.0%	Excellent coverage - well-tested code	View Report →

gemma_self_edit

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/1	4/4	96.2%	87.5%	Good line coverage but some branches untested	View Report →
HumanEval/17	3/5	88.2%	75.0%	2 test(s) failed - fix failing tests first	View Report →
HumanEval/26	3/3	100.0%	100.0%	Excellent coverage - well-tested code	View Report →
HumanEval/77	0/8	0.0%	N/A	8 test(s) failed - fix failing tests first	View Report →
HumanEval/84	0/5	100.0%	N/A	5 test(s) failed - fix failing tests first	View Report →
HumanEval/94	9/9	95.3%	92.9%	Excellent coverage - well-tested code	View Report →
HumanEval/101	6/6	100.0%	N/A	Good line coverage achieved	View Report →
HumanEval/107	8/8	100.0%	100.0%	Excellent coverage - well-tested code	View Report →

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/114	12/12	100.0%	100.0%	Excellent coverage - well-tested code	View Report →
HumanEval/156	14/14	100.0%	100.0%	Excellent coverage - well-tested code	View Report →

llama_self_edit

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/1	4/4	96.2%	87.5%	Good line coverage but some branches untested	View Report →
HumanEval/17	3/5	88.2%	75.0%	2 test(s) failed - fix failing tests first	View Report →
HumanEval/26	3/3	100.0%	100.0%	Excellent coverage - well-tested code	View Report →
HumanEval/77	0/8	0.0%	N/A	8 test(s) failed - fix failing tests first	View Report →
HumanEval/84	0/5	100.0%	N/A	5 test(s) failed - fix failing tests first	View Report →
HumanEval/94	9/9	95.3%	92.9%	Excellent coverage - well-tested code	View Report →
HumanEval/101	6/6	100.0%	N/A	Good line coverage achieved	View Report →
HumanEval/107	8/8	100.0%	100.0%	Excellent coverage - well-tested code	View Report →

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/114	12/12	100.0%	100.0%	Excellent coverage - well-tested code	View Report →
HumanEval/156	14/14	100.0%	100.0%	Excellent coverage - well-tested code	View Report →

llama_self_planning

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/1	4/4	96.2%	87.5%	Good line coverage but some branches untested	View Report →
HumanEval/17	3/5	88.2%	75.0%	2 test(s) failed - fix failing tests first	View Report →
HumanEval/26	3/3	100.0%	100.0%	Excellent coverage - well-tested code	View Report →
HumanEval/77	0/8	0.0%	N/A	8 test(s) failed - fix failing tests first	View Report →
HumanEval/84	0/5	100.0%	N/A	5 test(s) failed - fix failing tests first	View Report →
HumanEval/94	9/9	95.3%	92.9%	Excellent coverage - well-tested code	View Report →
HumanEval/101	6/6	100.0%	N/A	Good line coverage achieved	View Report →
HumanEval/107	8/8	100.0%	100.0%	Excellent coverage - well-tested code	View Report →

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/114	12/12	100.0%	100.0%	Excellent coverage - well-tested code	View Report →
HumanEval/156	14/14	100.0%	100.0%	Excellent coverage - well-tested code	View Report →

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report