

APPENDIX

LLMorpheus: Mutation Testing using Large Language Models

SUPPLEMENTAL MATERIALS

Frank Tip, Jonathan Bell, Max Schäfer

This appendix contains the following experimental results:

- Section A expands on the results reported in Sections IV-C and IV-H by including the results for 5 experiments using the *codellama-34b-instruct* LLM at temperature 0.0.
- Sections B, C, and D expand on the results reported in Section IV-E by including the results for 5 experiments using the *codellama-34b-instruct* LLM at temperatures 0.25, 0.50, and 1.0, respectively.
- Section E expands on the discussion in Sections IV-C and IV-H about the variability of the mutants observed in 5 experiments using the *codellama-34b-instruct* LLM, at temperatures 0.25, 0.50, and 1.0, respectively.
- Section F expands on the results reported in Section IV-G by including results for 5 experiments using the *codellama-13b-instruct* LLM at temperature 0.0.
- Section G expands on the results reported in Section IV-G by including results for 5 experiments using the *mixtral-8x7b-instruct* LLM at temperature 0.0.
- Section H expands on the results reported in Section IV-G by including results for 5 experiments using the *gpt-4o-mini* LLM at temperature 0.0.
- Section I expands on the results reported in Section IV-G by including results for 5 experiments using the *llama-3.3-70b-instruct* LLM at temperature 0.0.
- Section J expands on the results reported in Section IV-F by including results for 5 experiments using the `onemutation` template.
- Section K expands on the results reported in Section IV-F by including results for 5 experiments using the `noexplanation` template.
- Section L expands on the results reported in Section IV-F by including results for 5 experiments using the `noinstructions` template.
- Section M expands on the results reported in Section IV-F by including results for 5 experiments using a generic system prompt.
- Section N expands on the results reported in Section IV-F by including results for 5 experiments using the `basic` template.
- Section P includes the results of running the standard mutation operators of *StrykerJS* on the subject applications.
- Section Q reports measurements of the average string edit distance observed when using different LLMs.

All experimental data associated with the experiments reported on in this paper can be found at <https://github.com/neu-se/mutation-testing-data>.

A. Results for *codellama-34b-instruct*, full template, temperature 0.0

Tables X–XIX show the results for 5 experiments with the *codellama-34b-instruct* model at temperature 0.0 using the default prompt and system prompt shown in Figure 7.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,451	194	13	45	1,199	725	473	1	60.55
<i>countries-and-timezones</i>	106	318	89	0	12	217	188	29	0	86.64
<i>crawler-url-parser</i>	176	521	205	14	17	285	157	128	0	55.09
<i>delta</i>	462	1,367	565	10	25	767	634	101	32	86.83
<i>image-downloader</i>	42	124	33	2	0	89	72	17	0	80.90
<i>node-dirty</i>	154	450	153	15	7	275	163	100	12	63.64
<i>node-geo-point</i>	140	408	93	0	13	302	223	79	0	73.84
<i>node-jsonfile</i>	68	199	42	3	0	154	49	48	57	68.83
<i>plural</i>	153	442	101	42	18	281	205	75	1	73.31
<i>pull-stream</i>	351	1,028	238	12	9	769	441	271	57	64.76
<i>q</i>	1,051	3,121	1,000	34	52	2,035	158	1,792	85	11.94
<i>spacel-core</i>	134	395	140	10	6	239	199	39	1	83.68
<i>zip-a-folder</i>	49	143	41	1	1	100	23	3	74	97.00
<i>Total</i>	3,376	9,967	2,894	156	205	6,712	3,237	3,155	320	—

TABLE X

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #312). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,450	194	13	45	1,198	726	471	1	60.68
<i>countries-and-timezones</i>	106	318	89	0	12	217	188	29	0	86.64
<i>crawler-url-parser</i>	176	521	208	14	17	282	156	126	0	55.32
<i>delta</i>	462	1,367	565	10	24	768	636	100	32	86.98
<i>image-downloader</i>	42	124	34	0	0	90	73	17	0	81.11
<i>node-dirty</i>	154	450	154	15	7	274	161	101	12	63.14
<i>node-geo-point</i>	140	410	95	0	13	302	223	79	0	73.84
<i>node-jsonfile</i>	68	199	43	3	0	153	49	47	57	69.28
<i>plural</i>	153	441	101	42	18	280	204	75	1	73.21
<i>pull-stream</i>	351	1,028	236	12	9	771	441	273	57	64.59
<i>q</i>	1,051	3,123	1,004	34	54	2,031	159	1,788	84	11.96
<i>spacel-core</i>	134	394	138	10	7	239	197	41	1	82.85
<i>zip-a-folder</i>	49	143	41	1	1	100	23	3	74	97.00
<i>Total</i>	3,376	9,968	2,902	154	207	6,705	3,236	3,150	319	—

TABLE XI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #314). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,452	195	13	45	1,199	724	474	1	60.47
<i>countries-and-timezones</i>	106	318	89	0	12	217	188	29	0	86.64
<i>crawler-url-parser</i>	176	522	208	14	17	283	157	126	0	55.48
<i>delta</i>	462	1,367	566	10	26	765	633	100	32	86.93
<i>image-downloader</i>	42	124	35	0	0	89	72	17	0	80.90
<i>node-dirty</i>	154	450	154	14	7	275	161	102	12	62.91
<i>node-geo-point</i>	140	408	93	0	13	302	223	79	0	73.84
<i>node-jsonfile</i>	68	199	42	3	0	154	49	48	57	68.83
<i>plural</i>	153	442	101	42	18	281	205	75	1	73.31
<i>pull-stream</i>	351	1,028	237	12	9	770	441	272	57	64.68
<i>q</i>	1,051	3,121	1,000	34	52	2,035	159	1,793	83	11.89
<i>spacel-core</i>	134	393	137	10	7	239	197	41	1	82.85
<i>zip-a-folder</i>	49	143	41	1	1	100	23	3	74	97.00
<i>Total</i>	3,376	9,967	2,898	153	207	6,709	3,232	3,159	318	—

TABLE XII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #315). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,452	195	13	45	1,199	725	473	1	60.55
<i>countries-and-timezones</i>	106	318	89	0	12	217	188	29	0	86.64
<i>crawler-url-parser</i>	176	521	207	14	17	283	157	125	1	55.83
<i>delta</i>	462	1,367	565	10	26	766	634	100	32	86.95
<i>image-downloader</i>	42	124	35	0	0	89	72	17	0	80.90
<i>node-dirty</i>	154	450	153	15	8	274	160	102	12	62.77
<i>node-geo-point</i>	140	410	95	0	13	302	222	80	0	73.51
<i>node-jsonfile</i>	68	200	43	3	0	154	49	48	57	68.83
<i>plural</i>	153	441	101	43	18	279	203	75	1	73.12
<i>pull-stream</i>	351	1,028	237	13	9	769	444	268	57	65.15
<i>q</i>	1,051	3,121	1,001	34	54	2,032	158	1,790	84	11.91
<i>spacl-core</i>	134	394	139	10	7	238	197	40	1	83.19
<i>zip-a-folder</i>	49	143	41	1	1	100	23	3	74	97.00
<i>Total</i>	3,376	9,969	2,901	156	210	6,702	3,232	3,150	320	—

TABLE XIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #316). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,450	194	13	44	1,199	726	472	1	60.63
<i>countries-and-timezones</i>	106	317	89	0	12	216	187	29	0	86.57
<i>crawler-url-parser</i>	176	521	205	14	17	285	157	128	0	55.09
<i>delta</i>	462	1,367	565	10	25	767	636	100	31	86.96
<i>image-downloader</i>	42	124	32	2	0	90	73	17	0	81.11
<i>node-dirty</i>	154	450	153	15	7	275	161	102	12	62.91
<i>node-geo-point</i>	140	409	94	0	13	302	223	79	0	73.84
<i>node-jsonfile</i>	68	199	42	3	0	154	49	48	57	68.83
<i>plural</i>	153	442	101	43	18	280	204	75	1	73.21
<i>pull-stream</i>	351	1,028	236	12	9	771	441	273	57	64.59
<i>q</i>	1,051	3,122	1,003	34	54	2,031	159	1,788	84	11.96
<i>spacl-core</i>	134	394	138	10	7	239	198	40	1	83.26
<i>zip-a-folder</i>	49	143	41	1	1	100	23	3	74	97.00
<i>Total</i>	3,376	9,966	2,893	157	207	6,709	3,237	3,154	318	—

TABLE XIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #317). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	total
<i>Complex.js</i>	3,050.00	637.85	967,508	102,517	1,070,025
<i>countries-and-timezones</i>	1,070.89	313.86	105,828	23,441	129,269
<i>crawler-url-parser</i>	1,642.70	929.43	386,223	39,175	425,398
<i>delta</i>	2,961.66	3,839.60	890,252	98,974	989,226
<i>image-downloader</i>	430.53	379.25	24,655	9,134	33,789
<i>node-dirty</i>	1,526.20	241.81	246,248	33,070	279,318
<i>node-geo-point</i>	1,411.11	987.17	316,333	30,013	346,346
<i>node-jsonfile</i>	690.61	474.78	57,516	14,797	72,313
<i>plural</i>	1,521.32	155.24	265,602	34,174	299,776
<i>pull-stream</i>	2,492.50	1,608.97	208,130	76,513	284,643
<i>q</i>	5,241.46	14,034.67	2,127,655	220,215	2,347,870
<i>spacl-core</i>	1,351.08	798.96	162,705	29,236	191,941
<i>zip-a-folder</i>	500.57	1,156.11	82,457	10,725	93,182
<i>Total</i>	23,890.64	25,557.70	5,841,112	721,984	6,563,096

TABLE XV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #312). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,087.58	637.10	967,508	102,428	1,069,936
<i>countries-and-timezones</i>	1,070.89	313.12	105,828	23,427	129,255
<i>crawler-url-parser</i>	1,645.11	679.89	386,223	39,210	425,433
<i>delta</i>	2,941.55	3,838.23	890,252	98,951	989,203
<i>image-downloader</i>	430.54	377.12	24,655	9,175	33,830
<i>node-dirty</i>	1,526.11	248.57	246,248	33,038	279,286
<i>node-geo-point</i>	1,411.06	999.59	316,333	29,959	346,292
<i>node-jsonfile</i>	690.69	478.32	57,516	14,829	72,345
<i>plural</i>	1,521.04	147.85	265,602	34,164	299,766
<i>pull-stream</i>	2,477.99	1,400.76	208,130	76,398	284,528
<i>q</i>	5,231.88	14,004.86	2,127,655	220,252	2,347,907
<i>spacl-core</i>	1,351.05	805.30	162,705	29,283	191,988
<i>zip-a-folder</i>	500.57	1,152.62	82,457	10,705	93,162
<i>Total</i>	23,886.05	25,083.32	5,841,112	721,819	6,562,931

TABLE XVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #314). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,000.46	630.00	967,508	102,314	1,069,822
<i>countries-and-timezones</i>	1,070.90	314.35	105,828	23,438	129,266
<i>crawler-url-parser</i>	1,644.58	1,051.08	386,223	39,105	425,328
<i>delta</i>	3,006.51	3,795.29	890,252	98,978	989,230
<i>image-downloader</i>	430.54	376.08	24,655	9,186	33,841
<i>node-dirty</i>	1,526.69	247.49	246,248	33,089	279,337
<i>node-geo-point</i>	1,411.05	1,003.93	316,333	30,010	346,343
<i>node-jsonfile</i>	690.67	478.93	57,516	14,803	72,319
<i>plural</i>	1,521.23	148.33	265,602	34,082	299,684
<i>pull-stream</i>	2,492.02	1,392.68	208,130	76,599	284,729
<i>q</i>	5,296.20	14,072.49	2,127,655	220,395	2,348,050
<i>spacl-core</i>	1,351.09	802.30	162,705	29,334	192,039
<i>zip-a-folder</i>	500.56	1,155.04	82,457	10,764	93,221
<i>Total</i>	23,942.51	25,467.99	5,841,112	722,097	6,563,209

TABLE XVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #315). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,034.16	631.23	967,508	102,497	1,070,005
<i>countries-and-timezones</i>	1,070.91	309.40	105,828	23,444	129,272
<i>crawler-url-parser</i>	1,644.09	1,022.89	386,223	39,174	425,397
<i>delta</i>	2,983.03	3,969.11	890,252	99,003	989,255
<i>image-downloader</i>	430.51	374.45	24,655	9,148	33,803
<i>node-dirty</i>	1,563.30	251.17	246,248	33,068	279,316
<i>node-geo-point</i>	1,411.00	1,001.75	316,333	30,041	346,374
<i>node-jsonfile</i>	690.69	474.83	57,516	14,750	72,266
<i>plural</i>	1,521.09	151.19	265,602	34,132	299,734
<i>pull-stream</i>	2,541.67	1,398.89	208,130	76,567	284,697
<i>q</i>	5,399.09	13,959.40	2,127,655	220,191	2,347,846
<i>spacl-core</i>	1,351.08	959.37	162,705	29,287	191,992
<i>zip-a-folder</i>	510.58	1,154.14	82,457	10,725	93,182
<i>Total</i>	24,151.21	25,657.83	5,841,112	722,027	6,563,139

TABLE XVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #316). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	<i>LLMorpheus</i>	<i>StrykerJS</i>			
<i>Complex.js</i>	3,019.69	669.52	967,508	102,524	1,070,032
<i>countries-and-timezones</i>	1,070.82	313.99	105,828	23,425	129,253
<i>crawler-url-parser</i>	1,641.20	958.26	386,223	39,160	425,383
<i>delta</i>	3,013.74	3,867.52	890,252	99,031	989,283
<i>image-downloader</i>	430.56	378.47	24,655	9,117	33,772
<i>node-dirty</i>	1,527.49	251.95	246,248	33,113	279,361
<i>node-geo-point</i>	1,451.19	1,042.75	316,333	29,894	346,227
<i>node-jsonfile</i>	690.66	479.95	57,516	14,803	72,319
<i>plural</i>	1,521.18	150.87	265,602	34,163	299,765
<i>pull-stream</i>	2,486.78	1,384.03	208,130	76,520	284,650
<i>q</i>	5,250.40	14,006.76	2,127,655	220,193	2,347,848
<i>spacel-core</i>	1,350.99	808.62	162,705	29,297	192,002
<i>zip-a-folder</i>	500.63	1,171.49	82,457	10,705	93,162
<i>Total</i>	23,955.33	25,484.18	5,841,112	721,945	6,563,057

TABLE XIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #317). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

B. Results for *codellama-34b-instruct*, full template, temperature 0.25

Tables XX–XXIX show the results for 5 experiments with the *codellama-34b-instruct* model at temperature 0.25 using the default prompt and system prompt shown in Figure 7.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,460	210	9	44	1,197	730	466	1	61.07
<i>countries-and-timezones</i>	106	315	86	0	10	219	181	38	0	82.65
<i>crawler-url-parser</i>	176	515	216	7	12	260	167	93	0	64.23
<i>delta</i>	462	1,366	550	10	25	781	642	111	28	85.79
<i>image-downloader</i>	42	125	38	1	0	86	71	15	0	82.56
<i>node-dirty</i>	154	457	157	9	12	279	175	93	11	66.67
<i>node-geo-point</i>	140	381	76	1	8	293	225	68	0	76.79
<i>node-jsonfile</i>	68	202	45	4	2	151	52	41	58	72.85
<i>plural</i>	153	440	104	44	19	273	208	63	2	76.92
<i>pull-stream</i>	351	1,030	236	7	8	779	452	270	57	65.34
<i>q</i>	1,051	3,126	1,011	23	42	2,050	153	1,813	84	11.56
<i>spacel-core</i>	134	396	140	11	4	223	187	36	0	83.86
<i>zip-a-folder</i>	49	141	41	1	1	97	24	4	69	95.88
<i>Total</i>	3,376	9,954	2,910	127	187	6,688	3,267	3,111	310	—

TABLE XX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #348). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,460	202	13	45	1,200	729	471	0	60.75
<i>countries-and-timezones</i>	106	316	87	1	5	223	198	25	0	88.79
<i>crawler-url-parser</i>	176	518	206	13	15	269	163	106	0	60.59
<i>delta</i>	462	1,362	576	5	29	752	608	110	34	85.37
<i>image-downloader</i>	42	123	35	3	0	85	67	18	0	78.82
<i>node-dirty</i>	154	457	165	17	4	271	163	96	12	64.58
<i>node-geo-point</i>	140	412	85	0	13	311	243	68	0	78.14
<i>node-jsonfile</i>	68	204	42	4	0	158	50	54	54	65.82
<i>plural</i>	153	438	101	35	19	283	213	68	2	75.97
<i>pull-stream</i>	351	1,027	241	9	5	772	450	265	57	65.67
<i>q</i>	1,051	3,125	1,037	45	63	1,980	144	1,756	80	11.31
<i>spacel-core</i>	134	393	134	10	6	225	194	30	1	86.67
<i>zip-a-folder</i>	49	140	41	0	1	95	36	53	6	44.21
<i>Total</i>	3,376	9,975	2,952	155	205	6,624	3,258	3,120	246	—

TABLE XXI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #350). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,451	199	12	45	1,195	737	457	1	61.76
<i>countries-and-timezones</i>	106	315	88	0	12	215	189	26	0	87.91
<i>crawler-url-parser</i>	176	512	211	16	19	247	144	103	0	58.30
<i>delta</i>	462	1,362	554	8	24	776	647	100	29	87.11
<i>image-downloader</i>	42	123	36	0	0	87	71	16	0	81.61
<i>node-dirty</i>	154	454	156	14	10	274	162	100	12	63.50
<i>node-geo-point</i>	140	413	103	0	13	294	218	76	0	74.15
<i>node-jsonfile</i>	68	201	43	4	0	154	55	43	56	72.08
<i>plural</i>	153	439	98	36	19	286	204	81	1	71.68
<i>pull-stream</i>	351	1,029	239	12	5	773	442	280	51	63.78
<i>q</i>	1,051	3,118	1,038	32	46	2,002	146	1,780	76	11.09
<i>spacel-core</i>	134	396	142	9	3	228	193	35	0	84.65
<i>zip-a-folder</i>	49	145	43	2	2	98	24	3	71	96.94
<i>Total</i>	3,376	9,958	2,950	145	198	6,629	3,232	3,100	297	—

TABLE XXII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #351). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,457	222	7	36	1,192	721	471	0	60.49
<i>countries-and-timezones</i>	106	317	86	2	2	227	200	27	0	88.11
<i>crawler-url-parser</i>	176	513	213	15	12	260	151	109	0	58.08
<i>delta</i>	462	1,370	576	7	22	765	619	109	37	85.75
<i>image-downloader</i>	42	122	37	0	3	82	62	20	0	75.61
<i>node-dirty</i>	154	454	160	12	15	267	158	97	12	63.67
<i>node-geo-point</i>	140	414	92	2	9	308	226	82	0	73.38
<i>node-jsonfile</i>	68	201	48	1	1	151	48	47	56	68.87
<i>plural</i>	153	440	101	36	24	279	212	66	1	76.34
<i>pull-stream</i>	351	1,033	243	8	9	773	443	277	53	64.17
<i>q</i>	1,051	3,121	1,009	17	51	2,044	141	1,822	81	10.86
<i>spac-core</i>	134	397	145	14	8	214	183	30	1	85.98
<i>zip-a-folder</i>	49	142	35	1	1	105	27	2	76	98.10
<i>Total</i>	3,376	9,981	2,967	122	193	6,667	3,191	3,159	317	—

TABLE XXIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #352). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,451	226	11	46	1,168	717	451	0	61.39
<i>countries-and-timezones</i>	106	313	80	0	8	225	199	26	0	88.44
<i>crawler-url-parser</i>	176	523	204	15	19	268	161	107	0	60.07
<i>delta</i>	462	1,368	574	11	23	760	620	106	34	86.05
<i>image-downloader</i>	42	126	34	3	1	85	71	14	0	83.53
<i>node-dirty</i>	154	450	161	17	7	265	167	86	12	67.55
<i>node-geo-point</i>	140	412	93	1	19	296	220	76	0	74.32
<i>node-jsonfile</i>	68	203	41	3	1	158	55	49	54	68.99
<i>plural</i>	153	441	100	30	20	291	215	75	1	74.23
<i>pull-stream</i>	351	1,029	237	9	8	775	448	276	51	64.39
<i>q</i>	1,051	3,132	1,018	20	53	2,041	151	1,803	87	11.66
<i>spac-core</i>	134	395	132	10	9	226	189	36	1	84.07
<i>zip-a-folder</i>	49	140	35	0	1	104	48	50	6	51.92
<i>Total</i>	3,376	9,983	2,935	130	215	6,662	3,261	3,155	246	—

TABLE XXIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #353). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	3,044.19	631.79	967,508	101,588	1,069,096
<i>countries-and-timezones</i>	1,070.90	327.25	105,828	23,471	129,299
<i>crawler-url-parser</i>	1,646.30	818.05	386,223	39,000	425,223
<i>delta</i>	2,954.37	3,844.00	890,252	99,341	989,593
<i>image-downloader</i>	430.54	348.86	24,655	9,217	33,872
<i>node-dirty</i>	1,532.06	232.35	246,248	32,400	278,648
<i>node-geo-point</i>	1,708.63	961.35	291,061	26,301	317,362
<i>node-jsonfile</i>	740.73	502.52	57,516	14,400	71,916
<i>plural</i>	1,537.89	149.38	265,602	33,182	298,784
<i>pull-stream</i>	2,522.38	1,395.14	208,130	76,091	284,221
<i>q</i>	5,294.80	14,085.10	2,127,655	218,620	2,346,275
<i>spac-core</i>	1,351.05	728.32	162,705	29,167	191,872
<i>zip-a-folder</i>	500.56	1,086.06	82,457	10,557	93,014
<i>Total</i>	24,334.42	25,110.17	5,815,840	713,335	6,529,175

TABLE XXV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #348). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,029.62	629.66	967,508	100,540	1,068,048
countries-and-timezones	1,070.85	314.98	105,828	23,186	129,014
crawler-url-parser	1,777.23	867.78	386,223	38,916	425,139
delta	2,978.88	3,758.68	890,252	99,176	989,428
image-downloader	430.55	366.05	24,655	9,223	33,878
node-dirty	1,528.73	237.16	246,248	32,776	279,024
node-geo-point	1,411.08	1,021.61	316,333	29,301	345,634
node-jsonfile	690.69	485.24	57,516	14,071	71,587
plural	1,521.09	154.20	265,602	33,560	299,162
pull-stream	2,503.60	1,383.12	208,130	76,551	284,681
q	5,379.31	13,584.78	2,127,655	217,699	2,345,354
spacl-core	1,350.98	739.03	162,705	29,184	191,889
zip-a-folder	500.57	531.10	82,457	10,753	93,210
Total	24,173.19	24,073.41	5,841,112	714,936	6,556,048

TABLE XXVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #350). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,026.34	650.01	967,508	101,118	1,068,626
countries-and-timezones	1,070.91	309.33	105,828	23,331	129,159
crawler-url-parser	1,644.86	811.88	386,223	39,215	425,438
delta	2,962.02	3,896.15	890,252	99,274	989,526
image-downloader	430.54	373.38	24,655	9,163	33,818
node-dirty	1,526.45	247.44	246,248	32,894	279,142
node-geo-point	1,421.09	935.39	316,333	29,830	346,163
node-jsonfile	690.68	504.57	57,516	14,702	72,218
plural	1,521.32	151.42	265,602	33,298	298,900
pull-stream	2,497.56	1,351.06	208,130	76,100	284,230
q	5,341.32	13,704.14	2,127,655	218,805	2,346,460
spacl-core	1,351.03	756.66	162,705	28,939	191,644
zip-a-folder	500.57	1,119.70	82,457	10,786	93,243
Total	23,984.67	24,811.13	5,841,112	717,455	6,558,567

TABLE XXVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #351). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,031.72	631.91	967,508	102,075	1,069,583
countries-and-timezones	1,070.83	324.27	105,828	23,502	129,330
crawler-url-parser	1,642.15	835.28	386,223	39,240	425,463
delta	2,969.01	3,912.89	890,252	99,383	989,635
image-downloader	430.57	486.28	24,655	9,228	33,883
node-dirty	1,527.16	237.65	246,248	32,850	279,098
node-geo-point	1,411.10	1,015.42	316,333	28,895	345,228
node-jsonfile	690.64	496.41	57,516	14,557	72,073
plural	1,521.03	148.10	265,602	33,162	298,764
pull-stream	2,509.05	1,370.63	208,130	75,917	284,047
q	5,300.14	14,131.12	2,127,655	218,921	2,346,576
spacl-core	1,351.01	712.33	162,705	28,809	191,514
zip-a-folder	500.63	1,266.61	82,457	10,707	93,164
Total	23,955.04	25,568.93	5,841,112	717,246	6,558,358

TABLE XXVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #352). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
<i>Complex.js</i>	3,082.97	613.65	967,508	101,316	1,068,824
<i>countries-and-timezones</i>	1,070.86	326.66	105,828	22,979	128,807
<i>crawler-url-parser</i>	1,641.17	853.01	386,223	38,790	425,013
<i>delta</i>	2,952.50	3,773.99	890,252	99,524	989,776
<i>image-downloader</i>	470.58	359.76	24,655	8,898	33,553
<i>node-dirty</i>	1,526.98	239.71	246,248	32,476	278,724
<i>node-geo-point</i>	1,411.01	1,001.71	316,333	29,427	345,760
<i>node-jsonfile</i>	690.74	500.76	57,516	14,495	72,011
<i>plural</i>	1,521.11	158.00	265,602	33,838	299,440
<i>pull-stream</i>	2,510.21	1,355.46	208,130	75,432	283,562
<i>q</i>	5,390.06	14,133.05	2,127,655	217,855	2,345,510
<i>spacel-core</i>	1,350.98	752.72	162,705	29,399	192,104
<i>zip-a-folder</i>	520.57	564.11	82,457	10,749	93,206
<i>Total</i>	24,139.74	24,632.57	5,841,112	715,178	6,556,290

TABLE XXIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #353). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.25, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

C. Results for *codellama-34b-instruct*, full template, temperature 0.5

Tables XXX–XXXIX show the results for 5 experiments with the *codellama-34b-instruct* model at temperature 0.5 using the default prompt and system prompt shown in Figure 7.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,448	211	7	39	1,191	739	452	0	62.05
<i>countries-and-timezones</i>	106	311	81	3	3	224	194	30	0	86.61
<i>crawler-url-parser</i>	176	510	185	14	13	298	166	108	24	63.76
<i>delta</i>	462	1,369	560	13	27	769	642	93	34	87.91
<i>image-downloader</i>	42	126	35	1	1	89	68	21	0	76.40
<i>node-dirty</i>	154	454	162	8	7	277	158	107	12	61.37
<i>node-geo-point</i>	140	409	87	2	18	302	230	72	0	76.16
<i>node-jsonfile</i>	68	201	41	3	4	153	51	43	59	71.90
<i>plural</i>	153	438	105	23	21	289	219	69	1	76.12
<i>pull-stream</i>	351	1,037	225	10	6	796	465	278	53	65.08
<i>q</i>	1,051	3,114	977	25	39	2,073	163	1,823	87	12.06
<i>spacel-core</i>	134	393	132	3	8	250	210	39	1	84.40
<i>zip-a-folder</i>	49	122	33	1	1	87	48	33	6	62.07
<i>Total</i>	3,376	9,932	2,834	113	187	6,798	3,353	3,168	277	—

TABLE XXX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #318). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,456	250	7	33	1,166	714	452	0	61.23
<i>countries-and-timezones</i>	106	316	81	4	10	221	192	29	0	86.88
<i>crawler-url-parser</i>	176	515	176	11	18	310	173	137	0	55.81
<i>delta</i>	462	1,366	548	8	18	792	668	89	35	88.76
<i>image-downloader</i>	42	126	36	3	5	82	54	28	0	65.85
<i>node-dirty</i>	154	456	147	8	13	288	173	104	11	63.89
<i>node-geo-point</i>	140	414	91	2	12	309	230	79	0	74.43
<i>node-jsonfile</i>	68	200	44	3	1	152	53	38	61	75.00
<i>plural</i>	153	440	102	30	23	285	215	70	0	75.44
<i>pull-stream</i>	351	1,028	218	9	17	784	463	270	51	65.56
<i>q</i>	1,051	3,125	1,017	36	52	2,020	166	1,774	80	12.18
<i>spacel-core</i>	134	392	120	10	5	257	221	35	1	86.38
<i>zip-a-folder</i>	49	143	36	3	1	103	29	5	69	95.15
<i>Total</i>	3,376	9,977	2,866	134	208	6,769	3,351	3,110	308	—

TABLE XXXI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #319). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,459	232	10	35	1,182	730	452	0	61.76
<i>countries-and-timezones</i>	106	315	86	4	3	222	199	23	0	89.64
<i>crawler-url-parser</i>	176	516	201	15	16	284	166	117	1	58.80
<i>delta</i>	462	1,357	567	9	21	760	621	104	35	86.32
<i>image-downloader</i>	42	125	33	3	4	85	60	25	0	70.59
<i>node-dirty</i>	154	457	156	9	13	279	166	101	12	63.80
<i>node-geo-point</i>	140	371	88	3	11	269	211	58	0	78.44
<i>node-jsonfile</i>	68	200	46	4	0	150	54	39	57	74.00
<i>plural</i>	153	417	88	22	14	293	208	84	1	71.33
<i>pull-stream</i>	351	1,034	205	5	11	813	476	278	59	65.81
<i>q</i>	1,051	3,127	987	30	53	2,057	143	1,838	76	10.65
<i>spacel-core</i>	134	395	124	7	4	260	217	42	1	83.85
<i>zip-a-folder</i>	49	143	27	0	2	114	31	4	79	96.49
<i>Total</i>	3,376	9,916	2,840	121	187	6,768	3,282	3,165	321	—

TABLE XXXII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #320). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,456	216	9	41	1,190	716	474	0	60.17
<i>countries-and-timezones</i>	106	306	70	6	9	221	191	29	1	86.88
<i>crawler-url-parser</i>	176	512	202	18	13	279	158	121	0	56.63
<i>delta</i>	462	1,358	554	12	17	775	634	105	36	86.45
<i>image-downloader</i>	42	125	36	2	2	85	68	17	0	80.00
<i>node-dirty</i>	154	454	166	12	7	269	152	105	12	60.97
<i>node-geo-point</i>	140	409	81	3	13	312	229	83	0	73.40
<i>node-jsonfile</i>	68	201	40	1	0	160	58	36	66	77.50
<i>plural</i>	153	435	96	30	13	296	232	63	1	78.72
<i>pull-stream</i>	351	1,034	232	5	8	789	466	268	55	66.03
<i>q</i>	1,051	3,126	1,015	23	55	2,033	152	1,784	97	12.25
<i>spacl-core</i>	134	391	118	5	6	262	223	38	1	85.50
<i>zip-a-folder</i>	49	144	41	1	2	100	45	50	5	50.00
<i>Total</i>	3,376	9,951	2,867	127	186	6,771	3,324	3,173	274	—

TABLE XXXIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #321). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,451	219	7	39	1,186	725	461	0	61.13
<i>countries-and-timezones</i>	106	311	78	1	8	224	197	27	0	87.95
<i>crawler-url-parser</i>	176	518	200	8	13	297	173	124	0	58.25
<i>delta</i>	462	1,362	535	8	28	791	660	92	39	88.37
<i>image-downloader</i>	42	125	40	1	4	80	68	12	0	85.00
<i>node-dirty</i>	154	455	159	8	5	283	161	110	12	61.13
<i>node-geo-point</i>	140	413	98	0	10	305	239	66	0	78.36
<i>node-jsonfile</i>	68	204	41	2	1	160	51	42	67	73.75
<i>plural</i>	153	439	97	30	23	289	225	64	0	77.85
<i>pull-stream</i>	351	1,032	216	3	7	806	490	255	61	68.36
<i>q</i>	1,051	3,120	1,002	31	59	2,028	141	1,801	86	11.19
<i>spacl-core</i>	134	397	133	8	4	252	212	39	1	84.52
<i>zip-a-folder</i>	49	144	39	1	1	103	27	5	71	95.15
<i>Total</i>	3,376	9,971	2,857	108	202	6,804	3,369	3,098	337	—

TABLE XXXIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #322). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	3,016.85	592.07	967,508	100,627	1,068,135
<i>countries-and-timezones</i>	1,070.93	319.44	105,828	22,932	128,760
<i>crawler-url-parser</i>	1,668.66	1,244.64	386,223	38,920	425,143
<i>delta</i>	2,980.99	3,869.42	890,252	98,347	988,599
<i>image-downloader</i>	430.55	365.90	24,655	8,925	33,580
<i>node-dirty</i>	1,532.57	247.29	246,248	32,995	279,243
<i>node-geo-point</i>	1,411.06	1,006.90	316,333	29,454	345,787
<i>node-jsonfile</i>	720.71	494.61	57,516	15,051	72,567
<i>plural</i>	1,522.56	152.03	265,602	33,550	299,152
<i>pull-stream</i>	2,506.81	1,378.63	208,130	75,239	283,369
<i>q</i>	5,332.46	14,311.02	2,127,655	217,886	2,345,541
<i>spacl-core</i>	1,351.01	850.39	162,705	28,919	191,624
<i>zip-a-folder</i>	900.98	471.62	72,362	9,438	81,800
<i>Total</i>	24,446.14	25,303.97	5,831,017	712,283	6,543,300

TABLE XXXV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #318). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,065.41	611.71	967,508	100,190	1,067,698
<i>countries-and-timezones</i>	1,070.87	313.43	105,828	22,897	128,725
<i>crawler-url-parser</i>	1,643.05	1,045.88	386,223	39,148	425,371
<i>delta</i>	3,028.61	4,054.65	890,252	99,121	989,373
<i>image-downloader</i>	430.56	496.08	24,655	8,793	33,448
<i>node-dirty</i>	1,527.72	252.26	246,248	33,054	279,302
<i>node-geo-point</i>	1,411.09	1,060.50	316,333	28,836	345,169
<i>node-jsonfile</i>	690.68	520.06	57,516	14,997	72,513
<i>plural</i>	1,522.17	146.72	265,602	33,944	299,546
<i>pull-stream</i>	2,517.82	1,365.08	208,130	75,400	283,530
<i>q</i>	5,232.53	13,865.94	2,127,655	217,305	2,344,960
<i>spacl-core</i>	1,351.05	848.71	162,705	28,593	191,298
<i>zip-a-folder</i>	500.59	1,105.81	82,457	10,544	93,001
<i>Total</i>	23,992.16	25,686.82	5,841,112	712,822	6,553,934

TABLE XXXVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #319). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,080.04	618.08	967,508	102,242	1,069,750
<i>countries-and-timezones</i>	1,070.88	311.67	105,828	22,556	128,384
<i>crawler-url-parser</i>	1,645.90	965.22	386,223	38,423	424,646
<i>delta</i>	2,992.70	3,820.99	890,252	99,184	989,436
<i>image-downloader</i>	430.53	500.27	24,655	9,240	33,895
<i>node-dirty</i>	1,527.92	252.60	246,248	32,911	279,159
<i>node-geo-point</i>	1,740.11	890.54	289,389	26,285	315,674
<i>node-jsonfile</i>	690.70	487.19	57,516	14,355	71,871
<i>plural</i>	1,677.14	158.69	249,979	30,944	280,923
<i>pull-stream</i>	2,510.30	1,455.39	208,130	75,369	283,499
<i>q</i>	5,453.21	14,010.80	2,127,655	217,999	2,345,654
<i>spacl-core</i>	1,351.08	861.29	162,705	28,654	191,359
<i>zip-a-folder</i>	500.56	1,278.26	82,457	10,677	93,134
<i>Total</i>	24,671.06	25,610.98	5,798,545	708,839	6,507,384

TABLE XXXVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #320). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,025.70	617.91	967,508	101,251	1,068,759
<i>countries-and-timezones</i>	1,070.86	333.07	105,828	23,224	129,052
<i>crawler-url-parser</i>	1,686.57	939.88	386,223	39,014	425,237
<i>delta</i>	3,020.99	3,992.69	890,252	99,683	989,935
<i>image-downloader</i>	430.56	347.98	24,655	9,059	33,714
<i>node-dirty</i>	1,526.16	227.18	246,248	32,693	278,941
<i>node-geo-point</i>	1,411.04	1,057.16	316,333	29,723	346,056
<i>node-jsonfile</i>	690.68	565.63	57,516	14,528	72,044
<i>plural</i>	1,521.10	156.75	265,602	34,049	299,651
<i>pull-stream</i>	2,516.94	1,385.89	208,130	75,071	283,201
<i>q</i>	5,280.14	14,131.09	2,127,655	217,911	2,345,566
<i>spacl-core</i>	1,350.99	879.83	162,705	28,767	191,472
<i>zip-a-folder</i>	510.57	542.42	82,457	10,709	93,166
<i>Total</i>	24,042.30	25,177.49	5,841,112	715,682	6,556,794

TABLE XXXVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #321). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	
	<i>LLMorpheus</i>	<i>StrykerJS</i>		compl.	total
<i>Complex.js</i>	3,058.33	648.14	967,508	100,962	1,068,470
<i>countries-and-timezones</i>	1,070.88	328.85	105,828	23,165	128,993
<i>crawler-url-parser</i>	1,645.45	966.36	386,223	38,649	424,872
<i>delta</i>	2,954.29	4,010.15	890,252	97,995	988,247
<i>image-downloader</i>	430.57	330.02	24,655	9,182	33,837
<i>node-dirty</i>	1,528.13	239.77	246,248	32,568	278,816
<i>node-geo-point</i>	1,411.08	1,024.88	316,333	29,330	345,663
<i>node-jsonfile</i>	690.67	543.04	57,516	14,833	72,349
<i>plural</i>	1,522.09	148.56	265,602	33,906	299,508
<i>pull-stream</i>	2,509.37	1,433.54	208,130	75,574	283,704
<i>q</i>	5,254.98	13,947.69	2,127,655	216,579	2,344,234
<i>spacel-core</i>	1,351.03	834.35	162,705	29,103	191,808
<i>zip-a-folder</i>	500.60	1,144.50	82,457	10,347	92,804
<i>Total</i>	23,927.46	25,599.86	5,841,112	712,193	6,553,305

TABLE XXXIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #322). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.5, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

D. Results for *codellama-34b-instruct*, full template, temperature 1.0

Tables XL–XLIX show the results for 5 experiments with the *codellama-34b-instruct* model at temperature 1.0 using the default prompt and system prompt shown in Figure 7.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,405	345	5	27	1,028	648	379	1	63.13
<i>countries-and-timezones</i>	106	300	108	2	4	186	156	30	0	83.87
<i>crawler-url-parser</i>	176	490	202	4	6	278	202	76	0	72.66
<i>delta</i>	462	1,250	519	11	22	698	583	83	32	88.11
<i>image-downloader</i>	42	118	42	1	0	75	53	22	0	70.67
<i>node-dirty</i>	154	438	175	9	8	246	150	84	12	65.85
<i>node-geo-point</i>	140	405	118	4	10	273	213	60	0	78.02
<i>node-jsonfile</i>	68	187	52	1	2	132	50	22	60	83.33
<i>plural</i>	153	417	96	8	14	299	229	69	1	76.92
<i>pull-stream</i>	351	999	239	8	9	743	461	235	47	68.37
<i>q</i>	1,051	3,012	1,057	23	33	1,899	147	1,671	81	12.01
<i>spacel-core</i>	134	381	147	9	7	218	180	38	0	82.57
<i>zip-a-folder</i>	49	142	45	0	1	96	54	38	4	60.42
<i>Total</i>	3,376	9,544	3,145	85	143	6,171	3,126	2,807	238	—

TABLE XL

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #341). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,382	363	9	23	987	611	376	0	61.90
<i>countries-and-timezones</i>	106	254	79	7	4	164	145	19	0	88.41
<i>crawler-url-parser</i>	176	497	228	6	6	257	177	80	0	68.87
<i>delta</i>	462	1,307	580	7	19	701	596	75	30	89.30
<i>image-downloader</i>	42	115	52	1	3	59	46	13	0	77.97
<i>node-dirty</i>	154	438	159	4	7	268	164	96	8	64.18
<i>node-geo-point</i>	140	404	123	1	10	270	209	61	0	77.41
<i>node-jsonfile</i>	68	199	42	4	3	150	67	16	67	89.33
<i>plural</i>	153	442	114	11	12	304	245	58	1	80.92
<i>pull-stream</i>	351	1,012	255	11	5	741	460	238	43	67.88
<i>q</i>	1,051	3,025	1,080	16	35	1,894	169	1,652	73	12.78
<i>spacel-core</i>	134	372	162	0	5	205	173	32	0	84.39
<i>zip-a-folder</i>	49	130	44	2	0	84	23	4	57	95.24
<i>Total</i>	3,376	9,577	3,281	79	132	6,084	3,085	2,720	279	—

TABLE XLI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #342). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,378	346	7	23	1,002	637	365	0	63.57
<i>countries-and-timezones</i>	106	297	94	1	4	198	177	21	0	89.39
<i>crawler-url-parser</i>	176	491	207	6	9	269	173	96	0	64.31
<i>delta</i>	462	1,296	535	11	23	727	611	85	31	88.31
<i>image-downloader</i>	42	118	39	1	0	78	70	8	0	89.74
<i>node-dirty</i>	154	437	180	4	8	245	136	97	12	60.41
<i>node-geo-point</i>	140	399	132	3	5	259	194	65	0	74.90
<i>node-jsonfile</i>	68	190	41	2	1	146	68	14	64	90.41
<i>plural</i>	153	429	109	9	16	295	236	58	1	80.34
<i>pull-stream</i>	351	1,008	249	12	7	740	443	248	49	66.49
<i>q</i>	1,051	3,048	1,040	19	47	1,941	147	1,707	87	12.06
<i>spacel-core</i>	134	378	143	5	5	225	196	29	0	87.11
<i>zip-a-folder</i>	49	140	44	2	0	94	28	4	62	95.74
<i>Total</i>	3,376	9,609	3,159	82	148	6,219	3,116	2,797	306	—

TABLE XLII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #343). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,385	318	3	22	1,042	667	374	1	64.11
<i>countries-and-timezones</i>	106	290	84	1	3	202	183	19	0	90.59
<i>crawler-url-parser</i>	176	504	221	4	3	276	200	76	0	72.46
<i>delta</i>	462	1,296	531	5	16	744	639	71	34	90.46
<i>image-downloader</i>	42	116	38	2	1	75	49	26	0	65.33
<i>node-dirty</i>	154	412	142	7	3	260	145	100	15	61.54
<i>node-geo-point</i>	140	398	127	2	10	259	205	54	0	79.15
<i>node-jsonfile</i>	68	192	46	1	2	143	56	24	63	83.22
<i>plural</i>	153	439	120	5	11	303	242	60	1	80.20
<i>pull-stream</i>	351	1,006	237	9	8	752	456	238	58	68.35
<i>q</i>	1,051	3,030	1,109	18	24	1,879	130	1,659	90	11.71
<i>spacl-core</i>	134	377	129	3	5	240	221	19	0	92.08
<i>zip-a-folder</i>	49	139	40	1	2	96	38	3	55	96.88
<i>Total</i>	3,376	9,584	3,142	61	110	6,271	3,231	2,723	317	—

TABLE XLIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #345). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,366	318	6	24	1,018	640	378	0	62.87
<i>countries-and-timezones</i>	106	298	88	1	2	207	177	30	0	85.51
<i>crawler-url-parser</i>	176	494	232	4	6	252	178	74	0	70.63
<i>delta</i>	462	1,315	617	4	13	679	564	83	32	87.78
<i>image-downloader</i>	42	122	41	0	6	75	56	17	2	77.33
<i>node-dirty</i>	154	435	165	2	1	267	155	103	9	61.42
<i>node-geo-point</i>	140	400	123	3	11	263	198	65	0	75.29
<i>node-jsonfile</i>	68	195	46	1	7	141	60	25	56	82.27
<i>plural</i>	153	438	109	11	10	308	245	60	3	80.52
<i>pull-stream</i>	351	999	246	10	2	741	457	237	47	68.02
<i>q</i>	1,051	3,024	1,106	14	36	1,868	130	1,652	86	11.56
<i>spacl-core</i>	134	379	149	3	3	224	202	21	1	90.63
<i>zip-a-folder</i>	49	145	40	1	2	102	28	1	73	99.02
<i>Total</i>	3,376	9,610	3,280	60	123	6,145	3,090	2,746	309	—

TABLE XLIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #347). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	3,090.47	545.31	967,508	101,516	1,069,024
<i>countries-and-timezones</i>	1,070.88	268.04	105,828	23,041	128,869
<i>crawler-url-parser</i>	1,786.52	902.86	377,644	36,825	414,469
<i>delta</i>	3,134.68	3,531.25	877,266	93,814	971,080
<i>image-downloader</i>	430.55	457.54	24,655	9,002	33,657
<i>node-dirty</i>	1,526.84	215.86	246,248	32,721	278,969
<i>node-geo-point</i>	1,411.05	897.39	316,333	29,494	345,827
<i>node-jsonfile</i>	710.73	477.34	57,516	14,447	71,963
<i>plural</i>	1,533.27	157.43	265,602	31,993	297,595
<i>pull-stream</i>	2,504.20	1,268.62	208,130	73,625	281,755
<i>q</i>	5,355.10	13,120.09	2,127,655	213,824	2,341,479
<i>spacl-core</i>	1,351.04	679.82	162,705	28,574	191,279
<i>zip-a-folder</i>	500.58	493.11	82,457	10,747	93,204
<i>Total</i>	24,405.89	23,014.65	5,819,547	699,623	6,519,170

TABLE XLV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #341). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,099.95	515.00	967,508	99,831	1,067,339
<i>countries-and-timezones</i>	1,537.50	234.01	96,352	19,743	116,095
<i>crawler-url-parser</i>	1,643.69	843.29	386,223	36,746	422,969
<i>delta</i>	3,119.28	3,448.69	890,252	95,930	986,182
<i>image-downloader</i>	430.53	351.02	24,655	8,785	33,440
<i>node-dirty</i>	1,526.21	216.70	246,248	31,856	278,104
<i>node-geo-point</i>	1,422.53	877.31	316,333	30,037	346,370
<i>node-jsonfile</i>	690.70	537.75	57,516	14,738	72,254
<i>plural</i>	1,521.12	159.03	265,602	32,288	297,890
<i>pull-stream</i>	2,659.93	1,222.06	208,130	73,902	282,032
<i>q</i>	5,264.59	12,818.11	2,127,655	212,677	2,340,332
<i>spacl-core</i>	1,361.03	682.26	162,705	27,970	190,675
<i>zip-a-folder</i>	500.58	908.74	82,457	9,890	92,347
<i>Total</i>	24,777.63	22,813.96	5,831,636	694,393	6,526,029

TABLE XLVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #342). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,024.06	525.74	967,508	98,687	1,066,195
<i>countries-and-timezones</i>	1,070.86	293.27	105,828	23,318	129,146
<i>crawler-url-parser</i>	1,644.85	845.43	386,223	36,638	422,861
<i>delta</i>	2,948.43	3,663.87	890,252	96,381	986,633
<i>image-downloader</i>	430.52	274.68	24,655	9,228	33,883
<i>node-dirty</i>	1,526.11	216.05	246,248	31,882	278,130
<i>node-geo-point</i>	1,411.07	852.26	316,333	29,152	345,485
<i>node-jsonfile</i>	690.73	549.73	57,516	13,670	71,186
<i>plural</i>	1,521.17	153.33	265,602	31,946	297,548
<i>pull-stream</i>	2,506.18	1,294.61	208,130	72,667	280,797
<i>q</i>	5,178.82	13,300.53	2,127,655	213,258	2,340,913
<i>spacl-core</i>	1,351.08	743.18	162,705	27,865	190,570
<i>zip-a-folder</i>	500.55	1,038.21	82,457	10,518	92,975
<i>Total</i>	23,804.42	23,750.90	5,841,112	695,210	6,536,322

TABLE XLVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #343). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,070.12	559.72	967,508	100,998	1,068,506
<i>countries-and-timezones</i>	1,070.86	285.32	105,828	22,109	127,937
<i>crawler-url-parser</i>	1,644.31	872.80	386,223	38,267	424,490
<i>delta</i>	2,974.80	3,703.16	890,252	96,594	986,846
<i>image-downloader</i>	430.54	438.80	24,655	8,660	33,315
<i>node-dirty</i>	1,527.36	248.76	246,248	31,479	277,727
<i>node-geo-point</i>	1,411.07	843.77	316,333	29,660	345,993
<i>node-jsonfile</i>	690.69	512.58	57,516	14,276	71,792
<i>plural</i>	1,521.47	156.91	265,602	32,805	298,407
<i>pull-stream</i>	2,495.10	1,354.27	208,130	73,802	281,932
<i>q</i>	5,350.94	13,107.50	2,127,655	213,504	2,341,159
<i>spacl-core</i>	1,351.11	811.62	162,705	28,193	190,898
<i>zip-a-folder</i>	500.61	968.74	82,457	10,354	92,811
<i>Total</i>	24,038.96	23,863.96	5,841,112	700,701	6,541,813

TABLE XLVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #345). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
<i>Complex.js</i>	3,021.09	531.51	967,508	100,183	1,067,691
<i>countries-and-timezones</i>	1,070.89	297.61	105,828	23,229	129,057
<i>crawler-url-parser</i>	1,642.94	776.31	386,223	36,920	423,143
<i>delta</i>	2,984.22	3,374.88	890,252	96,698	986,950
<i>image-downloader</i>	430.54	435.23	24,655	9,024	33,679
<i>node-dirty</i>	1,528.77	227.23	246,248	32,731	278,979
<i>node-geo-point</i>	1,411.05	858.36	316,333	29,774	346,107
<i>node-jsonfile</i>	690.66	494.30	57,516	14,127	71,643
<i>plural</i>	1,521.30	172.76	265,602	32,844	298,446
<i>pull-stream</i>	2,481.70	1,251.69	208,130	73,022	281,152
<i>q</i>	5,205.03	13,013.51	2,127,655	214,495	2,342,150
<i>spacel-core</i>	1,351.07	731.40	162,705	27,723	190,428
<i>zip-a-folder</i>	500.58	1,168.69	82,457	10,656	93,113
<i>Total</i>	23,839.84	23,333.49	5,841,112	701,426	6,542,538

TABLE XLIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #347). MODEL: *codellama-34b-instruct*, TEMPERATURE: 1.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

E. Results for variability analysis for *codellama-34b-instruct*

Tables L–LIII show the variability of the mutants observed in 5 runs using *codellama-34b-instruct*, for each of the temperature settings that were previously discussed in Sections A–D.

application	#min	#max	#distinct	#common
<i>Complex.js</i>	1,198	1,199	1,217	1,181 (97.04%)
<i>countries-and-timezones</i>	216	217	218	215 (98.62%)
<i>crawler-url-parser</i>	282	285	289	278 (96.19%)
<i>delta</i>	765	768	787	746 (94.79%)
<i>image-downloader</i>	89	90	90	89 (98.89%)
<i>node-dirty</i>	274	275	283	266 (93.99%)
<i>node-geo-point</i>	302	302	307	297 (96.74%)
<i>node-jsonfile</i>	153	154	158	149 (94.30%)
<i>plural</i>	279	281	288	274 (95.14%)
<i>pull-stream</i>	758	760	770	746 (96.88%)
<i>q</i>	2,031	2,035	2,053	2,014 (98.10%)
<i>spacel-core</i>	238	239	252	225 (89.29%)
<i>zip-a-folder</i>	100	100	102	98 (96.08%)

TABLE L

VARIABILITY OF THE MUTANTS GENERATED IN 5 RUNS OF *LLMorpheus* USING THE *codellama-34b-instruct* LLM AT TEMPERATURE 0.0 (RUN #312,#314,#315,#316,#317). THE COLUMNS OF THE TABLE SHOW, FROM LEFT TO RIGHT: (I) THE MINIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (II) THE MAXIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (III) THE TOTAL NUMBER OF DISTINCT MUTANTS OBSERVED IN ALL RUNS, AND (IV) THE NUMBER (PERCENTAGE) OF MUTANTS ARE OBSERVED IN ALL RUNS.

application	#min	#max	#distinct	#common
<i>Complex.js</i>	1,168	1,200	2,354	447 (18.99%)
<i>countries-and-timezones</i>	215	227	551	53 (9.62%)
<i>crawler-url-parser</i>	266	285	700	67 (9.57%)
<i>delta</i>	752	781	1,706	247 (14.48%)
<i>image-downloader</i>	82	88	202	27 (13.37%)
<i>node-dirty</i>	265	279	600	82 (13.67%)
<i>node-geo-point</i>	296	314	660	99 (15.00%)
<i>node-jsonfile</i>	151	158	369	42 (11.38%)
<i>plural</i>	273	291	678	79 (11.65%)
<i>pull-stream</i>	764	770	1,664	280 (16.83%)
<i>q</i>	1,980	2,050		

application	#min	#max	#distinct	#common
<i>Complex.js</i>	987	1,042	4,200	17 (0.40%)
<i>countries-and-timezones</i>	164	207	835	2 (0.24%)
<i>crawler-url-parser</i>	252	278	1,171	1 (0.09%)
<i>delta</i>	681	744	3,013	2 (0.07%)
<i>image-downloader</i>	59	78	298	1 (0.34%)
<i>node-dirty</i>	245	268	1,075	2 (0.19%)
<i>node-geo-point</i>	259	273	1,111	2 (0.18%)
<i>node-jsonfile</i>	132	150	606	1 (0.17%)
<i>plural</i>	295	308	1,316	5 (0.38%)
<i>pull-stream</i>	736	749	2,814	30 (1.07%)
<i>q</i>	1,868	1,942	7,695	43 (0.56%)
<i>spacl-core</i>	205	240	958	2 (0.21%)
<i>zip-a-folder</i>	84	102	366	6 (1.64%)

TABLE LIII

VARIABILITY OF THE MUTANTS GENERATED IN 5 RUNS OF *LLMorpheus* USING THE *codellama-34b-instruct* LLM AT TEMPERATURE 1.0 (RUN #341,#342,#343,#345,#347). THE COLUMNS OF THE TABLE SHOW, FROM LEFT TO RIGHT: (I) THE MINIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (II) THE MAXIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (III) THE TOTAL NUMBER OF DISTINCT MUTANTS OBSERVED IN ALL RUNS, AND (IV) THE NUMBER (PERCENTAGE) OF MUTANTS ARE OBSERVED IN ALL RUNS.

F. Results for *codellama-13b-instruct*

Tables LIV–LXIII show the results for 5 experiments with the *codellama-13b-instruct* model at temperature 0.0 using the default prompt and system prompt shown in Figure 7. Table LXIV shows the variability of the mutants observed in 5 runs using *codellama-13b-instruct* at temperature 0.0.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,438	339	116	28	955	553	401	1	58.01
<i>countries-and-timezones</i>	106	306	83	15	1	207	177	30	0	85.51
<i>crawler-url-parser</i>	176	506	186	51	12	247	129	118	0	52.23
<i>delta</i>	462									

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,439	340	116	28	955	553	401	1	58.01
<i>countries-and-timezones</i>	106	306	83	15	1	207	177	30	0	85.51
<i>crawler-url-parser</i>	176	506	186	51	12	247	129	118	0	52.23
<i>delta</i>	462	1,350	530	92	16	712	583	107	22	84.97
<i>image-downloader</i>	42	124	40	5	2	77	48	29	0	62.34
<i>node-dirty</i>	154	450	161	33	11	245	142	92	11	62.45
<i>node-geo-point</i>	140	406	64	21	16	304	237	67	0	77.96
<i>node-jsonfile</i>	68	198	45	10	6	137	43	45	49	67.15
<i>plural</i>	153	424	100	99	17	208	154	53	1	74.52
<i>pull-stream</i>	351	1,015	280	54	13	668	386	236	46	64.67
<i>q</i>	1,051	3,046	899	379	55	1,713	122	1,518	73	11.38
<i>spack-core</i>	134	384	142	40	7	185	160	25	0	86.49
<i>zip-a-folder</i>	49	138	43	7	1	87	27	55	5	36.78
<i>Total</i>	3,376	9,786	2,913	922	185	5,745	2,761	2,776	208	—

TABLE LVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #356). MODEL: *codellama-13b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts
-------------	----------

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,041.69	525.25	967,508	104,246	1,071,754
<i>countries-and-timezones</i>	1,070.79	310.80	105,828	23,971	129,799
<i>crawler-url-parser</i>	1,637.98	803.17	386,223	39,906	426,129
<i>delta</i>	2,993.70	3,529.05	890,252	103,085	993,337
<i>image-downloader</i>	430.56	460.64	24,655	9,339	33,994
<i>node-dirty</i>	1,526.96	211.68	246,248	34,892	281,140
<i>node-geo-point</i>	1,411.04	1,000.74	316,333	30,715	347,048
<i>node-jsonfile</i>	690.65	425.39	57,516	15,398	72,914
<i>plural</i>	1,521.00	111.98	265,602	34,926	300,528
<i>pull-stream</i>	2,489.93	1,188.59	208,130	77,308	285,438
<i>q</i>	5,187.67	11,850.22	2,127,655	231,175	2,358,830
<i>spacl-core</i>	1,350.97	616.43	162,705	30,694	193,399
<i>zip-a-folder</i>	500.51	496.32	82,457	11,494	93,951
<i>Total</i>	23,853.45	21,530.25	5,841,112	747,149	6,588,261

TABLE LIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #354). MODEL: *codellama-13b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,065.29	504.22	967,508	104,246	1,071,754
<i>countries-and-timezones</i>	1,070.80	296.35	105,828	23,971	129,799
<i>crawler-url-parser</i>	1,646.62	799.72	386,223	39,938	426,161
<i>delta</i>					

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,045.49	511.92	967,508	104,246	1,071,754
<i>countries-and-timezones</i>	1,070.84	308.11	105,828	23,971	129,799
<i>crawler-url-parser</i>	1,639.50	877.27	386,223	39,906	426,129
<i>delta</i>	2,975.67	3,533.28	890,252	103,025	993,277
<i>image-downloader</i>	430.50	458.40	24,655	9,339	33,994
<i>node-dirty</i>	1,527.47	209.67	246,248	34,892	281,140
<i>node-geo-point</i>	1,411.06	994.07	316,333	30,715	347,048
<i>node-jsonfile</i>	690.64	421.39	57,516	15,398	72,914
<i>plural</i>	1,521.06	112.22	265,602	34,926	300,528
<i>pull-stream</i>	2,493.00	1,182.21	208,130	77,307	285,437
<i>q</i>	5,177.80	11,862.47	2,127,655	231,214	2,358,869
<i>spacl-core</i>	1,350.98	606.49	162,705	30,694	193,399
<i>zip-a-folder</i>	500.52	486.21	82,457	11,494	93,951
<i>Total</i>	23,834.54	21,563.71	5,841,112	747,127	6,588,239

TABLE LXII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #358). MODEL: *codellama-13b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,065.97	514.43	967,508	104,246	1,071,754
<i>countries-and-timezones</i>	1,070.82	328.45	105,828	23,951	129,779
<i>crawler-url-parser</i>	1,638.27	905.63	386,223	39,906	426,129
<i>delta</i>	2,954.39				

G. Results for mixtral-8x7b-instruct

Tables LXV–LXXIV show the results for 5 experiments with the *mixtral-8x7b-instruct* model at temperature 0.0 using the default prompt and system prompt shown in Figure 7. Table LXXV shows the variability of the mutants observed in 5 runs using *mixtral-8x7b-instruct* at temperature 0.0.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,287	310	0	15	962	589	373	0	61.23
<i>countries-and-timezones</i>	106	277	65	2	5	205	166	39	0	80.98
<i>crawler-url-parser</i>	176	414	165	0	3	234	130	104	0	55.56
<i>delta</i>	462	1,156	452	0	24	680	516	128	36	81.18
<i>image-downloader</i>	42	108	38	0	1	69	46	23	0	66.67
<i>node-dirty</i>	154	310	109	0	10	191	111	72	8	62.30
<i>node-geo-point</i>	140	352	88	0	11	247	166	81	0	67.21
<i>node-jsonfile</i>	68	159	23	0	4	132	54	32	46	75.76
<i>plural</i>	153	307	73	0	8	226	166	60	0	73.45
<i>pull-stream</i>	351	940	255	1	6	678	386	248	44	63.42
<i>q</i>	1,051	2,468	772	3	50	1,643	112	1,460	71	11.14
<i>spacel-core</i>	134	333	152	0	3	157	134	22	1	85.99
<i>zip-a-folder</i>	49	117	38	0	0	78	24	44	10	43.59
<i>Total</i>	3,376	8,228	2,540	6	140	5,502	2,600	2,68		

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,299	305	0	25	969	587	382	0	60.58
<i>countries-and-timezones</i>	106	270	61	2	4	203	160	43	0	78.82
<i>crawler-url-parser</i>	176	401	167	0	1	220	118	102	0	53.64
<i>delta</i>	462	1,122	440	0	22	660	500	126	34	80.91
<i>image-downloader</i>	42	103	39	0	0	64	41	23	0	64.06
<i>node-dirty</i>	154	338	125	0	10	203	120	77	6	62.07
<i>node-geo-point</i>	140	358	81	0	15	258	178	80	0	68.99
<i>node-jsonfile</i>	68	166	25	1	5	135	56	32	47	76.30
<i>plural</i>	153	310	71	0	7	232	171	61	0	73.71
<i>pull-stream</i>	351	946	251	1	7	687	398	245	44	64.34
<i>q</i>	1,051	2,504	790	3	51	1,660	116	1,477	67	11.02
<i>spac-core</i>	134	329	148	0	3	157	134	22	1	85.99
<i>zip-a-folder</i>	49	116	38	0	0	78	27	42	9	46.15
<i>Total</i>	3,376	8,262	2,541	7	150	5,526	2,606	2,712	208	—

TABLE LXVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #363). MODEL: *mixtral-8x7b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt</*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,372.45	520.01	960,545	98,260	1,058,805
countries-and-timezones	1,071.57	285.14	104,291	22,228	126,519
crawler-url-parser	1,669.78	698.95	384,404	32,771	417,175
delta	3,199.74	3,398.17	882,477	89,050	971,527
image-downloader	432.08	378.08	24,140	8,005	32,145
node-dirty	1,554.61	185.18	242,671	26,837	269,508
node-geo-point	1,416.99	850.15	318,251	28,316	346,567
node-jsonfile	740.68	431.76	56,273	12,101	68,374
plural	1,546.99	120.70	261,626	25,293	286,919
pull-stream	2,660.01	1,259.31	204,431	70,142	274,573
q	6,149.24	11,173.93	2,103,232	188,223	2,291,455
spacl-core	1,416.02	534.70	162,695	26,751	189,446
zip-a-folder	500.53	489.08	81,279	9,372	90,651
Total	25,730.68	20,325.15	5,786,315	637,349	6,423,664

TABLE LXXI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #361). MODEL: *mixtral-8x7b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,359.29	519.50	960,545	96,727	1,057,272
countries-and-timezones	1,074.42	277.16	104,291	22,353	126,644
crawler-url-parser	1,661.26	731.82	384,404	32,772	417,176
delta	3,170.76	3,384.81	882,477	89,334	971,811
image-downloader	430.53	386.39	24,140	7,934	32,074
node-dirty	1,530.96	182.12	244,297	27,524	271,821
node-geo-point	1,413.38	829.76	318,251	27,995	346,246
node-jsonfile	690.64	444.51	56,273	11,970	68,243
plural	1,524.56	117.41	261,626	25,277	286,903
pull-stream	2,644.38	1,205.37	204,431	69,081	273,512
q	6,079.84	11,465.34	2,103,232	192,672	2,295,904
spacl-core	1,354.60	540.52	162,695	26,151	188,846
zip-a-folder	500.60	464.06	81,279	9,340	90,619
Total	25,435.21	20,548.77	5,787,941	639,130	6,427,071

TABLE LXXII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #362). MODEL: *mixtral-8x7b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,511.86	522.92	960,545	96,846	1,057,391
countries-and-timezones	1,073.78	299.00	104,291	22,090	126,381
crawler-url-parser	1,664.52	748.35	384,404	32,721	417,125
delta	3,343.71	3,332.02	882,477	88,421	970,898
image-downloader	440.52	362.20	24,140	7,972	32,112
node-dirty	1,532.40	164.86	244,297	26,801	271,098
node-geo-point	1,436.26	842.23	318,251	28,074	346,325
node-jsonfile	692.85	440.06	56,273	12,731	69,004
plural	1,525.78	117.87	261,626	25,198	286,824
pull-stream	2,725.42	1,227.21	204,431	70,751	275,182
q	6,622.34	11,507.75	2,103,232	194,705	2,297,937
spacl-core	1,359.32	532.62	162,695	26,100	188,795
zip-a-folder	500.56	463.12	81,279	9,118	90,397
Total	26,429.31	20,560.22	5,787,941	641,528	6,429,469

TABLE LXXIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #363). MODEL: *mixtral-8x7b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		#tokens		
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	total
<i>Complex.js</i>	3,649.24	528.79	960,545	98,038	1,058,583
<i>countries-and-timezones</i>	1,087.49	295.42	104,291	22,259	126,550
<i>crawler-url-parser</i>	1,672.78	705.16	384,404	32,640	417,044
<i>delta</i>	3,494.16	3,421.43	882,477	90,244	972,721
<i>image-downloader</i>	430.50	378.73	24,140	8,213	32,353
<i>node-dirty</i>	1,530.34	197.97	244,297	26,982	271,279
<i>node-geo-point</i>	1,432.72	865.00	318,251	28,015	346,266
<i>node-jsonfile</i>	702.02	419.40	56,273	11,348	67,621
<i>plural</i>	1,533.26	118.22	261,626	25,664	287,290
<i>pull-stream</i>	2,763.15	1,206.16	204,431	69,471	273,902
<i>q</i>	6,879.90	11,339.61	2,103,232	191,046	2,294,278
<i>spacl-core</i>	1,409.48	527.44	162,695	26,807	189,502
<i>zip-a-folder</i>	500.54	444.67	81,279	9,009	90,288
<i>Total</i>	27,085.58	20,447.99	5,787,941	639,736	6,427,677

TABLE LXXIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #364). MODEL: *mixtral-8x7b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

application	#min	#max	#distinct	#common
<i>Complex.js</i>	962	989	1,425	604 (42.39%)
<i>countries-and-timezones</i>	196	205	287	132 (45.99%)
<i>crawler-url-parser</i>	225	246	349	144 (41.26%)
<i>delta</i>	660	680	958	431 (44.99%)
<i>image-downloader</i>	64	69	95	42 (44.21%)
<i>node-dirty</i>	191	213	307	120 (39.09%)
<i>node-geo-point</i>	253	263	368	169 (45.92%)
<i>node-jsonfile</i>	121	135	187	81 (43.32%)
<i>plural</i>	226	232	374	128 (34.22%)
<i>pull-stream</i>	669	692	981	451 (45.97%)
<i>q</i>	1,609	1,660	2,438	999 (40.98%)
<i>spacl-core</i>	178	181	261	113 (43.30%)
<i>zip-a-folder</i>	75	86	112	56 (50.00%)

TABLE LXXV

VARIABILITY OF THE MUTANTS GENERATED IN 5 RUNS OF *LLMorpheus* USING THE *mixtral-8x7b-instruct* LLM AT TEMPERATURE 0.0 (RUN #360,#361,#362,#363,#364). THE COLUMNS OF THE TABLE SHOW, FROM LEFT TO RIGHT: (I) THE MINIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (II) THE MAXIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (III) THE TOTAL NUMBER OF DISTINCT MUTANTS OBSERVED IN ALL RUNS, AND (IV) THE NUMBER (PERCENTAGE) OF MUTANTS ARE OBSERVED IN ALL RUNS.

H. Results for gpt-4o-mini

Tables LXXVI–LXXXV show the results for 5 experiments with the *gpt-4o-mini* model at temperature 0.0 using the default prompt and system prompt shown in Figure 7. Table LXXXVI shows the variability of the mutants observed in 5 runs using *gpt-4o-mini* at temperature 0.0.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,470	446	0	38	986	596	390	0	60.45
<i>countries-and-timezones</i>	106	317	99	0	9	209	171	38	0	81.82
<i>crawler-url-parser</i>	176	527	227	0	11	275	181	94	0	65.82
<i>delta</i>	462	1,385	636	2	40	707	564	108	35	84.72
<i>image-downloader</i>	42	126	58	0	3	65	45	20	0	69.23
<i>node-dirty</i>	154	462	188	0	9	265	154	103	8	61.13
<i>node-geo-point</i>	140	419	86	0	20	311	225	86	0	72.35
<i>node-jsonfile</i>	68	204	44	0	6	154	64	26	64	83.12
<i>plural</i>	153	454	110	5	26	313	257	55	1	82.43
<i>pull-stream</i>	351	1,053	302	0	16	735	420	247	68	66.39
<i>q</i>	1,051	3,153	1,287	2	69	1,795	137	1,597	61	11.03
<i>spacel-core</i>	134	402	158	0	10	215	195	20	0	90.70
<i>zip-a-folder</i>	49	147	62	0	4	81	14	7	60	91.36
<i>Total</i>	3,376	10,119	3,703	9	261	6,111	3,023	2,791	297	—

TABLE LXXVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #58). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,470	453	0	35	982	597	385	0	60.79
<i>countries-and-timezones</i>	106	318	105	0	9	204	166	38	0	81.37
<i>crawler-url-parser</i>	176	527	238	0	9	266	175	91	0	65.79
<i>delta</i>	462	1,385	651	1	35	698	566	97	35	86.10
<i>image-downloader</i>	42	126	58	0	2	66	41	25	0	62.12
<i>node-dirty</i>	154	462	195	0	8	259	153	99	7	61.78
<i>node-geo-point</i>	140	420	88	0	25	305	217	88	0	71.15
<i>node-jsonfile</i>	68	204	43	0	6	155	67	24	64	84.52
<i>plural</i>	153	453	114	6	22	311	255	55	1	82.32
<i>pull-stream</i>	351	1,053	300	0	14	739	436	240	63	67.52
<i>q</i>	1,051	3,153	1,289	7	71	1,786	132	1,588	66	11.09
<i>spacel-core</i>	134	402	157	0	10	216	198	18	0	91.67
<i>zip-a-folder</i>	49	147	63	0	2	82	22	4	56	95.12
<i>Total</i>	3,376	10,120	3,754	14	248	6,069	3,025	2,752	292	—

TABLE LXXVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #59). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,470	455	0	33	982	594	388	0	60.49
<i>countries-and-timezones</i>	106	318	97	0	13	208	174	34	0	83.65
<i>crawler-url-parser</i>	176	527	236	0	9	268	178	90	0	66.42
<i>delta</i>	462	1,385	644	2	39	700	561	105	34	85.00
<i>image-downloader</i>	42	126	57	0	3	66	45	21	0	68.18
<i>node-dirty</i>	154	462	196	0	9	257	146	103	8	59.92
<i>node-geo-point</i>	140	419	83	0	23	310	221	89	0	71.29
<i>node-jsonfile</i>	68	204	45	0	7	152	67	25	60	83.55
<i>plural</i>	153	455	110	5	27	313	254	58	1	81.47
<i>pull-stream</i>	351	1,053	299	0	14	740	431	244	65	67.03
<i>q</i>	1,051	3,153	1,296	3	81	1,773	128	1,580	65	10.89
<i>spacel-core</i>	134	401	160	0	12	211	195	16	0	92.42
<i>zip-a-folder</i>	49	147	62	0	2	82	17	4	61	95.12
<i>Total</i>	3,376	10,120	3,740	10	272	6,062	3,011	2,757	294	—

TABLE LXXVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #60). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,470	443	0	36	991	603	388	0	60.85
<i>countries-and-timezones</i>	106	318	99	0	11	208	173	35	0	83.17
<i>crawler-url-parser</i>	176	528	235	0	12	268	176	92	0	65.67
<i>delta</i>	462	1,385	641	1	40	703	563	103	37	85.35
<i>image-downloader</i>	42	126	55	0	3	68	44	24	0	64.71
<i>node-dirty</i>	154	462	190	0	9	263	157	99	7	62.36
<i>node-geo-point</i>	140	419	85	0	27	305	218	87	0	71.48
<i>node-jsonfile</i>	68	204	46	0	7	151	66	24	61	84.11
<i>plural</i>	153	455	114	4	22	315	262	53	0	83.17
<i>pull-stream</i>	351	1,053	299	0	11	743	432	244	67	67.16
<i>q</i>	1,051	3,153	1,298	7	70	1,778	123	1,595	60	10.29
<i>spacel-core</i>	134	401	153	0	12	220	200	20	0	90.91
<i>zip-a-folder</i>	49	147	59	0	4	83	18	4	61	95.18
<i>Total</i>	3,376	10,121	3,717	12	264	6,096	3,035	2,768	293	—

TABLE LXXIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #61). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,468	441	0	39	988	603	385	0	61.03
<i>countries-and-timezones</i>	106	318	103	0	7	208	173	35	0	83.17
<i>crawler-url-parser</i>	176	528	235	0	9	271	181	90	0	66.79
<i>delta</i>	462	1,385	644	1	42	698	558	104	36	85.10
<i>image-downloader</i>	42	126	56	0	2	68	49	19	0	72.06
<i>node-dirty</i>	154	462	187	0	10	265	152	106	7	60.00
<i>node-geo-point</i>	140	419	86	0	24	307	216	91	0	70.36
<i>node-jsonfile</i>	68	204	46	0	7	151	60	30	61	80.13
<i>plural</i>	153	455	112	5	26	312	261	51	0	83.65
<i>pull-stream</i>	351	1,053	300	0	14	739	432	246	61	66.71
<i>q</i>	1,051	3,153	1,288	4	74	1,787	132	1,594	61	10.80
<i>spacel-core</i>	134	402	158	0	12	216	200	16	0	92.59
<i>zip-a-folder</i>	49	147	63	0	3	81	18	4	59	95.06
<i>Total</i>	3,376	10,120	3,719	10	269	6,091	3,035	2,771	285	—

TABLE LXXX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #63). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	3,126.98	305.91	788,189	96,972	885,161
<i>countries-and-timezones</i>	1,070.51	223.26	82,293	21,679	103,972
<i>crawler-url-parser</i>	1,663.34	529.03	288,452	36,368	324,820
<i>delta</i>	2,954.87	2,245.12	705,441	91,727	797,168
<i>image-downloader</i>	430.37	363.94	19,697	8,185	27,882
<i>node-dirty</i>	1,526.46	149.87	198,716	31,214	229,930
<i>node-geo-point</i>	1,534.64	731.35	254,246	26,052	280,298
<i>node-jsonfile</i>	692.89	470.97	46,918	13,532	60,450
<i>plural</i>	1,522.37	93.31	218,037	32,037	250,074
<i>pull-stream</i>	2,570.39	1,189.58	171,725	70,357	242,082
<i>q</i>	5,661.54	11,513.31	1,694,668	207,545	1,902,213
<i>spacel-core</i>	1,350.59	881.31	137,862	26,728	164,590
<i>zip-a-folder</i>	500.39	819.97	65,869	10,140	76,009
<i>Total</i>	24,605.34	19,516.91	4,672,113	672,536	5,344,649

TABLE LXXXI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #58). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,231.94	306.23	788,189	97,504	885,693
countries-and-timezones	1,071.30	212.00	82,293	21,875	104,168
crawler-url-parser	1,667.41	521.15	288,452	36,277	324,729
delta	3,101.18	2,248.14	705,441	91,878	797,319
image-downloader	430.36	374.81	19,697	8,207	27,904
node-dirty	1,536.82	141.77	198,716	31,186	229,902
node-geo-point	1,411.39	711.52	254,246	25,773	280,019
node-jsonfile	690.45	478.21	46,918	13,386	60,304
plural	1,589.51	93.88	218,037	32,039	250,076
pull-stream	2,599.70	1,147.49	171,725	70,365	242,090
q	5,669.28	11,541.55	1,694,668	207,750	1,902,418
spacl-core	1,353.11	900.14	137,862	26,648	164,510
zip-a-folder	502.44	765.68	65,869	10,124	75,993
Total	24,854.90	19,442.55	4,672,113	673,012	5,345,125

TABLE LXXXII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #59). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	2,839.62	303.29	788,189	97,203	885,392
countries-and-timezones	1,070.52	210.16	82,293	21,852	104,145
crawler-url-parser	1,646.04	509.26	288,452	36,258	324,710
delta	2,774.38	2,206.72	705,441	91,952	797,393
image-downloader	430.38	370.79	19,697	8,198	27,895
node-dirty	1,527.61	142.70	198,716	31,200	229,916
node-geo-point	1,410.64	716.58	254,246	25,833	280,079
node-jsonfile	690.45	453.09	46,918	13,455	60,373
plural	1,520.69	94.22	218,037	31,898	249,935
pull-stream	2,438.84	1,171.72	171,725	69,914	241,639
q	4,712.10	11,452.86	1,694,668	207,508	1,902,176
spacl-core	1,353.02	848.87	137,862	26,698	164,560
zip-a-folder	500.38	816.75	65,869	10,016	75,885
Total	22,914.67	19,297.01	4,672,113	671,985	5,344,098

TABLE LXXXIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #60). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	2,952.22	307.21	788,189	96,710	884,899
countries-and-timezones	1,070.49	220.94	82,293	21,677	103,970
crawler-url-parser	1,647.14	494.72	288,452	36,287	324,739
delta	2,877.05	2,256.62	705,441	91,533	796,974
image-downloader	430.38	371.91	19,697	8,175	27,872
node-dirty	1,527.58	142.81	198,716	31,147	229,863
node-geo-point	1,410.65	713.98	254,246	25,849	280,095
node-jsonfile	690.43	464.79	46,918	13,608	60,526
plural	1,527.23	88.79	218,037	31,950	249,987
pull-stream	2,481.90	1,186.52	171,725	70,548	242,273
q	4,829.57	11,485.59	1,694,668	207,651	1,902,319
spacl-core	1,350.62	891.52	137,862	26,626	164,488
zip-a-folder	500.41	820.40	65,869	10,073	75,942
Total	23,295.67	19,445.81	4,672,113	671,834	5,343,947

TABLE LXXXIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #61). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		#tokens		
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	total
<i>Complex.js</i>	2,993.67	305.57	788,189	96,526	884,715
<i>countries-and-timezones</i>	1,070.53	212.32	82,293	21,735	104,028
<i>crawler-url-parser</i>	1,650.74	510.52	288,452	36,399	324,851
<i>delta</i>	3,461.17	2,263.13	705,441	92,146	797,587
<i>image-downloader</i>	430.40	376.81	19,697	8,150	27,847
<i>node-dirty</i>	1,531.95	145.81	198,716	31,238	229,954
<i>node-geo-point</i>	1,619.27	717.56	254,246	25,924	280,170
<i>node-jsonfile</i>	691.27	455.30	46,918	13,572	60,490
<i>plural</i>	1,527.63	89.76	218,037	31,989	250,026
<i>pull-stream</i>	2,527.67	1,150.99	171,725	70,511	242,236
<i>q</i>	5,162.96	11,468.80	1,694,668	207,007	1,901,675
<i>spacel-core</i>	1,350.59	885.50	137,862	26,516	164,378
<i>zip-a-folder</i>	501.32	792.67	65,869	10,026	75,895
<i>Total</i>	24,519.16	19,374.74	4,672,113	671,739	5,343,852

TABLE LXXXV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #63). MODEL: *gpt-4o-mini*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

application	#min	#max	#distinct	#common
<i>Complex.js</i>	982	991	1,484	613 (41.31%)
<i>countries-and-timezones</i>	204	209	362	95 (26.24%)
<i>crawler-url-parser</i>	280	289	472	153 (32.42%)
<i>delta</i>	698	707	1,094	413 (37.75%)
<i>image-downloader</i>	65	68	127	31 (24.41%)
<i>node-dirty</i>	257	265	417	147 (35.25%)
<i>node-geo-point</i>	307	313	492	189 (38.41%)
<i>node-jsonfile</i>	151	155	250	80 (32.00%)
<i>plural</i>	311	315	585	145 (24.79%)
<i>pull-stream</i>	728	734	1,176	411 (34.95%)
<i>q</i>	1,773	1,795	2,740	1,093 (39.89%)
<i>spacel-core</i>	229	236	408	109 (26.72%)
<i>zip-a-folder</i>	81	84	131	50 (38.17%)

TABLE LXXXVI

VARIABILITY OF THE MUTANTS GENERATED IN 5 RUNS OF *LLMorpheus* USING THE *gpt-4o-mini* LLM AT TEMPERATURE 0.0 (RUN #58,#59,#60,#61,#63). THE COLUMNS OF THE TABLE SHOW, FROM LEFT TO RIGHT: (I) THE MINIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (II) THE MAXIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (III) THE TOTAL NUMBER OF DISTINCT MUTANTS OBSERVED IN ALL RUNS, AND (IV) THE NUMBER (PERCENTAGE) OF MUTANTS ARE OBSERVED IN ALL RUNS.

I. Results for llama-3.3-70b-instruct

Tables LXXXVII–XCVI show the results for 5 experiments with the *llama-3.3-70b-instruct* model at temperature 0.0 using the default prompt and system prompt shown in Figure 7. Table XCVII shows the variability of the mutants observed in 5 runs using *gpt-4o-mini* at temperature 0.0.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,470	279	0	53	1,138	690	447	1	60.72
<i>countries-and-timezones</i>	106	318	64	0	14	240	207	33	0	86.25
<i>crawler-url-parser</i>	176	528	175	0	22	319	208	111	0	65.20
<i>delta</i>	462	1,387	539	0	54	794	626	130	38	83.63
<i>image-downloader</i>	42	127	43	0	5	79	54	25	0	68.35
<i>node-dirty</i>	154	463	121	1	10	331	168	142	21	57.10
<i>node-geo-point</i>	140	421	39	0	22	358	255	103	0	71.23
<i>node-jsonfile</i>	68	204	25	0	6	173	64	37	72	78.61
<i>plural</i>	153	456	96	0	29	331	244	87	0	73.72
<i>pull-stream</i>	351	1,054	262	0	10	782	465	265	52	66.11
<i>q</i>	1,051	3,154	855	0	80	2,219	127	2,006	86	9.60
<i>spacel-core</i>	134	401	134	0	18	236	203	32	1	86.44
<i>zip-a-folder</i>	49	147	24	0	2	121	87	5	29	95.87
<i>Total</i>	3,376	10,130	2,656	1	325	7,121	3,398	3,423	300	—

TABLE LXXXVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #23). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,471	292	0	56	1,123	677	445	1	60.37
<i>countries-and-timezones</i>	106	318	65	0	13	240	205	35	0	85.42
<i>crawler-url-parser</i>	176	529	182	0	21	315	201	114	0	63.81
<i>delta</i>	462	1,388	537	0	48	803	640	126	37	84.31
<i>image-downloader</i>	42	126	44	0	5	77	45	32	0	58.44
<i>node-dirty</i>	154	463	128	1	14	320	160	141	19	55.94
<i>node-geo-point</i>	140	420	40	0	26	351	256	95	0	72.93
<i>node-jsonfile</i>	68	204	28	0	5	171	63	39	69	77.19
<i>plural</i>	153	455	91	0	25	339	239	100	0	70.50
<i>pull-stream</i>	351	1,053	265	0	12	776	462	262	52	66.24
<i>q</i>	1,051	3,155	863	0	87	2,205	124	1,995	86	9.52
<i>spacel-core</i>	134	401	131	0	16	244	210	33	1	86.48
<i>zip-a-folder</i>	49	147	25	0	2	120	18	5	97	95.83
<i>Total</i>	3,376	10,130	2,691	1	330	7,084	3,300	3,422	362	—

TABLE LXXXVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #24). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,470	285	0	59	1,126	678	447	1	60.30
<i>countries-and-timezones</i>	106	318	67	0	11	240	199	41	0	82.92
<i>crawler-url-parser</i>	176	528	179	0	24	315	203	112	0	64.44
<i>delta</i>	462	1,387	542	0	47	798	625	134	39	83.21
<i>image-downloader</i>	42	126	46	0	4	76	45	31	0	59.21
<i>node-dirty</i>	154	465	126	2	11	326	160	146	20	55.21
<i>node-geo-point</i>	140	421	39	0	19	360	257	103	0	71.39
<i>node-jsonfile</i>	68	204	28	0	5	171	62	41	68	76.02
<i>plural</i>	153	455	93	0	26	336	245	91	0	72.92
<i>pull-stream</i>	351	1,053	270	0	8	775	467	256	52	66.97
<i>q</i>	1,051	3,155	837	1	80	2,237	129	2,027	81	9.39
<i>spacel-core</i>	134	401	131	0	15	241	205	35	1	85.48
<i>zip-a-folder</i>	49	147	27	0	2	118	17	5	96	95.76
<i>Total</i>	3,376	10,130	2,670	3	311	7,119	3,292	3,469	358	—

TABLE LXXXIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #25). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,473	273	0	64	1,136	683	452	1	60.21
<i>countries-and-timezones</i>	106	318	62	0	13	243	208	35	0	85.60
<i>crawler-url-parser</i>	176	527	178	0	21	316	201	115	0	63.61
<i>delta</i>	462	1,388	545	0	50	793	634	128	31	83.86
<i>image-downloader</i>	42	128	43	0	4	81	57	24	0	70.37
<i>node-dirty</i>	154	464	121	1	10	332	174	140	18	57.83
<i>node-geo-point</i>	140	421	40	0	17	361	256	105	0	70.91
<i>node-jsonfile</i>	68	204	27	0	6	171	63	38	70	77.78
<i>plural</i>	153	460	92	0	29	339	248	91	0	73.16
<i>pull-stream</i>	351	1,053	266	0	9	778	464	266	48	65.81
<i>q</i>	1,051	3,155	853	1	83	2,218	131	2,007	80	9.51
<i>spacl-core</i>	134	401	124	0	14	250	215	34	1	86.40
<i>zip-a-folder</i>	49	147	28	0	2	117	85	5	27	95.73
<i>Total</i>	3,376	10,139	2,652	2	322	7,135	3,419	3,440	276	—

TABLE XC

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #26). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,477	283	1	65	1,128	680	447	1	60.37
<i>countries-and-timezones</i>	106	318	67	0	12	239	201	38	0	84.10
<i>crawler-url-parser</i>	176	528	179	0	21	318	203	115	0	63.84
<i>delta</i>	462	1,387	539	0	47	801	643	120	38	85.02
<i>image-downloader</i>	42	127	42	0	5	80	57	23	0	71.25
<i>node-dirty</i>	154	463	124	1	10	328	172	140	16	57.32
<i>node-geo-point</i>	140	421	36	0	22	360	264	96	0	73.33
<i>node-jsonfile</i>	68	204	25	0	4	175	67	36	72	79.43
<i>plural</i>	153	453	89	0	26	338	249	89	0	73.67
<i>pull-stream</i>	351	1,053	269	0	8	776	467	259	50	66.62
<i>q</i>	1,051	3,155	850	0	80	2,225	124	2,022	79	9.12
<i>spacl-core</i>	134	401	128	0	16	244	206	37	1	84.84
<i>zip-a-folder</i>	49	148	26	0	2	120	18	5	97	95.83
<i>Total</i>	3,376	10,135	2,657	2	318	7,132	3,351	3,427	354	—

TABLE XCI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #27). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	2,035.43	359.18	785,856	99,437	885,293
<i>countries-and-timezones</i>	503.27	250.43	82,111	20,809	102,920
<i>crawler-url-parser</i>	810.25	606.98	283,056	35,138	318,194
<i>delta</i>	1,881.70	2,576.74	700,458	90,055	790,513
<i>image-downloader</i>	199.98	362.58	19,515	7,900	27,415
<i>node-dirty</i>	739.35	257.29	197,447	30,494	227,941
<i>node-geo-point</i>	688.13	840.31	252,643	27,340	279,983
<i>node-jsonfile</i>	334.39	527.30	46,029	13,579	59,608
<i>plural</i>	748.74	93.78	218,344	32,058	250,402
<i>pull-stream</i>	1,483.02	1,141.94	170,717	68,203	238,920
<i>q</i>	3,563.55	14,535.02	1,690,971	205,818	1,896,789
<i>spacl-core</i>	631.83	898.43	136,724	26,473	163,197
<i>zip-a-folder</i>	260.98	451.69	64,841	10,049	74,890
<i>Total</i>	13,880.61	22,901.67	4,648,712	667,353	5,316,065

TABLE XCII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #23). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	1,935.55	354.56	785,902	98,758	884,660
countries-and-timezones	517.60	248.81	82,111	21,228	103,339
crawler-url-parser	862.45	606.56	283,056	35,246	318,302
delta	1,894.76	2,611.62	700,458	90,310	790,768
image-downloader	197.05	453.31	19,515	7,908	27,423
node-dirty	714.96	253.07	197,447	30,659	228,106
node-geo-point	700.66	831.74	252,643	27,406	280,049
node-jsonfile	325.93	518.65	46,029	13,248	59,277
plural	791.02	95.56	218,344	32,292	250,636
pull-stream	1,417.60	1,132.73	170,717	68,708	239,425
q	3,506.09	14,435.66	1,690,971	205,518	1,896,489
spacl-core	665.00	997.34	136,724	26,424	163,148
zip-a-folder	262.84	1,296.03	64,841	10,226	75,067
Total	13,791.51	23,835.63	4,648,758	667,931	5,316,689

TABLE XCIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #24). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	1,951.37	356.92	785,856	99,172	885,028
countries-and-timezones	522.56	246.91	82,111	21,238	103,349
crawler-url-parser	804.63	604.74	283,056	35,252	318,308
delta	1,862.31	2,566.04	700,458	90,059	790,517
image-downloader	195.17	449.66	19,515	7,845	27,360
node-dirty	724.87	254.24	197,447	31,135	228,582
node-geo-point	664.03	860.50	252,643	27,500	280,143
node-jsonfile	353.83	501.73	46,029	13,505	59,534
plural	742.45	96.06	218,344	32,077	250,421
pull-stream	1,454.40	1,137.23	170,717	68,279	238,996
q	3,808.92	14,578.83	1,690,971	204,821	1,895,792
spacl-core	632.50	974.79	136,724	26,240	162,964
zip-a-folder	261.77	1,272.87	64,841	10,055	74,896
Total	13,978.80	23,900.50	4,648,712	667,178	5,315,890

TABLE XCIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #25). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	1,965.89	356.55	785,856	99,239	885,095
countries-and-timezones	520.83	252.88	82,111	21,050	103,161
crawler-url-parser	845.23	415.43	283,056	35,204	318,260
delta	1,820.88	2,526.25	700,458	90,336	790,794
image-downloader	199.63	376.57	19,515	8,115	27,630
node-dirty	732.92	248.37	197,447	30,614	228,061
node-geo-point	648.30	852.65	252,643	27,289	279,932
node-jsonfile	337.73	504.46	46,029	13,500	59,529
plural	771.86	96.38	218,344	32,276	250,620
pull-stream	1,482.72	1,117.82	170,717	68,160	238,877
q	3,818.36	14,500.22	1,690,971	205,244	1,896,215
spacl-core	661.68	1,006.28	136,724	26,127	162,851
zip-a-folder	261.75	423.42	64,841	10,119	74,960
Total	14,067.79	22,677.29	4,648,712	667,273	5,315,985

TABLE XCV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #26). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	1,949.16	359.45	785,856	98,758	884,614
<i>countries-and-timezones</i>	518.42	249.03	82,111	21,290	103,401
<i>crawler-url-parser</i>	843.83	610.56	283,056	35,543	318,599
<i>delta</i>	1,930.29	2,578.87	700,458	90,277	790,735
<i>image-downloader</i>	199.97	357.91	19,515	7,985	27,500
<i>node-dirty</i>	711.46	241.03	197,447	30,092	227,539
<i>node-geo-point</i>	664.09	878.84	252,643	27,234	279,877
<i>node-jsonfile</i>	334.95	526.44	46,029	13,288	59,317
<i>plural</i>	756.46	95.20	218,344	31,877	250,221
<i>pull-stream</i>	1,444.48	1,125.62	170,717	68,975	239,692
<i>q</i>	3,636.06	14,547.79	1,690,971	205,379	1,896,350
<i>spacl-core</i>	641.07	1,006.59	136,724	26,304	163,028
<i>zip-a-folder</i>	255.10	1,275.79	64,841	9,904	74,745
<i>Total</i>	13,885.34	23,853.13	4,648,712	666,906	5,315,618

TABLE XCVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #27). MODEL: *meta-llama/llama-3.3-70b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

application	#min	#max	#distinct	#common
<i>Complex.js</i>	1,123	1,138	1,712	698 (40.77%)
<i>countries-and-timezones</i>	239	243	374	139 (37.17%)
<i>crawler-url-parser</i>	325	331	559	186 (33.27%)
<i>delta</i>	793	803	1,222	492 (40.26%)
<i>image-downloader</i>	76	81	131	40 (30.53%)
<i>node-dirty</i>	320	332	511	189 (36.99%)
<i>node-geo-point</i>	354	364	533	232 (43.53%)
<i>node-jsonfile</i>	171	175	273	97 (35.53%)
<i>plural</i>	331	339	598	169 (28.26%)
<i>pull-stream</i>	768	775	1,158	500 (43.18%)
<i>q</i>	2,205	2,237	2,898	1,708 (58.94%)
<i>spacl-core</i>	249	263	445	131 (29.44%)
<i>zip-a-folder</i>	117	121	172	78 (45.35%)

TABLE XCVII

VARIABILITY OF THE MUTANTS GENERATED IN 5 RUNS OF *LLMorpheus* USING THE *meta-llama/llama-3.3-70b-instruct* LLM AT TEMPERATURE 0.0 (RUN #23,#24,#25,#26,#27). THE COLUMNS OF THE TABLE SHOW, FROM LEFT TO RIGHT: (I) THE MINIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (II) THE MAXIMUM NUMBER OF MUTANTS OBSERVED IN ANY OF THE RUNS, (III) THE TOTAL NUMBER OF DISTINCT MUTANTS OBSERVED IN ALL RUNS, AND (IV) THE NUMBER (PERCENTAGE) OF MUTANTS ARE OBSERVED IN ALL RUNS.

J. Results for *template-onemutation-0.0*

Tables XCVIII–CVII show the results for 5 experiments with the *codellama-34b-instruct* model at temperature 0.0 using the prompt template of Figure 15 and using the system prompt shown in Figure 7.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	490	67	6	11	406	245	161	0	60.34
<i>countries-and-timezones</i>	106	106	24	1	2	79	65	14	0	82.28
<i>crawler-url-parser</i>	176	175	70	8	5	86	50	36	0	58.14
<i>delta</i>	462	461	182	4	9	266	221	37	8	86.09
<i>image-downloader</i>	42	42	8	0	0	34	26	8	0	76.47
<i>node-dirty</i>	154	155	50	3	3	99	55	41	3	58.59
<i>node-geo-point</i>	140	140	32	0	3	104	74	30	0	71.15
<i>node-jsonfile</i>	68	68	11	0	0	57	18	18	21	68.42
<i>plural</i>	153	153	39	8	6	100	70	30	0	70.00
<i>pull-stream</i>	351	351	67	3	1	280	165	95	20	66.07
<i>q</i>	1,051	1,052	306	26	17	703	46	630	27	10.38
<i>spacel-core</i>	134	134	41	3	2	80	63	17	0	78.75
<i>zip-a-folder</i>	49	49	9	0	1	39	19	17	3	56.41
<i>Total</i>	3,376	3,376	906	62	60	2,333	1,117	1,134	82	—

TABLE XCVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #365). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	490	67	6	11	406	244	162	0	60.10
<i>countries-and-timezones</i>	106	106	24	1	2	79	66	13	0	83.54
<i>crawler-url-parser</i>	176	175	70	8	5	86	50	36	0	58.14
<i>delta</i>	462	461	181	4	9	267	222	37	8	86.14
<i>image-downloader</i>	42	42	8	0	0	34	26	8	0	76.47
<i>node-dirty</i>	154	155	50	3	3	99	55	41	3	58.59
<i>node-geo-point</i>	140	140	32	0	3	104	74	30	0	71.15
<i>node-jsonfile</i>	68	68	11	0	0	57	18	18	21	68.42
<i>plural</i>	153	153	39	8	6	100	70	30	0	70.00
<i>pull-stream</i>	351	351	67	3	1	280	164	96	20	65.71
<i>q</i>	1,051	1,052	306	26	17	703	46	630	27	10.38
<i>spacel-core</i>	134	134	41	3	2	80	63	17	0	78.75
<i>zip-a-folder</i>	49	49	9	0	1	39	19	17	3	56.41
<i>Total</i>	3,376	3,376	905	62	60	2,334	1,117	1,135	82	—

TABLE XCIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #366). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	485	67	6	11	401	241	160	0	60.10
<i>countries-and-timezones</i>	106	106	24	1	2	79	66	13	0	83.54
<i>crawler-url-parser</i>	176	175	71	8	5	85	50	35	0	58.82
<i>delta</i>	462	461	181	4	9	267	222	37	8	86.14
<i>image-downloader</i>	42	42	8	0	0	34	26	8	0	76.47
<i>node-dirty</i>	154	155	50	3	3	99	55	41	3	58.59
<i>node-geo-point</i>	140	140	32	0	3	104	74	30	0	71.15
<i>node-jsonfile</i>	68	68	11	0	0	57	18	18	21	68.42
<i>plural</i>	153	153	39	7	6	101	71	30	0	70.30
<i>pull-stream</i>	351	351	67	3	1	280	165	95	20	66.07
<i>q</i>	1,051	1,052	306	26	16	704	45	632	27	10.23
<i>spacel-core</i>	134	134	41	3	2	80	63	17	0	78.75
<i>zip-a-folder</i>	49	49	9	0	1	39	19	17	3	56.41
<i>Total</i>	3,376	3,371	906	61	59	2,330	1,115	1,133	82	—

TABLE C

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #369). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	490	67	6	11	406	245	161	0	60.34
<i>countries-and-timezones</i>	106	106	24	1	2	79	66	13	0	83.54
<i>crawler-url-parser</i>	176	175	69	8	5	87	51	36	0	58.62
<i>delta</i>	462	461	182	4	9	266	221	37	8	86.09
<i>image-downloader</i>	42	42	8	0	0	34	26	8	0	76.47
<i>node-dirty</i>	154	155	50	4	3	98	55	40	3	59.18
<i>node-geo-point</i>	140	140	32	0	3	104	74	30	0	71.15
<i>node-jsonfile</i>	68	68	11	0	0	57	18	18	21	68.42
<i>plural</i>	153	153	39	7	6	101	71	30	0	70.30
<i>pull-stream</i>	351	351	67	3	1	280	165	95	20	66.07
<i>q</i>	1,051	1,052	306	26	17	703	46	630	27	10.38
<i>spacel-core</i>	134	134	41	3	2	80	63	17	0	78.75
<i>zip-a-folder</i>	49	49	9	0	1	39	19	17	3	56.41
<i>Total</i>	3,376	3,376	905	62	60	2,334	1,120	1,132	82	—

TABLE CI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #370). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	490	67	6	11	406	245	161	0	60.34
<i>countries-and-timezones</i>	106	106	24	1	2	79	66	13	0	83.54
<i>crawler-url-parser</i>	176	175	69	8	5	87	51	36	0	58.62
<i>delta</i>	462	461	182	4	9	266	221	37	8	86.09
<i>image-downloader</i>	42	42	8	0	0	34	26	8	0	76.47
<i>node-dirty</i>	154	155	50	4	3	98	55	40	3	59.18
<i>node-geo-point</i>	140	140	32	0	3	104	74	30	0	71.15
<i>node-jsonfile</i>	68	68	11	0	0	57	18	18	21	68.42
<i>plural</i>	153	153	39	7	6	101	71	30	0	70.30
<i>pull-stream</i>	351	351	67	3	1	280	165	95	20	66.07
<i>q</i>	1,051	1,052	306	26	17	703	46	630	27	10.38
<i>spacel-core</i>	134	134	41	3	2	80	63	17	0	78.75
<i>zip-a-folder</i>	49	49	9	0	1	39	19	17	3	56.41
<i>Total</i>	3,376	3,376	905	62	60	2,334	1,120	1,132	82	—

TABLE CII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #371). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	total
<i>Complex.js</i>	2,784.11	210.86	927,818	39,567	967,385
<i>countries-and-timezones</i>	1,071.07	117.59	97,242	8,518	105,760
<i>crawler-url-parser</i>	1,636.44	292.18	371,967	15,504	387,471
<i>delta</i>	2,676.03	1,251.08	852,830	37,401	890,231
<i>image-downloader</i>	430.61	139.23	21,253	3,459	24,712
<i>node-dirty</i>	1,526.39	77.95	233,774	12,906	246,680
<i>node-geo-point</i>	1,411.29	330.28	304,993	11,192	316,185
<i>node-jsonfile</i>	690.81	183.43	52,008	5,846	57,854
<i>plural</i>	1,521.37	54.03	253,209	13,450	266,659
<i>pull-stream</i>	2,400.61	499.06	179,699	30,228	209,927
<i>q</i>	4,195.04	4,866.38	2,042,524	82,318	2,124,842
<i>spacel-core</i>	1,351.30	271.81	151,851	10,803	162,654
<i>zip-a-folder</i>	500.63	219.06	78,488	4,405	82,893
<i>Total</i>	22,195.69	8,512.94	5,567,656	275,597	5,843,253

TABLE CIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #365). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
Complex.js	2,757.00	215.00	927,818	39,486	967,304
countries-and-timezones	1,091.07	116.98	97,242	8,527	105,769
crawler-url-parser	1,636.38	300.85	371,967	15,532	387,499
delta	2,681.00	1,235.39	852,830	37,383	890,213
image-downloader	460.67	139.42	21,253	3,476	24,729
node-dirty	1,526.36	75.52	233,774	12,907	246,681
node-geo-point	1,411.28	327.69	304,993	11,211	316,204
node-jsonfile	730.85	184.62	52,008	5,779	57,787
plural	1,521.30	53.73	253,209	13,418	266,627
pull-stream	2,397.86	497.98	179,699	30,310	210,009
q	4,204.99	4,839.22	2,042,524	82,262	2,124,786
spacl-core	1,351.25	273.58	151,851	10,809	162,660
zip-a-folder	500.62	219.57	78,488	4,403	82,891
Total	22,270.63	8,479.55	5,567,656	275,503	5,843,159

TABLE CIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #366). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
Complex.js	2,820.16	211.48	916,945	39,061	956,006
countries-and-timezones	1,071.10	117.43	97,242	8,548	105,790
crawler-url-parser	1,636.46	288.17	371,967	15,519	387,486
delta	2,686.73	1,239.88	852,830	37,432	890,262
image-downloader	430.69	142.46	21,253	3,475	24,728
node-dirty	1,536.37	75.72	233,774	12,869	246,643
node-geo-point	1,411.33	338.80	304,993	11,209	316,202
node-jsonfile	690.78	183.73	52,008	5,845	57,853
plural	1,521.35	54.18	253,209	13,392	266,601
pull-stream	2,398.19	499.73	179,699	30,182	209,881
q	4,188.82	4,807.59	2,042,524	82,120	2,124,644
spacl-core	1,351.23	272.94	151,851	10,813	162,664
zip-a-folder	500.64	222.35	78,488	4,405	82,893
Total	22,243.85	8,454.47	5,556,783	274,870	5,831,653

TABLE CV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #369). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
Complex.js	2,754.08	218.66	927,818	39,566	967,384
countries-and-timezones	1,071.03	119.64	97,242	8,579	105,821
crawler-url-parser	1,636.41	281.15	371,967	15,525	387,492
delta	2,676.11	1,249.77	852,830	37,449	890,279
image-downloader	430.60	138.45	21,253	3,475	24,728
node-dirty	1,526.42	73.92	233,774	12,859	246,633
node-geo-point	1,411.28	329.40	304,993	11,210	316,203
node-jsonfile	690.77	183.71	52,008	5,787	57,795
plural	1,521.36	53.81	253,209	13,434	266,643
pull-stream	2,397.58	499.25	179,699	30,160	209,859
q	4,211.46	4,819.77	2,042,524	82,203	2,124,727
spacl-core	1,361.19	269.34	151,851	10,818	162,669
zip-a-folder	500.61	217.85	78,488	4,400	82,888
Total	22,188.89	8,454.73	5,567,656	275,465	5,843,121

TABLE CVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #370). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	
	<i>LLMorpheus</i>	<i>StrykerJS</i>		compl.	total
<i>Complex.js</i>	2,763.09	214.80	927,818	39,505	967,323
<i>countries-and-timezones</i>	1,071.04	115.36	97,242	8,565	105,807
<i>crawler-url-parser</i>	1,636.32	285.19	371,967	15,616	387,583
<i>delta</i>	2,676.35	1,249.33	852,830	37,349	890,179
<i>image-downloader</i>	430.63	137.29	21,253	3,461	24,714
<i>node-dirty</i>	1,526.34	74.50	233,774	12,868	246,642
<i>node-geo-point</i>	1,411.31	337.02	304,993	11,183	316,176
<i>node-jsonfile</i>	690.79	183.54	52,008	5,774	57,782
<i>plural</i>	1,521.35	52.33	253,209	13,401	266,610
<i>pull-stream</i>	2,403.01	497.08	179,699	30,238	209,937
<i>q</i>	4,196.27	4,832.34	2,042,524	82,120	2,124,644
<i>spack-core</i>	1,351.29	269.13	151,851	10,793	162,644
<i>zip-a-folder</i>	500.65	220.66	78,488	4,403	82,891
<i>Total</i>	22,178.44	8,468.57	5,567,656	275,276	5,842,932

TABLE CVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #371). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-onemutation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

```

Your task is to apply mutation testing to the following code:
'''
{{{code}}}
'''

by replacing the PLACEHOLDER with a buggy code fragment that has different
behavior than the original code fragment, which was:
'''
{{{orig}}}
'''

Please consider changes such as using different operators, changing constants,
referring to different variables, object properties, functions, or methods.

Provide your answer as a fenced code block containing a single line of code,
using the following template:

The PLACEHOLDER can be replaced with:
'''
<code fragment>
'''

This would result in different behavior because <brief explanation>.

Please conclude your response with "DONE."

```

Fig. 15. Variation on the template of Figure 7 that requests only one mutation.

K. Results for *template-noexplanation-0.0*

Tables CVIII–CXIII show the results for 5 experiments with the *codellama-34b-instruct* model at temperature 0.0 using the prompt template of Figure 16 and using the system prompt shown in Figure 7.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,465	276	27	37	1,125	676	448	1	60.18
<i>countries-and-timezones</i>	106	307	84	2	10	211	183	28	0	86.73
<i>crawler-url-parser</i>	176	514	216	23	19	239	140	99	0	58.58
<i>delta</i>	462	1,375	592	29	20	734	598	110	26	85.01
<i>image-downloader</i>	42	126	41	4	2	77	62	15	0	80.52
<i>node-dirty</i>	154	458	160	30	10	258	146	99	13	61.63
<i>node-geo-point</i>	140	414	103	3	9	297	216	81	0	72.73
<i>node-jsonfile</i>	68	204	48	4	0	152	54	45	53	70.39
<i>plural</i>	153	449	110	50	16	273	198	74	1	72.89
<i>pull-stream</i>	351	1,037	236	16	11	774	440	278	56	64.08
<i>q</i>	1,051	3,143	1,106	122	59	1,856	138	1,635	83	11.91
<i>spacel-core</i>	134	395	143	17	5	211	175	35	1	83.41
<i>zip-a-folder</i>	49	143	41	3	1	98	27	3	68	96.94
<i>Total</i>	3,376	10,030	3,156	330	199	6,305	3,053	2,950	302	—

TABLE CVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #372). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,465	276	27	37	1,125	678	446	1	60.36
<i>countries-and-timezones</i>	106	307	84	2	10	211	183	28	0	86.73
<i>crawler-url-parser</i>	176	514	216	23	19	239	140	99	0	58.58
<i>delta</i>	462	1,375	595	28	19	733	598	108	27	85.27
<i>image-downloader</i>	42	126	41	4	2	77	62	15	0	80.52
<i>node-dirty</i>	154	458	160	30	10	258	146	99	13	61.63
<i>node-geo-point</i>	140	414	103	2	9	298	216	82	0	72.48
<i>node-jsonfile</i>	68	204	48	4	0	152	54	45	53	70.39
<i>plural</i>	153	449	110	50	16	273	198	74	1	72.89
<i>pull-stream</i>	351	1,037	236	16	11	774	439	279	56	63.95
<i>q</i>	1,051	3,143	1,107	122	60	1,854	138	1,633	83	11.92
<i>spacel-core</i>	134	395	143	18	6	209	175	33	1	84.21
<i>zip-a-folder</i>	49	143	41	3	1	98	26	3	69	96.94
<i>Total</i>	3,376	10,030	3,160	329	200	6,301	3,053	2,944	304	—

TABLE CIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #374). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,465	280	27	37	1,121	674	446	1	60.21
<i>countries-and-timezones</i>	106	307	84	2	10	211	183	28	0	86.73
<i>crawler-url-parser</i>	176	514	216	23	19	239	140	99	0	58.58
<i>delta</i>	462	1,375	595	28	19	733	597	110	26	84.99
<i>image-downloader</i>	42	126	41	4	2	77	62	15	0	80.52
<i>node-dirty</i>	154	458	160	30	10	258	146	99	13	61.63
<i>node-geo-point</i>	140	414	103	2	9	298	216	82	0	72.48
<i>node-jsonfile</i>	68	204	48	4	0	152	54	45	53	70.39
<i>plural</i>	153	449	110	50	16	273	198	74	1	72.89
<i>pull-stream</i>	351	1,037	236	16	11	774	439	279	56	63.95
<i>q</i>	1,051	3,143	1,107	122	59	1,855	138	1,634	83	11.91
<i>spacel-core</i>	134	396	144	18	6	209	175	33	1	84.21
<i>zip-a-folder</i>	49	143	41	3	1	98	26	3	69	96.94
<i>Total</i>	3,376	10,031	3,165	329	199	6,298	3,048	2,947	303	—

TABLE CX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #375). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,465	275	27	37	1,126	679	446	1	60.39
<i>countries-and-timezones</i>	106	307	84	2	10	211	183	28	0	86.73
<i>crawler-url-parser</i>	176	514	216	23	19	239	140	99	0	58.58
<i>delta</i>	462	1,375	598	25	19	733	597	109	27	85.13
<i>image-downloader</i>	42	126	41	4	2	77	62	15	0	80.52
<i>node-dirty</i>	154	458	160	30	10	258	146	99	13	61.63
<i>node-geo-point</i>	140	414	103	2	9	298	216	82	0	72.48
<i>node-jsonfile</i>	68	204	48	4	0	152	54	45	53	70.39
<i>plural</i>	153	449	110	50	16	273	198	74	1	72.89
<i>pull-stream</i>	351	1,037	239	16	9	773	440	277	56	64.17
<i>q</i>	1,051	3,143	1,108	121	60	1,854	136	1,634	84	11.87
<i>spacel-core</i>	134	396	144	18	6	209	177	31	1	85.17
<i>zip-a-folder</i>	49	143	41	3	1	98	27	3	68	96.94
<i>Total</i>	3,376	10,031	3,167	325	198	6,301	3,055	2,942	304	—

TABLE CXI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #376). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,465	276	27	37	1,125	678	446	1	60.36
<i>countries-and-timezones</i>	106	307	84	2	10	211	183	28	0	86.73
<i>crawler-url-parser</i>	176	514	213	26	19	239	140	99	0	58.58
<i>delta</i>	462	1,375	595	28	19	733	597	110	26	84.99
<i>image-downloader</i>	42	126	41	4	2	77	62	15	0	80.52
<i>node-dirty</i>	154	458	160	30	10	258	146	99	13	61.63
<i>node-geo-point</i>	140	414	106	2	9	295	213	82	0	72.20
<i>node-jsonfile</i>	68	204	48	4	0	152	54	45	53	70.39
<i>plural</i>	153	449	110	50	16	273	198	74	1	72.89
<i>pull-stream</i>	351	1,037	236	16	11	774	438	280	56	63.82
<i>q</i>	1,051	3,143	1,107	121	59	1,856	137	1,635	84	11.91
<i>spacel-core</i>	134	396	144	18	6	209	175	33	1	84.21
<i>zip-a-folder</i>	49	143	41	3	1	98	27	3	68	96.94
<i>Total</i>	3,376	10,031	3,161	331	199	6,300	3,048	2,949	303	—

TABLE CXII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #377). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	3,056.33	600.05	948,398	75,551	1,023,949
<i>countries-and-timezones</i>	1,070.68	309.71	101,694	23,742	125,436
<i>crawler-url-parser</i>	1,656.21	778.34	379,359	31,115	410,474
<i>delta</i>	2,870.28	3,625.25	872,234	64,880	937,114
<i>image-downloader</i>	430.47	329.74	23,017	9,110	32,127
<i>node-dirty</i>	1,526.57	243.81	240,242	24,279	264,521
<i>node-geo-point</i>	1,411.01	1,009.62	310,873	26,100	336,973
<i>node-jsonfile</i>	690.59	482.67	54,864	15,154	70,018
<i>plural</i>	1,522.48	146.41	259,635	26,465	286,100
<i>pull-stream</i>	2,632.74	1,388.06	194,441	73,821	268,262
<i>q</i>	4,694.78	12,908.57	2,086,666	127,647	2,214,313
<i>spacel-core</i>	1,350.87	711.26	157,479	28,201	185,680
<i>zip-a-folder</i>	500.54	1,097.23	80,546	10,243	90,789
<i>Total</i>	23,413.56	23,630.72	5,709,448	536,308	6,245,756

TABLE CXIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #372). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
<i>Complex.js</i>	3,088.22	620.07	948,398	75,464	1,023,862
<i>countries-and-timezones</i>	1,070.66	302.71	101,694	23,766	125,460
<i>crawler-url-parser</i>	1,653.80	799.42	379,359	31,089	410,448
<i>delta</i>	2,871.99	3,718.30	872,234	65,148	937,382
<i>image-downloader</i>	430.48	328.19	23,017	9,096	32,113
<i>node-dirty</i>	1,526.65	242.49	240,242	24,129	264,371
<i>node-geo-point</i>	1,410.97	985.29	310,873	26,143	337,016
<i>node-jsonfile</i>	690.54	481.86	54,864	15,125	69,989
<i>plural</i>	1,522.07	145.74	259,635	26,527	286,162
<i>pull-stream</i>	2,643.33	1,393.56	194,441	73,922	268,363
<i>q</i>	4,623.16	12,860.56	2,086,666	127,954	2,214,620
<i>spacl-core</i>	1,360.90	649.12	157,479	28,174	185,653
<i>zip-a-folder</i>	500.51	1,117.51	80,546	10,267	90,813
<i>Total</i>	23,393.28	23,644.84	5,709,448	536,804	6,246,252

TABLE CXIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #374). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
<i>Complex.js</i>	3,054.59	596.41	948,398	75,593	1,023,991
<i>countries-and-timezones</i>	1,070.75	306.48	101,694	23,740	125,434
<i>crawler-url-parser</i>	1,653.25	791.38	379,359	31,096	410,455
<i>delta</i>	2,869.41	3,656.14	872,234	64,872	937,106
<i>image-downloader</i>	430.47	328.44	23,017	9,109	32,126
<i>node-dirty</i>	1,526.59	236.98	240,242	24,096	264,338
<i>node-geo-point</i>	1,410.99	992.09	310,873	26,143	337,016
<i>node-jsonfile</i>	690.55	485.04	54,864	15,130	69,994
<i>plural</i>	1,521.62	145.62	259,635	26,482	286,117
<i>pull-stream</i>	2,631.99	1,391.93	194,441	73,754	268,195
<i>q</i>	4,695.78	12,866.72	2,086,666	127,918	2,214,584
<i>spacl-core</i>	1,350.86	706.03	157,479	28,159	185,638
<i>zip-a-folder</i>	500.51	1,109.02	80,546	10,227	90,773
<i>Total</i>	23,407.38	23,612.27	5,709,448	536,319	6,245,767

TABLE CXV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #375). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
<i>Complex.js</i>	3,053.84	617.90	948,398	75,377	1,023,775
<i>countries-and-timezones</i>	1,070.72	306.03	101,694	23,805	125,499
<i>crawler-url-parser</i>	1,656.36	838.63	379,359	31,102	410,461
<i>delta</i>	2,886.73	3,644.21	872,234	64,947	937,181
<i>image-downloader</i>	430.48	330.08	23,017	9,107	32,124
<i>node-dirty</i>	1,526.82	243.25	240,242	24,153	264,395
<i>node-geo-point</i>	1,410.93	999.44	310,873	26,143	337,016
<i>node-jsonfile</i>	690.54	480.47	54,864	15,130	69,994
<i>plural</i>	1,522.37	144.72	259,635	26,473	286,108
<i>pull-stream</i>	2,649.32	1,395.27	194,441	73,826	268,267
<i>q</i>	4,627.96	12,851.25	2,086,666	127,807	2,214,473
<i>spacl-core</i>	1,350.90	680.52	157,479	28,203	185,682
<i>zip-a-folder</i>	500.52	1,098.96	80,546	10,244	90,790
<i>Total</i>	23,377.51	23,630.74	5,709,448	536,317	6,245,765

TABLE CXVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #376). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
<i>Complex.js</i>	3,048.61	622.31	948,398	75,411	1,023,809
<i>countries-and-timezones</i>	1,070.74	302.32	101,694	23,740	125,434
<i>crawler-url-parser</i>	1,653.97	835.85	379,359	30,947	410,306
<i>delta</i>	2,880.16	3,648.84	872,234	65,086	937,320
<i>image-downloader</i>	430.49	326.72	23,017	9,110	32,127
<i>node-dirty</i>	1,526.71	230.04	240,242	24,142	264,384
<i>node-geo-point</i>	1,410.98	963.40	310,873	26,313	337,186
<i>node-jsonfile</i>	690.67	530.41	54,864	15,130	69,994
<i>plural</i>	1,522.04	148.93	259,635	26,465	286,100
<i>pull-stream</i>	2,630.34	1,397.17	194,441	73,763	268,204
<i>q</i>	4,627.86	12,869.01	2,086,666	127,790	2,214,456
<i>spacl-core</i>	1,350.91	714.31	157,479	28,174	185,653
<i>zip-a-folder</i>	500.51	1,073.84	80,546	10,244	90,790
<i>Total</i>	23,344.00	23,663.14	5,709,448	536,315	6,245,763

TABLE CXVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #377). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noexplanation.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

```

Your task is to apply mutation testing to the following code:
'''
{{{code}}}
'''

by replacing the PLACEHOLDER with a buggy code fragment that has different
behavior than the original code fragment, which was:
'''
{{{orig}}}
'''

Please consider changes such as using different operators, changing constants,
referring to different variables, object properties, functions, or methods.

Provide three answers as fenced code blocks containing a single line of code,
using the following template:

Option 1: The PLACEHOLDER can be replaced with:
'''
<code fragment>
'''

Option 2: The PLACEHOLDER can be replaced with:
'''
<code fragment>
'''

Option 3: The PLACEHOLDER can be replaced with:
'''
<code fragment>
'''

Please conclude your response with "DONE."

```

Fig. 16. Variation on the template of Figure 7 that does not request explanations for the suggested mutations.

L. Results for *template-noinstructions-0.0*

Tables CXVIII–CXXVII show the results for 5 experiments with the *codellama-34b-instruct* model at temperature 0.0 using the prompt template of Figure 17 and using the system prompt shown in Figure 7.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,453	235	36	45	1,137	696	440	1	61.30
<i>countries-and-timezones</i>	106	315	85	8	4	218	174	44	0	79.82
<i>crawler-url-parser</i>	176	518	211	23	20	246	134	112	0	54.47
<i>delta</i>	462	1,366	563	26	18	759	612	115	32	84.85
<i>image-downloader</i>	42	126	39	2	0	84	69	15	0	82.14
<i>node-dirty</i>	154	457	162	26	9	260	146	103	11	60.38
<i>node-geo-point</i>	140	413	85	7	13	306	230	76	0	75.16
<i>node-jsonfile</i>	68	200	42	9	1	148	45	51	52	65.54
<i>plural</i>	153	444	109	57	17	261	189	71	1	72.80
<i>pull-stream</i>	351	1,040	224	19	16	781	467	248	66	68.25
<i>q</i>	1,051	3,130	1,019	94	59	1,958	138	1,726	94	11.85
<i>spacel-core</i>	134	395	158	30	7	187	155	31	1	83.42
<i>zip-a-folder</i>	49	144	41	5	1	97	26	4	67	95.88
<i>Total</i>	3,376	10,001	2,973	342	210	6,442	3,081	3,036	325	—

TABLE CXVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #378). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,453	235	36	45	1,137	696	440	1	61.30
<i>countries-and-timezones</i>	106	315	85	8	4	218	174	44	0	79.82
<i>crawler-url-parser</i>	176	518	211	23	20	246	134	112	0	54.47
<i>delta</i>	462	1,366	563	26	18	759	610	117	32	84.58
<i>image-downloader</i>	42	126	39	2	0	84	70	14	0	83.33
<i>node-dirty</i>	154	457	162	26	9	260	146	103	11	60.38
<i>node-geo-point</i>	140	413	85	7	13	306	230	76	0	75.16
<i>node-jsonfile</i>	68	200	42	9	1	148	45	51	52	65.54
<i>plural</i>	153	444	109	57	17	261	189	71	1	72.80
<i>pull-stream</i>	351	1,040	224	19	16	781	467	248	66	68.25
<i>q</i>	1,051	3,130	1,017	93	60	1,960	137	1,728	95	11.84
<i>spacel-core</i>	134	395	157	30	7	188	156	31	1	83.51
<i>zip-a-folder</i>	49	144	41	5	1	97	26	4	67	95.88
<i>Total</i>	3,376	10,001	2,970	341	211	6,445	3,080	3,039	326	—

TABLE CXIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #379). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,453	235	36	45	1,137	696	440	1	61.30
<i>countries-and-timezones</i>	106	315	85	8	4	218	174	44	0	79.82
<i>crawler-url-parser</i>	176	518	211	23	20	246	134	112	0	54.47
<i>delta</i>	462	1,365	562	26	18	759	611	116	32	84.72
<i>image-downloader</i>	42	126	39	2	0	84	69	15	0	82.14
<i>node-dirty</i>	154	457	162	26	9	260	146	103	11	60.38
<i>node-geo-point</i>	140	412	84	7	13	306	230	76	0	75.16
<i>node-jsonfile</i>	68	200	42	9	1	148	45	51	52	65.54
<i>plural</i>	153	444	109	57	17	261	189	71	1	72.80
<i>pull-stream</i>	351	1,040	224	19	16	781	466	249	66	68.12
<i>q</i>	1,051	3,132	1,021	94	60	1,957	137	1,727	93	11.75
<i>spacel-core</i>	134	395	158	30	7	187	155	31	1	83.42
<i>zip-a-folder</i>	49	144	41	5	1	97	26	4	67	95.88
<i>Total</i>	3,376	10,001	2,973	342	211	6,441	3,078	3,039	324	—

TABLE CXX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #380). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,452	234	36	46	1,136	695	440	1	61.27
<i>countries-and-timezones</i>	106	315	85	8	4	218	174	44	0	79.82
<i>crawler-url-parser</i>	176	518	211	23	20	246	134	112	0	54.47
<i>delta</i>	462	1,366	563	26	18	759	612	115	32	84.85
<i>image-downloader</i>	42	126	39	2	0	84	70	14	0	83.33
<i>node-dirty</i>	154	457	162	26	9	260	147	102	11	60.77
<i>node-geo-point</i>	140	413	85	7	13	306	230	76	0	75.16
<i>node-jsonfile</i>	68	200	42	9	1	148	45	51	52	65.54
<i>plural</i>	153	445	109	57	17	262	189	72	1	72.52
<i>pull-stream</i>	351	1,040	226	19	16	779	464	249	66	68.04
<i>q</i>	1,051	3,131	1,019	96	60	1,956	138	1,725	93	11.81
<i>spacel-core</i>	134	395	158	30	7	187	155	31	1	83.42
<i>zip-a-folder</i>	49	144	41	5	1	97	26	4	67	95.88
<i>Total</i>	3,376	10,002	2,974	344	212	6,438	3,079	3,035	324	—

TABLE CXXI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #381). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,453	235	36	45	1,137	696	440	1	61.30
<i>countries-and-timezones</i>	106	315	85	8	4	218	174	44	0	79.82
<i>crawler-url-parser</i>	176	518	211	23	20	246	134	112	0	54.47
<i>delta</i>	462	1,366	563	26	18	759	612	115	32	84.85
<i>image-downloader</i>	42	126	39	2	0	84	69	15	0	82.14
<i>node-dirty</i>	154	457	159	26	10	262	149	102	11	61.07
<i>node-geo-point</i>	140	413	85	7	13	306	230	76	0	75.16
<i>node-jsonfile</i>	68	200	42	9	1	148	45	51	52	65.54
<i>plural</i>	153	443	108	57	17	261	189	71	1	72.80
<i>pull-stream</i>	351	1,040	226	19	16	779	465	248	66	68.16
<i>q</i>	1,051	3,130	1,019	93	60	1,958	137	1,728	93	11.75
<i>spacel-core</i>	134	395	158	30	7	187	155	31	1	83.42
<i>zip-a-folder</i>	49	144	41	5	1	97	26	4	67	95.88
<i>Total</i>	3,376	10,000	2,971	341	212	6,442	3,081	3,037	324	—

TABLE CXXII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #382). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	3,363.13	610.09	953,788	104,944	1,058,732
<i>countries-and-timezones</i>	1,070.70	333.60	102,860	23,506	126,366
<i>crawler-url-parser</i>	1,668.12	865.68	381,295	38,817	420,112
<i>delta</i>	3,242.91	4,060.50	877,316	99,522	976,838
<i>image-downloader</i>	430.47	361.71	23,479	8,905	32,384
<i>node-dirty</i>	1,530.58	228.43	241,936	33,033	274,969
<i>node-geo-point</i>	1,410.89	1,019.95	312,413	28,975	341,388
<i>node-jsonfile</i>	690.55	469.01	55,612	14,598	70,210
<i>plural</i>	1,523.05	141.16	261,318	34,491	295,809
<i>pull-stream</i>	2,608.43	1,433.25	198,302	74,144	272,446
<i>q</i>	5,802.92	13,526.38	2,098,227	218,277	2,316,504
<i>spacel-core</i>	1,350.79	626.46	158,953	29,519	188,472
<i>zip-a-folder</i>	500.51	1,087.03	81,085	10,694	91,779
<i>Total</i>	25,193.04	24,763.25	5,746,584	719,425	6,466,009

TABLE CXXIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #378). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,324.82	606.13	953,788	104,909	1,058,697
<i>countries-and-timezones</i>	1,070.79	314.67	102,860	23,502	126,362
<i>crawler-url-parser</i>	1,673.87	789.35	381,295	38,816	420,111
<i>delta</i>	3,186.56	3,835.89	877,316	99,463	976,779
<i>image-downloader</i>	430.46	362.97	23,479	8,961	32,440
<i>node-dirty</i>	1,530.82	229.51	241,936	33,036	274,972
<i>node-geo-point</i>	1,410.92	1,035.06	312,413	28,975	341,388
<i>node-jsonfile</i>	691.16	471.51	55,612	14,598	70,210
<i>plural</i>	1,523.31	142.93	261,318	34,484	295,802
<i>pull-stream</i>	2,610.16	1,428.74	198,302	74,195	272,497
<i>q</i>	5,806.23	13,570.79	2,098,227	218,057	2,316,284
<i>spacl-core</i>	1,350.78	598.84	158,953	29,457	188,410
<i>zip-a-folder</i>	500.50	1,065.03	81,085	10,694	91,779
<i>Total</i>	25,110.38	24,451.42	5,746,584	719,147	6,465,731

TABLE CXXIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #379). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,340.83	597.34	953,788	105,056	1,058,844
<i>countries-and-timezones</i>	1,070.75	323.54	102,860	23,502	126,362
<i>crawler-url-parser</i>	1,659.88	803.01	381,295	38,801	420,096
<i>delta</i>	3,211.07	3,830.91	877,316	99,521	976,837
<i>image-downloader</i>	430.47	362.10	23,479	8,905	32,384
<i>node-dirty</i>	1,530.67	228.98	241,936	33,033	274,969
<i>node-geo-point</i>	1,410.92	997.90	312,413	29,053	341,466
<i>node-jsonfile</i>	690.56	466.32	55,612	14,598	70,210
<i>plural</i>	1,522.82	141.80	261,318	34,484	295,802
<i>pull-stream</i>	2,609.79	1,444.02	198,302	74,220	272,522
<i>q</i>	6,945.19	13,543.90	2,098,227	218,309	2,316,536
<i>spacl-core</i>	1,350.87	605.57	158,953	29,527	188,480
<i>zip-a-folder</i>	500.50	1,075.42	81,085	10,694	91,779
<i>Total</i>	26,274.31	24,420.80	5,746,584	719,703	6,466,287

TABLE CXXV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #380). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
<i>Complex.js</i>	3,492.66	629.00	953,788	104,886	1,058,674
<i>countries-and-timezones</i>	1,177.99	323.34	102,860	23,502	126,362
<i>crawler-url-parser</i>	1,751.37	786.06	381,295	38,800	420,095
<i>delta</i>	3,365.20	3,872.79	877,316	99,562	976,878
<i>image-downloader</i>	441.26	358.96	23,479	8,961	32,440
<i>node-dirty</i>	1,608.97	228.53	241,936	33,053	274,989
<i>node-geo-point</i>	1,492.54	1,035.93	312,413	28,984	341,397
<i>node-jsonfile</i>	717.49	465.56	55,612	14,548	70,160
<i>plural</i>	1,630.97	140.39	261,318	34,444	295,762
<i>pull-stream</i>	2,738.37	1,433.35	198,302	74,222	272,524
<i>q</i>	6,155.10	13,521.99	2,098,227	218,163	2,316,390
<i>spacl-core</i>	1,440.32	585.59	158,953	29,512	188,465
<i>zip-a-folder</i>	509.75	1,063.09	81,085	10,694	91,779
<i>Total</i>	26,522.01	24,444.59	5,746,584	719,331	6,465,915

TABLE CXXVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #381). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		prompt	#tokens compl.	total
	LLMorpheus	StrykerJS			
<i>Complex.js</i>	3,378.96	601.89	953,788	104,940	1,058,728
<i>countries-and-timezones</i>	1,080.75	325.94	102,860	23,502	126,362
<i>crawler-url-parser</i>	1,660.11	741.99	381,295	38,801	420,096
<i>delta</i>	3,233.82	3,810.62	877,316	99,525	976,841
<i>image-downloader</i>	430.46	361.68	23,479	8,905	32,384
<i>node-dirty</i>	1,531.35	234.89	241,936	33,044	274,980
<i>node-geo-point</i>	1,410.90	1,014.28	312,413	28,969	341,382
<i>node-jsonfile</i>	690.54	466.32	55,612	14,598	70,210
<i>plural</i>	1,523.06	136.58	261,318	34,492	295,810
<i>pull-stream</i>	2,606.22	1,436.07	198,302	74,135	272,437
<i>q</i>	5,921.20	13,597.41	2,098,227	218,214	2,316,441
<i>spacl-core</i>	1,350.81	585.65	158,953	29,520	188,473
<i>zip-a-folder</i>	500.48	1,078.08	81,085	10,745	91,830
<i>Total</i>	25,318.67	24,391.40	5,746,584	719,390	6,465,974

TABLE CXXVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #382). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-noinstructions.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

Your task is to apply mutation testing to the following code:

```
'''
{{{code}}}
'''
```

by replacing the PLACEHOLDER with a buggy code fragment that has different behavior than the original code fragment, which was:

```
'''
{{{orig}}}
'''
```

Provide three answers as fenced code blocks containing a single line of code, using the following template:

Option 1: The PLACEHOLDER can be replaced with:

```
'''
<code fragment>
'''
```

This would result in different behavior because <brieif explanation>.

Option 2: The PLACEHOLDER can be replaced with:

```
'''
<code fragment>
'''
```

This would result in different behavior because <brieif explanation>.

Option 3: The PLACEHOLDER can be replaced with:

```
'''
<code fragment>
'''
```

This would result in different behavior because <brieif explanation>.

Please conclude your response with "DONE."

Fig. 17. Variation on the template of Figure 7 that does not provide instructions on how to create mutants.

M. Results for *template-full-genericssystemprompt-0.0*

Tables CXXVIII–CXXXVII show the results for 5 experiments with the *codellama-34b-instruct* model at temperature 0.0 using the prompt template of Figure 7 and the generic system prompt of Figure 18.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,459	201	9	50	1,199	740	458	1	61.80
<i>countries-and-timezones</i>	106	316	90	0	9	217	191	26	0	88.02
<i>crawler-url-parser</i>	176	522	226	16	19	246	143	103	0	58.13
<i>delta</i>	462	1,373	553	7	23	790	659	99	32	87.47
<i>image-downloader</i>	42	123	34	1	0	88	72	16	0	81.82
<i>node-dirty</i>	154	450	153	12	8	277	162	104	11	62.45
<i>node-geo-point</i>	140	414	95	1	11	305	229	76	0	75.08
<i>node-jsonfile</i>	68	199	47	2	0	150	49	49	52	67.33
<i>plural</i>	153	442	104	45	21	272	209	62	1	77.21
<i>pull-stream</i>	351	1,039	252	15	9	763	442	266	55	65.14
<i>q</i>	1,051	3,128	1,037	30	54	2,007	145	1,770	92	11.81
<i>spacel-core</i>	134	397	146	12	8	214	181	32	1	85.05
<i>zip-a-folder</i>	49	145	40	1	2	101	27	3	71	97.03
<i>Total</i>	3,376	10,007	2,978	151	214	6,629	3,249	3,064	316	—

TABLE CXXVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #384). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,459	201	9	50	1,199	740	458	1	61.80
<i>countries-and-timezones</i>	106	316	90	0	9	217	191	26	0	88.02
<i>crawler-url-parser</i>	176	522	226	16	19	246	143	103	0	58.13
<i>delta</i>	462	1,372	555	6	23	788	657	99	32	87.44
<i>image-downloader</i>	42	123	34	1	0	88	72	16	0	81.82
<i>node-dirty</i>	154	450	153	12	8	277	162	104	11	62.45
<i>node-geo-point</i>	140	414	95	1	11	305	229	76	0	75.08
<i>node-jsonfile</i>	68	200	47	2	0	151	49	49	53	67.55
<i>plural</i>	153	443	105	45	21	272	209	62	1	77.21
<i>pull-stream</i>	351	1,040	252	15	9	764	443	266	55	65.18
<i>q</i>	1,051	3,127	1,035	30	54	2,008	145	1,771	92	11.80
<i>spacel-core</i>	134	397	146	12	8	214	181	32	1	85.05
<i>zip-a-folder</i>	49	145	40	1	2	101	27	3	71	97.03
<i>Total</i>	3,376	10,008	2,979	150	214	6,630	3,248	3,065	317	—

TABLE CXXIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #385). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,459	201	9	50	1,199	740	458	1	61.80
<i>countries-and-timezones</i>	106	316	90	0	9	217	191	26	0	88.02
<i>crawler-url-parser</i>	176	522	226	16	19	246	143	103	0	58.13
<i>delta</i>	462	1,373	551	6	24	792	660	100	32	87.37
<i>image-downloader</i>	42	123	34	1	0	88	72	16	0	81.82
<i>node-dirty</i>	154	450	153	12	8	277	162	104	11	62.45
<i>node-geo-point</i>	140	415	96	1	11	305	229	76	0	75.08
<i>node-jsonfile</i>	68	200	47	2	0	151	49	49	53	67.55
<i>plural</i>	153	442	106	45	21	270	207	62	1	77.04
<i>pull-stream</i>	351	1,039	252	15	9	763	440	268	55	64.88
<i>q</i>	1,051	3,127	1,035	30	55	2,007	144	1,771	92	11.76
<i>spacel-core</i>	134	396	146	12	8	216	183	32	1	85.19
<i>zip-a-folder</i>	49	145	40	1	2	101	27	3	71	97.03
<i>Total</i>	3,376	10,007	2,977	150	216	6,632	3,247	3,068	317	—

TABLE CXXX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #386). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,459	201	9	49	1,200	741	458	1	61.83
<i>countries-and-timezones</i>	106	316	90	0	9	217	191	26	0	88.02
<i>crawler-url-parser</i>	176	522	226	16	19	246	144	102	0	58.54
<i>delta</i>	462	1,372	551	6	24	791	659	100	32	87.36
<i>image-downloader</i>	42	123	34	1	0	88	72	16	0	81.82
<i>node-dirty</i>	154	450	153	12	8	277	163	103	11	62.82
<i>node-geo-point</i>	140	414	95	1	11	305	229	76	0	75.08
<i>node-jsonfile</i>	68	200	47	2	0	151	49	49	53	67.55
<i>plural</i>	153	442	106	45	21	270	207	62	1	77.04
<i>pull-stream</i>	351	1,039	252	15	9	763	442	266	55	65.14
<i>q</i>	1,051	3,125	1,036	29	54	2,006	144	1,770	92	11.76
<i>spacl-core</i>	134	397	146	12	8	214	181	32	1	85.05
<i>zip-a-folder</i>	49	145	40	1	2	101	27	3	71	97.03
<i>Total</i>	3,376	10,004	2,977	149	214	6,629	3,249	3,063	317	—

TABLE CXXXI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #387). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	1,459	201	9	50	1,199	740	458	1	61.80
<i>countries-and-timezones</i>	106	292	81	0	9	202	179	23	0	88.61
<i>crawler-url-parser</i>	176	522	226	16	19	246	143	103	0	58.13
<i>delta</i>	462	1,373	551	6	24	792	660	100	32	87.37
<i>image-downloader</i>	42	123	34	1	0	88	72	16	0	81.82
<i>node-dirty</i>	154	450	153	12	8	277	162	104	11	62.45
<i>node-geo-point</i>	140	414	95	1	11	305	229	76	0	75.08
<i>node-jsonfile</i>	68	200	47	2	0	151	49	49	53	67.55
<i>plural</i>	153	442	104	45	21	272	209	62	1	77.21
<i>pull-stream</i>	351	1,039	252	15	9	763	442	266	55	65.14
<i>q</i>	1,051	3,127	1,039	30	53	2,005	146	1,768	91	11.82
<i>spacl-core</i>	134	397	145	12	8	215	182	32	1	85.12
<i>zip-a-folder</i>	49	145	40	1	2	101	27	3	71	97.03
<i>Total</i>	3,376	9,983	2,968	150	214	6,616	3,240	3,060	316	—

TABLE CXXXII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #388). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*.

project	time (sec)		#tokens		
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	total
<i>Complex.js</i>	3,198.91	631.03	943,498	97,397	1,040,895
<i>countries-and-timezones</i>	1,070.69	306.50	100,634	22,822	123,456
<i>crawler-url-parser</i>	1,666.72	773.61	377,599	38,968	416,567
<i>delta</i>	3,188.39	3,973.51	867,614	96,702	964,316
<i>image-downloader</i>	430.51	376.36	22,597	8,748	31,345
<i>node-dirty</i>	1,531.02	243.07	238,702	32,642	271,344
<i>node-geo-point</i>	1,410.94	993.80	309,473	28,703	338,176
<i>node-jsonfile</i>	690.59	470.19	54,184	13,966	68,150
<i>plural</i>	1,523.29	146.19	258,105	33,232	291,337
<i>pull-stream</i>	2,621.42	1,374.37	190,931	73,130	264,061
<i>q</i>	5,911.30	13,892.03	2,076,156	216,002	2,292,158
<i>spacl-core</i>	1,350.90	688.46	156,139	28,052	184,191
<i>zip-a-folder</i>	500.50	1,139.30	80,056	10,370	90,426
<i>Total</i>	25,095.20	25,008.45	5,675,688	700,734	6,376,422

TABLE CXXXIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #384). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*

You are a programming assistant. You are expected to be concise and precise and avoid any unnecessary examples, tests, and verbosity.

Fig. 18. Generic system prompt.

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,198.28	628.16	943,498	97,360	1,040,858
countries-and-timezones	1,070.75	313.88	100,634	22,817	123,451
crawler-url-parser	1,667.00	785.35	377,599	38,968	416,567
delta	3,193.53	4,061.32	867,614	96,672	964,286
image-downloader	430.50	373.49	22,597	8,748	31,345
node-dirty	1,530.73	243.51	238,702	32,641	271,343
node-geo-point	1,411.00	1,029.33	309,473	28,703	338,176
node-jsonfile	690.57	473.38	54,184	13,976	68,160
plural	1,523.28	142.88	258,105	33,183	291,288
pull-stream	2,622.56	1,368.80	190,931	73,002	263,933
q	5,761.33	13,943.75	2,076,156	216,075	2,292,231
spacl-core	1,350.88	668.94	156,139	28,048	184,187
zip-a-folder	500.52	1,170.72	80,056	10,370	90,426
Total	24,950.92	25,203.51	5,675,688	700,563	6,376,251

TABLE CXXXIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #385). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,246.58	637.01	943,498	97,351	1,040,849
countries-and-timezones	1,070.74	317.35	100,634	22,822	123,456
crawler-url-parser	1,666.54	828.48	377,599	38,968	416,567
delta	3,141.81	3,972.35	867,614	96,648	964,262
image-downloader	430.53	375.91	22,597	8,740	31,337
node-dirty	1,531.64	237.14	238,702	32,632	271,334
node-geo-point	1,410.80	1,007.52	309,473	28,670	338,143
node-jsonfile	690.60	479.57	54,184	13,982	68,166
plural	1,523.32	141.62	258,105	33,221	291,326
pull-stream	2,590.59	1,371.20	190,931	73,097	264,028
q	5,912.17	13,970.49	2,076,156	216,015	2,292,171
spacl-core	1,461.05	690.80	156,139	28,074	184,213
zip-a-folder	500.51	1,128.17	80,056	10,370	90,426
Total	25,176.88	25,157.62	5,675,688	700,590	6,376,278

TABLE CXXXV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #386). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*

project	time (sec)		prompt	#tokens	total
	LLMorpheus	StrykerJS		compl.	
Complex.js	3,229.79	627.53	943,498	97,331	1,040,829
countries-and-timezones	1,070.66	315.55	100,634	22,813	123,447
crawler-url-parser	1,677.85	830.94	377,599	39,015	416,614
delta	3,135.61	3,928.08	867,614	96,647	964,261
image-downloader	430.47	374.42	22,597	8,748	31,345
node-dirty	1,530.49	236.58	238,702	32,642	271,344
node-geo-point	1,410.83	1,012.70	309,473	28,703	338,176
node-jsonfile	690.54	473.16	54,184	13,999	68,183
plural	1,523.30	142.61	258,105	33,221	291,326
pull-stream	2,590.76	1,370.63	190,931	73,109	264,040
q	5,913.75	13,931.58	2,076,156	216,172	2,292,328
spacl-core	1,350.92	677.55	156,139	28,048	184,187
zip-a-folder	500.55	1,132.09	80,056	10,370	90,426
Total	25,055.51	25,053.41	5,675,688	700,818	6,376,506

TABLE CXXXVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #387). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	3,200.91	614.84	943,498	97,375	1,040,873
<i>countries-and-timezones</i>	1,501.26	297.33	95,530	21,102	116,632
<i>crawler-url-parser</i>	1,664.69	819.62	377,599	38,970	416,569
<i>delta</i>	3,159.72	4,099.01	867,614	96,624	964,238
<i>image-downloader</i>	430.49	373.69	22,597	8,748	31,345
<i>node-dirty</i>	1,555.91	239.75	238,702	32,632	271,334
<i>node-geo-point</i>	1,410.92	994.48	309,473	28,709	338,182
<i>node-jsonfile</i>	690.59	467.78	54,184	13,996	68,180
<i>plural</i>	1,523.33	146.45	258,105	33,232	291,337
<i>pull-stream</i>	2,596.08	1,354.67	190,931	73,098	264,029
<i>q</i>	5,912.23	13,931.75	2,076,156	216,129	2,292,285
<i>spacel-core</i>	1,350.85	689.90	156,139	28,037	184,176
<i>zip-a-folder</i>	530.56	1,147.59	80,056	10,370	90,426
<i>Total</i>	25,527.53	25,176.85	5,670,584	699,022	6,369,606

TABLE CXXXVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #388). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-full.hb*, SYSTEMPROMPT: *SystemPrompt-Generic.txt*

N. Results for *template-basic-0.0*

Tables CXXXVIII–CXLVII show the results for 5 experiments with the *codellama-34b-instruct* model at temperature 0.0 using the prompt template of Figure 19 and using the system prompt shown in Figure 7.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	490	212	92	1	185	120	65	0	64.86
<i>countries-and-timezones</i>	106	106	36	22	0	48	44	4	0	91.67
<i>crawler-url-parser</i>	176	176	75	27	1	67	49	18	0	73.13
<i>delta</i>	462	462	201	54	6	201	167	28	6	86.07
<i>image-downloader</i>	42	42	21	11	0	10	7	3	0	70.00
<i>node-dirty</i>	154	154	70	38	2	44	24	18	2	59.09
<i>node-geo-point</i>	140	140	39	33	6	62	54	8	0	87.10
<i>node-jsonfile</i>	68	68	17	28	1	22	11	3	8	86.36
<i>plural</i>	153	152	35	21	4	92	78	14	0	84.78
<i>pull-stream</i>	351	351	115	87	0	149	88	54	7	63.76
<i>q</i>	1,051	1,051	405	232	13	401	38	350	13	12.72
<i>spacel-core</i>	134	134	65	37	1	25	23	2	0	92.00
<i>zip-a-folder</i>	49	49	18	11	0	20	5	1	14	95.00
<i>Total</i>	3,376	3,375	1,309	693	35	1,326	708	568	50	—

TABLE CXXXVIII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #390). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	490	213	92	1	184	119	65	0	64.67
<i>countries-and-timezones</i>	106	106	36	22	0	48	43	5	0	89.58
<i>crawler-url-parser</i>	176	176	75	27	1	67	49	18	0	73.13
<i>delta</i>	462	462	200	54	6	202	168	28	6	86.14
<i>image-downloader</i>	42	42	21	11	0	10	7	3	0	70.00
<i>node-dirty</i>	154	154	70	38	2	44	24	18	2	59.09
<i>node-geo-point</i>	140	140	39	33	6	62	54	8	0	87.10
<i>node-jsonfile</i>	68	68	17	28	1	22	11	3	8	86.36
<i>plural</i>	153	152	35	21	4	92	78	14	0	84.78
<i>pull-stream</i>	351	351	115	87	0	149	88	54	7	63.76
<i>q</i>	1,051	1,051	405	232	13	401	38	350	13	12.72
<i>spacel-core</i>	134	134	65	37	1	25	23	2	0	92.00
<i>zip-a-folder</i>	49	49	18	11	0	20	5	1	14	95.00
<i>Total</i>	3,376	3,375	1,309	693	35	1,326	707	569	50	—

TABLE CXXXIX

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #391). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	490	213	92	1	184	119	65	0	64.67
<i>countries-and-timezones</i>	106	106	36	22	0	48	43	5	0	89.58
<i>crawler-url-parser</i>	176	176	75	27	1	67	49	18	0	73.13
<i>delta</i>	462	462	201	54	6	201	167	28	6	86.07
<i>image-downloader</i>	42	42	21	11	0	10	7	3	0	70.00
<i>node-dirty</i>	154	154	71	38	2	43	24	17	2	60.47
<i>node-geo-point</i>	140	140	39	33	6	62	54	8	0	87.10
<i>node-jsonfile</i>	68	68	17	28	1	22	11	3	8	86.36
<i>plural</i>	153	152	35	21	4	92	78	14	0	84.78
<i>pull-stream</i>	351	351	115	87	0	149	88	54	7	63.76
<i>q</i>	1,051	1,051	405	231	13	402	38	351	13	12.69
<i>spacel-core</i>	134	134	65	37	1	25	23	2	0	92.00
<i>zip-a-folder</i>	49	49	18	11	0	20	5	1	14	95.00
<i>Total</i>	3,376	3,375	1,311	692	35	1,325	706	569	50	—

TABLE CXL

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #392). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	490	213	92	1	184	119	65	0	64.67
<i>countries-and-timezones</i>	106	106	36	22	0	48	43	5	0	89.58
<i>crawler-url-parser</i>	176	176	75	27	1	67	49	18	0	73.13
<i>delta</i>	462	462	201	54	6	201	167	28	6	86.07
<i>image-downloader</i>	42	42	21	11	0	10	7	3	0	70.00
<i>node-dirty</i>	154	154	70	38	2	44	24	18	2	59.09
<i>node-geo-point</i>	140	140	39	33	6	62	54	8	0	87.10
<i>node-jsonfile</i>	68	68	17	28	1	22	11	3	8	86.36
<i>plural</i>	153	150	35	20	4	91	78	13	0	85.71
<i>pull-stream</i>	351	351	115	87	0	149	88	54	7	63.76
<i>q</i>	1,051	1,051	405	231	13	402	38	351	13	12.69
<i>spacel-core</i>	134	134	65	37	1	25	23	2	0	92.00
<i>zip-a-folder</i>	49	49	18	11	0	20	5	1	14	95.00
<i>Total</i>	3,376	3,373	1,310	691	35	1,325	706	569	50	—

TABLE CXLI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #393). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

application	#prompts	mutant candidates				#mutants	#killed	#survived	#timeout	mut. score
		total	invalid	identical	duplicate					
<i>Complex.js</i>	490	490	212	92	1	185	120	65	0	64.86
<i>countries-and-timezones</i>	106	106	36	22	0	48	44	4	0	91.67
<i>crawler-url-parser</i>	176	176	75	27	1	67	49	18	0	73.13
<i>delta</i>	462	462	202	54	6	200	166	28	6	86.00
<i>image-downloader</i>	42	42	21	11	0	10	7	3	0	70.00
<i>node-dirty</i>	154	154	71	38	2	43	24	17	2	60.47
<i>node-geo-point</i>	140	140	39	33	6	62	54	8	0	87.10
<i>node-jsonfile</i>	68	68	17	28	1	22	11	3	8	86.36
<i>plural</i>	153	152	35	21	4	92	78	14	0	84.78
<i>pull-stream</i>	351	351	115	87	0	149	88	54	7	63.76
<i>q</i>	1,051	1,051	404	232	13	402	38	351	13	12.69
<i>spacel-core</i>	134	134	65	37	1	25	23	2	0	92.00
<i>zip-a-folder</i>	49	49	18	11	0	20	5	1	14	95.00
<i>Total</i>	3,376	3,375	1,310	693	35	1,325	707	568	50	—

TABLE CXLII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #394). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*.

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	2,731.54	97.72	893,966	14,460	908,426
<i>countries-and-timezones</i>	1,071.22	73.08	89,939	3,087	93,026
<i>crawler-url-parser</i>	1,636.67	215.52	359,498	5,557	365,055
<i>delta</i>	2,659.93	910.49	820,541	13,472	834,013
<i>image-downloader</i>	430.67	65.98	18,348	1,448	19,796
<i>node-dirty</i>	1,526.59	39.53	223,071	4,425	227,496
<i>node-geo-point</i>	1,411.43	204.76	295,321	4,217	299,538
<i>node-jsonfile</i>	690.87	77.82	47,346	1,831	49,177
<i>plural</i>	1,521.52	48.57	241,953	5,075	247,028
<i>pull-stream</i>	2,382.28	245.31	156,016	9,287	165,303
<i>q</i>	4,158.17	2,697.98	1,970,359	30,059	2,000,418
<i>spacel-core</i>	1,351.47	85.42	142,466	4,013	146,479
<i>zip-a-folder</i>	500.75	235.31	75,033	1,594	76,627
<i>Total</i>	22,073.11	4,997.50	5,333,857	98,525	5,432,382

TABLE CXLIH

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #390). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		#tokens		
	LLMorpheus	StrykerJS	prompt	compl.	total
<i>Complex.js</i>	2,732.79	102.55	893,966	14,472	908,438
<i>countries-and-timezones</i>	1,071.27	72.88	89,939	3,113	93,052
<i>crawler-url-parser</i>	1,636.61	211.77	359,498	5,557	365,055
<i>delta</i>	2,667.43	897.69	820,541	13,458	833,999
<i>image-downloader</i>	430.64	65.54	18,348	1,448	19,796
<i>node-dirty</i>	1,526.57	38.93	223,071	4,422	227,493
<i>node-geo-point</i>	1,411.51	203.69	295,321	4,218	299,539
<i>node-jsonfile</i>	690.88	77.42	47,346	1,831	49,177
<i>plural</i>	1,521.47	47.92	241,953	5,075	247,028
<i>pull-stream</i>	2,382.21	245.17	156,016	9,288	165,304
<i>q</i>	4,159.06	2,700.88	1,970,359	30,059	2,000,418
<i>spacl-core</i>	1,351.31	85.36	142,466	4,007	146,473
<i>zip-a-folder</i>	500.73	227.68	75,033	1,594	76,627
<i>Total</i>	22,082.48	4,977.48	5,333,857	98,542	5,432,399

TABLE CXLIV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #391). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		#tokens		
	LLMorpheus	StrykerJS	prompt	compl.	total
<i>Complex.js</i>	2,730.84	97.59	893,966	14,461	908,427
<i>countries-and-timezones</i>	1,071.16	72.83	89,939	3,113	93,052
<i>crawler-url-parser</i>	1,636.68	220.83	359,498	5,576	365,074
<i>delta</i>	2,659.86	887.50	820,541	13,473	834,014
<i>image-downloader</i>	430.65	65.72	18,348	1,449	19,797
<i>node-dirty</i>	1,526.53	37.30	223,071	4,496	227,567
<i>node-geo-point</i>	1,411.50	195.42	295,321	4,230	299,551
<i>node-jsonfile</i>	690.89	77.90	47,346	1,831	49,177
<i>plural</i>	1,521.54	49.08	241,953	5,075	247,028
<i>pull-stream</i>	2,382.22	248.75	156,016	9,288	165,304
<i>q</i>	4,158.01	2,694.11	1,970,359	30,070	2,000,429
<i>spacl-core</i>	1,351.43	85.92	142,466	4,013	146,479
<i>zip-a-folder</i>	500.71	229.75	75,033	1,594	76,627
<i>Total</i>	22,072.01	4,962.71	5,333,857	98,669	5,432,526

TABLE CXLV

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #392). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		#tokens		
	LLMorpheus	StrykerJS	prompt	compl.	total
<i>Complex.js</i>	2,730.22	96.64	893,966	14,459	908,425
<i>countries-and-timezones</i>	1,071.13	74.16	89,939	3,112	93,051
<i>crawler-url-parser</i>	1,636.64	206.03	359,498	5,556	365,054
<i>delta</i>	2,659.88	897.06	820,541	13,471	834,012
<i>image-downloader</i>	430.66	65.63	18,348	1,449	19,797
<i>node-dirty</i>	1,526.57	38.82	223,071	4,425	227,496
<i>node-geo-point</i>	1,411.45	200.23	295,321	4,217	299,538
<i>node-jsonfile</i>	690.84	77.61	47,346	1,831	49,177
<i>plural</i>	1,556.64	47.70	238,779	5,029	243,808
<i>pull-stream</i>	2,382.19	247.28	156,016	9,288	165,304
<i>q</i>	4,156.62	2,695.05	1,970,359	30,071	2,000,430
<i>spacl-core</i>	1,351.45	84.84	142,466	4,008	146,474
<i>zip-a-folder</i>	500.75	235.51	75,033	1,594	76,627
<i>Total</i>	22,105.05	4,966.57	5,330,683	98,510	5,429,193

TABLE CXLVI

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #393). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

project	time (sec)		#tokens		total
	<i>LLMorpheus</i>	<i>StrykerJS</i>	prompt	compl.	
<i>Complex.js</i>	2,730.17	96.67	893,966	14,460	908,426
<i>countries-and-timezones</i>	1,071.19	73.83	89,939	3,086	93,025
<i>crawler-url-parser</i>	1,636.65	205.98	359,498	5,577	365,075
<i>delta</i>	2,659.91	887.25	820,541	13,475	834,016
<i>image-downloader</i>	430.71	65.63	18,348	1,449	19,797
<i>node-dirty</i>	1,526.60	38.47	223,071	4,500	227,571
<i>node-geo-point</i>	1,411.44	197.24	295,321	4,217	299,538
<i>node-jsonfile</i>	690.88	78.16	47,346	1,831	49,177
<i>plural</i>	1,522.37	47.69	241,953	5,075	247,028
<i>pull-stream</i>	2,382.20	245.09	156,016	9,290	165,306
<i>q</i>	4,154.13	2,691.66	1,970,359	30,055	2,000,414
<i>spacl-core</i>	1,351.41	83.75	142,466	4,013	146,479
<i>zip-a-folder</i>	500.72	232.15	75,033	1,594	76,627
<i>Total</i>	22,068.39	4,943.55	5,333,857	98,622	5,432,479

TABLE CXLVII

RESULTS FROM LLMORPHEUS EXPERIMENT (RUN #394). MODEL: *codellama-34b-instruct*, TEMPERATURE: 0.0, MAXTOKENS: 250, TEMPLATE: *template-basic.hb*, SYSTEMPROMPT: *SystemPrompt-MutationTestingExpert.txt*

Consider the following code fragment:

```
'''
{{{code}}}
'''
```

Please provide a code fragment that PLACEHOLDER can be replaced with.

Provide your answer as a fenced code block containing a single line of code, using the following template:

The PLACEHOLDER can be replaced with:

```
'''
<code fragment>
'''
```

Please conclude your response with "DONE."

Fig. 19. A minimal template for requesting a replacement for the PLACEHOLDER.

O. Summary of variability of results for all models examined in the paper

For each configuration of *LLMorpheus*, this section provides a summary of the mutation results, showing the average of each value across all five trials, along with the standard deviation for that value.

- Model codellama-13b-instruct with template full at temperature 0.0: Table CL
- Model codellama-34b-instruct with template basic at temperature 0.0: Table CXLIX
- Model codellama-34b-instruct with template full at temperature 0.0: Table ??
- Model codellama-34b-instruct with template full at temperature 0.25: Table CLI
- Model codellama-34b-instruct with template full at temperature 0.5: Table CLII
- Model codellama-34b-instruct with template full at temperature 1.0: Table CLIII
- Model codellama-34b-instruct with template full at temperature genericsystemprompt-0.0: Table CLIV
- Model codellama-34b-instruct with template noexplanation at temperature 0.0: Table CLV
- Model codellama-34b-instruct with template noinstructions at temperature 0.0: Table CLVI
- Model codellama-34b-instruct with template onemutation at temperature 0.0: Table CLVII
- Model gpt-4o-mini with template full at temperature 0.0: Table CLVIII
- Model llama-3.3-70b-instruct with template full at temperature 0.0: Table CLIX
- Model mixtral-8x7b-instruct with template full at temperature 0.0: Table CLX

TABLE CXLVIII

SUMMARY OF MUTANTS FOR CODELLAMA-13B-INSTRUCT-FULL-0.0 (RUNS 354, 355, 356, 358, 359). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,411 $\pm$ 0	340 $\pm$ 0	116 $\pm$ 0	28 $\pm$ 0	955 $\pm$ 0	553 $\pm$ 0	401 $\pm$ 0	1 $\pm$ 0	58.01 $\pm$ 0
<i>countries-and-timezones</i>	106 $\pm$ 0	305 $\pm$ 0	83 $\pm$ 0	15 $\pm$ 0	1 $\pm$ 0	207 $\pm$ 0	177 $\pm$ 0	30 $\pm$ 0	0 $\pm$ 0	85.51 $\pm$ 0
<i>crawler-url-parser</i>	176 $\pm$ 0	494 $\pm$ 0	186 $\pm$ 0	51 $\pm$ 0	12 $\pm$ 0	257 $\pm$ 0	128 $\pm$ 2	118 $\pm$ 0	1 $\pm$ 2	52.23 $\pm$ 0
<i>delta</i>	462 $\pm$ 0	1,334 $\pm$ 0	530 $\pm$ 0	92 $\pm$ 0	16 $\pm$ 0	712 $\pm$ 0	583 $\pm$ 0	107 $\pm$ 0	22 $\pm$ 0	84.97 $\pm$ 0
<i>image-downloader</i>	42 $\pm$ 0	122 $\pm$ 0	40 $\pm$ 0	5 $\pm$ 0	2 $\pm$ 0	77 $\pm$ 0	48 $\pm$ 0	29 $\pm$ 0	0 $\pm$ 0	62.34 $\pm$ 0
<i>node-dirty</i>	154 $\pm$ 0	439 $\pm$ 0	161 $\pm$ 0	33 $\pm$ 0	11 $\pm$ 0	245 $\pm$ 0	142 $\pm$ 0	92 $\pm$ 0	11 $\pm$ 0	62.45 $\pm$ 0
<i>node-geo-point</i>	140 $\pm$ 0	390 $\pm$ 0	64 $\pm$ 0	21 $\pm$ 0	16 $\pm$ 0	305 $\pm$ 0	237 $\pm$ 0	67 $\pm$ 0	0 $\pm$ 0	77.96 $\pm$ 0
<i>node-jsonfile</i>	68 $\pm$ 0	192 $\pm$ 0	45 $\pm$ 1	10 $\pm$ 0	6 $\pm$ 0	137 $\pm$ 0	43 $\pm$ 0	45 $\pm$ 0	49 $\pm$ 0	67.2 $\pm$ 0.11
<i>plural</i>	153 $\pm$ 0	404 $\pm$ 6	100 $\pm$ 1	98 $\pm$ 1	17 $\pm$ 0	206 $\pm$ 4	153 $\pm$ 3	53 $\pm$ 1	1 $\pm$ 0	74.52 $\pm$ 0.01
<i>pull-stream</i>	351 $\pm$ 0	1,002 $\pm$ 0	279 $\pm$ 0	54 $\pm$ 0	13 $\pm$ 0	669 $\pm$ 0	386 $\pm$ 1	236 $\pm$ 1	46 $\pm$ 0	64.65 $\pm$ 0.08
<i>q</i>	1,051 $\pm$ 0	2,992 $\pm$ 1	899 $\pm$ 1	379 $\pm$ 0	55 $\pm$ 0	1,714 $\pm$ 1	122 $\pm$ 0	1,519 $\pm$ 1	73 $\pm$ 0	11.38 $\pm$ 0
<i>spacel-core</i>	134 $\pm$ 0	377 $\pm$ 0	142 $\pm$ 0	40 $\pm$ 0	7 $\pm$ 0	195 $\pm$ 0	160 $\pm$ 0	25 $\pm$ 0	0 $\pm$ 0	86.38 $\pm$ 0.24
<i>zip-a-folder</i>	49 $\pm$ 0	137 $\pm$ 0	43 $\pm$ 0	7 $\pm$ 0	1 $\pm$ 0	87 $\pm$ 0	27 $\pm$ 0	55 $\pm$ 0	5 $\pm$ 0	36.78 $\pm$ 0
<i>Total</i>	3,376	9,599	2,911	921	185	5,766	2,759	2,777	209	63.41

TABLE CXLIX

SUMMARY OF MUTANTS FOR CODELLAMA-34B-INSTRUCT-BASIC-0.0 (RUNS 390, 391, 392, 393, 394). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	489 $\pm$ 0	213 $\pm$ 1	92 $\pm$ 0	1 $\pm$ 0	184 $\pm$ 1	119 $\pm$ 1	65 $\pm$ 0	0 $\pm$ 0	64.75 $\pm$ 0.1
<i>countries-and-timezones</i>	106 $\pm$ 0	106 $\pm$ 0	36 $\pm$ 0	22 $\pm$ 0	0 $\pm$ 0	48 $\pm$ 0	43 $\pm$ 1	5 $\pm$ 1	0 $\pm$ 0	90.42 $\pm$ 1.14
<i>crawler-url-parser</i>	176 $\pm$ 0	175 $\pm$ 0	75 $\pm$ 0	27 $\pm$ 0	1 $\pm$ 0	73 $\pm$ 0	49 $\pm$ 0	18 $\pm$ 0	0 $\pm$ 0	73.13 $\pm$ 0
<i>delta</i>	462 $\pm$ 0	456 $\pm$ 0	201 $\pm$ 1	54 $\pm$ 0	6 $\pm$ 0	201 $\pm$ 1	167 $\pm$ 1	28 $\pm$ 0	6 $\pm$ 0	86.07 $\pm$ 0.05
<i>image-downloader</i>	42 $\pm$ 0	42 $\pm$ 0	21 $\pm$ 0	11 $\pm$ 0	0 $\pm$ 0	10 $\pm$ 0	7 $\pm$ 0	3 $\pm$ 0	0 $\pm$ 0	70 $\pm$ 0
<i>node-dirty</i>	154 $\pm$ 0	152 $\pm$ 0	70 $\pm$ 1	38 $\pm$ 0	2 $\pm$ 0	44 $\pm$ 1	24 $\pm$ 0	18 $\pm$ 1	2 $\pm$ 0	59.64 $\pm$ 0.76
<i>node-geo-point</i>	140 $\pm$ 0	134 $\pm$ 0	39 $\pm$ 0	33 $\pm$ 0	6 $\pm$ 0	62 $\pm$ 0	54 $\pm$ 0	8 $\pm$ 0	0 $\pm$ 0	87.1 $\pm$ 0
<i>node-jsonfile</i>	68 $\pm$ 0	67 $\pm$ 0	17 $\pm$ 0	28 $\pm$ 0	1 $\pm$ 0	22 $\pm$ 0	11 $\pm$ 0	3 $\pm$ 0	8 $\pm$ 0	86.36 $\pm$ 0
<i>plural</i>	153 $\pm$ 0	148 $\pm$ 1	35 $\pm$ 0	21 $\pm$ 0	4 $\pm$ 0	92 $\pm$ 0	78 $\pm$ 0	14 $\pm$ 0	0 $\pm$ 0	84.97 $\pm$ 0.42
<i>pull-stream</i>	351 $\pm$ 0	351 $\pm$ 0	115 $\pm$ 0	87 $\pm$ 0	0 $\pm$ 0	149 $\pm$ 0	88 $\pm$ 0	54 $\pm$ 0	7 $\pm$ 0	63.76 $\pm$ 0
<i>q</i>	1,051 $\pm$ 0	1,038 $\pm$ 0	405 $\pm$ 0	232 $\pm$ 1	13 $\pm$ 0	402 $\pm$ 1	38 $\pm$ 0	351 $\pm$ 1	13 $\pm$ 0	12.7 $\pm$ 0.02
<i>spacl-core</i>	134 $\pm$ 0	133 $\pm$ 0	65 $\pm$ 0	37 $\pm$ 0	1 $\pm$ 0	31 $\pm$ 0	23 $\pm$ 0	2 $\pm$ 0	0 $\pm$ 0	92 $\pm$ 0
<i>zip-a-folder</i>	49 $\pm$ 0	49 $\pm$ 0	18 $\pm$ 0	11 $\pm$ 0	0 $\pm$ 0	20 $\pm$ 0	5 $\pm$ 0	1 $\pm$ 0	14 $\pm$ 0	95 $\pm$ 0
<i>Total</i>	3,376	3,340	1,310	692	35	1,337	707	569	50	74.3

TABLE CL

SUMMARY OF MUTANTS FOR CODELLAMA-13B-INSTRUCT-FULL-0.0 (RUNS 354, 355, 356, 358, 359). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,411 $\pm$ 0	340 $\pm$ 0	116 $\pm$ 0	28 $\pm$ 0	955 $\pm$ 0	553 $\pm$ 0	401 $\pm$ 0	1 $\pm$ 0	58.01 $\pm$ 0
<i>countries-and-timezones</i>	106 $\pm$ 0	305 $\pm$ 0	83 $\pm$ 0	15 $\pm$ 0	1 $\pm$ 0	207 $\pm$ 0	177 $\pm$ 0	30 $\pm$ 0	0 $\pm$ 0	85.51 $\pm$ 0
<i>crawler-url-parser</i>	176 $\pm$ 0	494 $\pm$ 0	186 $\pm$ 0	51 $\pm$ 0	12 $\pm$ 0	257 $\pm$ 0	128 $\pm$ 2	118 $\pm$ 0	1 $\pm$ 2	52.23 $\pm$ 0
<i>delta</i>	462 $\pm$ 0	1,334 $\pm$ 0	530 $\pm$ 0	92 $\pm$ 0	16 $\pm$ 0	712 $\pm$ 0	583 $\pm$ 0	107 $\pm$ 0	22 $\pm$ 0	84.97 $\pm$ 0
<i>image-downloader</i>	42 $\pm$ 0	122 $\pm$ 0	40 $\pm$ 0	5 $\pm$ 0	2 $\pm$ 0	77 $\pm$ 0	48 $\pm$ 0	29 $\pm$ 0	0 $\pm$ 0	62.34 $\pm$ 0
<i>node-dirty</i>	154 $\pm$ 0	439 $\pm$ 0	161 $\pm$ 0	33 $\pm$ 0	11 $\pm$ 0	245 $\pm$ 0	142 $\pm$ 0	92 $\pm$ 0	11 $\pm$ 0	62.45 $\pm$ 0
<i>node-geo-point</i>	140 $\pm$ 0	390 $\pm$ 0	64 $\pm$ 0	21 $\pm$ 0	16 $\pm$ 0	305 $\pm$ 0	237 $\pm$ 0	67 $\pm$ 0	0 $\pm$ 0	77.96 $\pm$ 0
<i>node-jsonfile</i>	68 $\pm$ 0	192 $\pm$ 0	45 $\pm$ 1	10 $\pm$ 0	6 $\pm$ 0	137 $\pm$ 0	43 $\pm$ 0	45 $\pm$ 0	49 $\pm$ 0	67.2 $\pm$ 0.11
<i>plural</i>	153 $\pm$ 0	404 $\pm$ 6	100 $\pm$ 1	98 $\pm$ 1	17 $\pm$ 0	206 $\pm$ 4	153 $\pm$ 3	53 $\pm$ 1	1 $\pm$ 0	74.52 $\pm$ 0.01
<i>pull-stream</i>	351 $\pm$ 0	1,002 $\pm$ 0	279 $\pm$ 0	54 $\pm$ 0	13 $\pm$ 0	669 $\pm$ 0	386 $\pm$ 1	236 $\pm$ 1	46 $\pm$ 0	64.65 $\pm$ 0.08
<i>q</i>	1,051 $\pm$ 0	2,992 $\pm$ 1	899 $\pm$ 1	379 $\pm$ 0	55 $\pm$ 0	1,714 $\pm$ 1	122 $\pm$ 0	1,519 $\pm$ 1	73 $\pm$ 0	11.38 $\pm$ 0
<i>spacl-core</i>	134 $\pm$ 0	377 $\pm$ 0	142 $\pm$ 0	40 $\pm$ 0	7 $\pm$ 0	195 $\pm$ 0	160 $\pm$ 0	25 $\pm$ 0	0 $\pm$ 0	86.38 $\pm$ 0.24
<i>zip-a-folder</i>	49 $\pm$ 0	137 $\pm$ 0	43 $\pm$ 0	7 $\pm$ 0	1 $\pm$ 0	87 $\pm$ 0	27 $\pm$ 0	55 $\pm$ 0	5 $\pm$ 0	36.78 $\pm$ 0
<i>Total</i>	3,376	9,599	2,911	921	185	5,766	2,759	2,777	209	63.41

TABLE CLI

SUMMARY OF MUTANTS FOR CODELLAMA-34B-INSTRUCT-FULL-0.25 (RUNS 348, 350, 351, 352, 353). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,413 $\pm$ 7	212 $\pm$ 12	10 $\pm$ 2	43 $\pm$ 4	1,190 $\pm$ 13	727 $\pm$ 8	463 $\pm$ 9	0 $\pm$ 1	61.09 $\pm$ 0.5
<i>countries-and-timezones</i>	106 $\pm$ 0	308 $\pm$ 5	85 $\pm$ 3	1 $\pm$ 1	7 $\pm$ 4	222 $\pm$ 5	193 $\pm$ 8	28 $\pm$ 5	0 $\pm$ 0	87.18 $\pm$ 2.55
<i>crawler-url-parser</i>	176 $\pm$ 0	501 $\pm$ 4	210 $\pm$ 5	13 $\pm$ 4	15 $\pm$ 4	278 $\pm$ 8	157 $\pm$ 9	104 $\pm$ 6	0 $\pm$ 0	60.25 $\pm$ 2.47
<i>delta</i>	462 $\pm$ 0	1,341 $\pm$ 6	566 $\pm$ 13	8 $\pm$ 2	25 $\pm$ 3	767 $\pm$ 12	627 $\pm$ 17	107 $\pm$ 4	32 $\pm$ 4	86.01 $\pm$ 0.66
<i>image-downloader</i>	42 $\pm$ 0	123 $\pm$ 2	36 $\pm$ 2	1 $\pm$ 2	1 $\pm$ 1	86 $\pm$ 2	68 $\pm$ 4	17 $\pm$ 2	0 $\pm$ 0	80.43 $\pm$ 3.22
<i>node-dirty</i>	154 $\pm$ 0	445 $\pm$ 5	160 $\pm$ 4	14 $\pm$ 3	10 $\pm$ 4	271 $\pm$ 6	165 $\pm$ 6	94 $\pm$ 5	12 $\pm$ 0	65.19 $\pm$ 1.82
<i>node-geo-point</i>	140 $\pm$ 0	394 $\pm$ 12	90 $\pm$ 10	1 $\pm$ 1	12 $\pm$ 4	303 $\pm$ 8	226 $\pm$ 10	74 $\pm$ 6	0 $\pm$ 0	75.36 $\pm$ 2.01
<i>node-jsonfile</i>	68 $\pm$ 0	201 $\pm$ 2	44 $\pm$ 3	3 $\pm$ 1	1 $\pm$ 1	154 $\pm$ 4	52 $\pm$ 3	47 $\pm$ 5	56 $\pm$ 2	69.72 $\pm$ 2.82
<i>plural</i>	153 $\pm$ 0	419 $\pm$ 2	101 $\pm$ 2	36 $\pm$ 5	20 $\pm$ 2	282 $\pm$ 7	210 $\pm$ 4	71 $\pm$ 7	1 $\pm$ 1	75.03 $\pm$ 2.12
<i>pull-stream</i>	351 $\pm$ 0	1,023 $\pm$ 1	239 $\pm$ 3	9 $\pm$ 2	7 $\pm$ 2	774 $\pm$ 3	447 $\pm$ 4	274 $\pm$ 6	54 $\pm$ 3	64.67 $\pm$ 0.8
<i>q</i>	1,051 $\pm$ 0	3,073 $\pm$ 8	1,023 $\pm$ 14	27 $\pm$ 11	51 $\pm$ 8	2,023 $\pm$ 31	147 $\pm$ 5	1,795 $\pm$ 27	82 $\pm$ 4	11.3 $\pm$ 0.33
<i>spacl-core</i>	134 $\pm$ 0	389 $\pm$ 3	139 $\pm$ 5	11 $\pm$ 2	6 $\pm$ 3	240 $\pm$ 6	189 $\pm$ 4	33 $\pm$ 3	1 $\pm$ 1	85.05 $\pm$ 1.23
<i>zip-a-folder</i>	49 $\pm$ 0	140 $\pm$ 2	39 $\pm$ 4	1 $\pm$ 1	1 $\pm$ 0	101 $\pm$ 4	32 $\pm$ 10	22 $\pm$ 27	46 $\pm$ 36	77.41 $\pm$ 26.94
<i>Total</i>	3,376	9,771	2,943	136	200	6,692	3,242	3,129	283	69.13

TABLE CLII

SUMMARY OF MUTANTS FOR CODELLAMA-34B-INSTRUCT-FULL-0.5 (RUNS 318, 319, 320, 321, 322). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,417 $\pm$ 7	226 $\pm$ 16	8 $\pm$ 1	37 $\pm$ 3	1,183 $\pm$ 10	725 $\pm$ 10	458 $\pm$ 10	0 $\pm$ 0	61.27 $\pm$ 0.72
<i>countries-and-timezones</i>	106 $\pm$ 0	305 $\pm$ 6	79 $\pm$ 6	4 $\pm$ 2	7 $\pm$ 3	222 $\pm$ 2	195 $\pm$ 3	28 $\pm$ 3	0 $\pm$ 0	87.59 $\pm$ 1.26
<i>crawler-url-parser</i>	176 $\pm$ 0	500 $\pm$ 3	193 $\pm$ 12	13 $\pm$ 4	15 $\pm$ 2	294 $\pm$ 12	167 $\pm$ 6	121 $\pm$ 11	5 $\pm$ 11	58.65 $\pm$ 3.1
<i>delta</i>	462 $\pm$ 0	1,340 $\pm$ 5	553 $\pm$ 12	10 $\pm$ 2	22 $\pm$ 5	777 $\pm$ 14	645 $\pm$ 19	97 $\pm$ 7	36 $\pm$ 2	87.56 $\pm$ 1.12
<i>image-downloader</i>	42 $\pm$ 0	122 $\pm$ 2	36 $\pm$ 3	2 $\pm$ 1	3 $\pm$ 2	84 $\pm$ 3	64 $\pm$ 6	21 $\pm$ 6	0 $\pm$ 0	75.57 $\pm$ 7.56
<i>node-dirty</i>	154 $\pm$ 0	446 $\pm$ 3	158 $\pm$ 7	9 $\pm$ 2	9 $\pm$ 4	279 $\pm$ 7	162 $\pm$ 8	105 $\pm$ 3	12 $\pm$ 0	62.23 $\pm$ 1.48
<i>node-geo-point</i>	140 $\pm$ 0	390 $\pm$ 18	89 $\pm$ 6	2 $\pm$ 1	13 $\pm$ 3	299 $\pm$ 17	228 $\pm$ 10	72 $\pm$ 10	0 $\pm$ 0	76.16 $\pm$ 2.27
<i>node-jsonfile</i>	68 $\pm$ 0	200 $\pm$ 2	42 $\pm$ 3	3 $\pm$ 1	1 $\pm$ 2	155 $\pm$ 5	53 $\pm$ 3	40 $\pm$ 3	62 $\pm$ 4	74.43 $\pm$ 2.05
<i>plural</i>	153 $\pm$ 0	415 $\pm$ 7	98 $\pm$ 7	27 $\pm$ 4	19 $\pm$ 5	290 $\pm$ 4	220 $\pm$ 9	70 $\pm$ 8	1 $\pm$ 1	75.89 $\pm$ 2.87
<i>pull-stream</i>	351 $\pm$ 0	1,023 $\pm$ 7	219 $\pm$ 10	6 $\pm$ 3	10 $\pm$ 4	798 $\pm$ 12	472 $\pm$ 11	270 $\pm$ 9	56 $\pm$ 4	66.17 $\pm$ 1.28
<i>q</i>	1,051 $\pm$ 0	3,071 $\pm$ 6	1,000 $\pm$ 17	29 $\pm$ 5	52 $\pm$ 8	2,042 $\pm$ 22	153 $\pm$ 11	1,804 $\pm$ 27	85 $\pm$ 8	11.67 $\pm$ 0.71
<i>spac-core</i>	134 $\pm$ 0	388 $\pm$ 4	125 $\pm$ 7	7 $\pm$ 3	5 $\pm$ 2	256 $\pm$ 5	217 $\pm$ 6	39 $\pm$ 3	1 $\pm$ 0	84.93 $\pm$ 1.01
<i>zip-a-folder</i>	49 $\pm$ 0	138 $\pm$ 9	35 $\pm$ 5	1 $\pm$ 1	1 $\pm$ 1	101 $\pm$ 10	36 $\pm$ 10	19 $\pm$ 21	46 $\pm$ 37	79.77 $\pm$ 22.09
<i>Total</i>	3,376	9,755	2,853	121	194	6,782	3,336	3,143	303	69.38

TABLE CLIII

SUMMARY OF MUTANTS FOR CODELLAMA-34B-INSTRUCT-FULL-1.0 (RUNS 341, 342, 343, 345, 347). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,359 $\pm$ 13	338 $\pm$ 20	6 $\pm$ 2	24 $\pm$ 2	1,015 $\pm$ 22	641 $\pm$ 20	374 $\pm$ 6	0 $\pm$ 1	63.12 $\pm$ 0.83
<i>countries-and-timezones</i>	106 $\pm$ 0	284 $\pm$ 20	91 $\pm$ 11	2 $\pm$ 3	3 $\pm$ 1	191 $\pm$ 17	168 $\pm$ 16	24 $\pm$ 6	0 $\pm$ 0	87.55 $\pm$ 2.79
<i>crawler-url-parser</i>	176 $\pm$ 0	489 $\pm$ 7	218 $\pm$ 13	5 $\pm$ 1	6 $\pm$ 2	266 $\pm$ 12	186 $\pm$ 14	80 $\pm$ 9	0 $\pm$ 0	69.79 $\pm$ 3.43
<i>delta</i>	462 $\pm$ 0	1,274 $\pm$ 28	556 $\pm$ 41	8 $\pm$ 3	19 $\pm$ 4	710 $\pm$ 25	599 $\pm$ 28	79 $\pm$ 6	32 $\pm$ 1	88.79 $\pm$ 1.09
<i>image-downloader</i>	42 $\pm$ 0	116 $\pm$ 2	42 $\pm$ 6	1 $\pm$ 1	2 $\pm$ 3	72 $\pm$ 8	55 $\pm$ 9	17 $\pm$ 7	0 $\pm$ 1	76.21 $\pm$ 9.17
<i>node-dirty</i>	154 $\pm$ 0	427 $\pm$ 10	164 $\pm$ 15	5 $\pm$ 3	5 $\pm$ 3	257 $\pm$ 11	150 $\pm$ 11	96 $\pm$ 7	11 $\pm$ 3	62.68 $\pm$ 2.25
<i>node-geo-point</i>	140 $\pm$ 0	392 $\pm$ 3	125 $\pm$ 5	3 $\pm$ 1	9 $\pm$ 2	265 $\pm$ 6	204 $\pm$ 8	61 $\pm$ 5	0 $\pm$ 0	76.95 $\pm$ 1.81
<i>node-jsonfile</i>	68 $\pm$ 0	190 $\pm$ 4	45 $\pm$ 4	2 $\pm$ 1	3 $\pm$ 2	142 $\pm$ 7	60 $\pm$ 8	20 $\pm$ 5	62 $\pm$ 4	85.71 $\pm$ 3.84
<i>plural</i>	153 $\pm$ 0	420 $\pm$ 12	110 $\pm$ 9	9 $\pm$ 2	13 $\pm$ 2	302 $\pm$ 5	239 $\pm$ 7	61 $\pm$ 5	1 $\pm$ 1	79.78 $\pm$ 1.62
<i>pull-stream</i>	351 $\pm$ 0	999 $\pm$ 6	245 $\pm$ 7	10 $\pm$ 2	6 $\pm$ 3	743 $\pm$ 5	455 $\pm$ 7	239 $\pm$ 5	49 $\pm$ 6	67.82 $\pm$ 0.77
<i>q</i>	1,051 $\pm$ 0	2,993 $\pm$ 11	1,078 $\pm$ 30	18 $\pm$ 3	35 $\pm$ 8	1,896 $\pm$ 28	145 $\pm$ 16	1,668 $\pm$ 23	83 $\pm$ 7	12.02 $\pm$ 0.47
<i>spac-core</i>	134 $\pm$ 0	372 $\pm$ 3	146 $\pm$ 12	4 $\pm$ 3	5 $\pm$ 1	222 $\pm$ 13	194 $\pm$ 19	28 $\pm$ 8	0 $\pm$ 0	87.36 $\pm$ 4.02
<i>zip-a-folder</i>	49 $\pm$ 0	138 $\pm$ 5	43 $\pm$ 2	1 $\pm$ 1	1 $\pm$ 1	94 $\pm$ 7	34 $\pm$ 12	10 $\pm$ 16	50 $\pm$ 27	89.46 $\pm$ 16.3
<i>Total</i>	3,376	9,454	3,201	73	131	6,179	3,130	2,759	290	72.86

TABLE CLIV

SUMMARY OF MUTANTS FOR CODELLAMA-34B-INSTRUCT-FULL-GENERICSYSTEMPROMPT-0.0 (RUNS 384, 385, 386, 387, 388). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,409 $\pm$ 0	201 $\pm$ 0	9 $\pm$ 0	50 $\pm$ 0	1,199 $\pm$ 0	740 $\pm$ 0	458 $\pm$ 0	1 $\pm$ 0	61.81 $\pm$ 0.01
<i>countries-and-timezones</i>	106 $\pm$ 0	302 $\pm$ 11	88 $\pm$ 4	0 $\pm$ 0	9 $\pm$ 0	214 $\pm$ 7	189 $\pm$ 5	25 $\pm$ 1	0 $\pm$ 0	88.14 $\pm$ 0.26
<i>crawler-url-parser</i>	176 $\pm$ 0	503 $\pm$ 0	226 $\pm$ 0	16 $\pm$ 0	19 $\pm$ 0	261 $\pm$ 0	143 $\pm$ 0	103 $\pm$ 0	0 $\pm$ 0	58.21 $\pm$ 0.18
<i>delta</i>	462 $\pm$ 0	1,349 $\pm$ 1	552 $\pm$ 2	6 $\pm$ 0	24 $\pm$ 1	791 $\pm$ 2	659 $\pm$ 1	100 $\pm$ 1	32 $\pm$ 0	87.4 $\pm$ 0.05
<i>image-downloader</i>	42 $\pm$ 0	123 $\pm$ 0	34 $\pm$ 0	1 $\pm$ 0	0 $\pm$ 0	88 $\pm$ 0	72 $\pm$ 0	16 $\pm$ 0	0 $\pm$ 0	81.82 $\pm$ 0
<i>node-dirty</i>	154 $\pm$ 0	442 $\pm$ 0	153 $\pm$ 0	12 $\pm$ 0	8 $\pm$ 0	277 $\pm$ 0	162 $\pm$ 0	104 $\pm$ 0	11 $\pm$ 0	62.52 $\pm$ 0.17
<i>node-geo-point</i>	140 $\pm$ 0	403 $\pm$ 0	95 $\pm$ 0	1 $\pm$ 0	11 $\pm$ 0	307 $\pm$ 0	229 $\pm$ 0	76 $\pm$ 0	0 $\pm$ 0	75.08 $\pm$ 0
<i>node-jsonfile</i>	68 $\pm$ 0	200 $\pm$ 0	47 $\pm$ 0	2 $\pm$ 0	0 $\pm$ 0	151 $\pm$ 0	49 $\pm$ 0	49 $\pm$ 0	53 $\pm$ 0	67.51 $\pm$ 0.1
<i>plural</i>	153 $\pm$ 0	421 $\pm$ 0	105 $\pm$ 1	45 $\pm$ 0	21 $\pm$ 0	271 $\pm$ 1	208 $\pm$ 1	62 $\pm$ 0	1 $\pm$ 0	77.14 $\pm$ 0.09
<i>pull-stream</i>	351 $\pm$ 0	1,030 $\pm$ 0	252 $\pm$ 0	15 $\pm$ 0	9 $\pm$ 0	763 $\pm$ 0	442 $\pm$ 1	266 $\pm$ 1	55 $\pm$ 0	65.1 $\pm$ 0.12
<i>q</i>	1,051 $\pm$ 0	3,073 $\pm$ 1	1,036 $\pm$ 2	30 $\pm$ 0	54 $\pm$ 1	2,007 $\pm$ 1	145 $\pm$ 1	1,770 $\pm$ 1	92 $\pm$ 0	11.79 $\pm$ 0.03
<i>spac-core</i>	134 $\pm$ 0	389 $\pm$ 0	146 $\pm$ 0	12 $\pm$ 0	8 $\pm$ 0	231 $\pm$ 1	182 $\pm$ 1	32 $\pm$ 0	1 $\pm$ 0	85.09 $\pm$ 0.06
<i>zip-a-folder</i>	49 $\pm$ 0	143 $\pm$ 0	40 $\pm$ 0	1 $\pm$ 0	2 $\pm$ 0	102 $\pm$ 0	27 $\pm$ 0	3 $\pm$ 0	71 $\pm$ 0	97.03 $\pm$ 0
<i>Total</i>	3,376	9,787	2,976	150	214	6,662	3,247	3,064	317	70.66

TABLE CLV

SUMMARY OF MUTANTS FOR CODELLAMA-34B-INSTRUCT-NOEXPLANATION-0.0 (RUNS 372, 374, 375, 376, 377). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,428 $\pm$ 0	277 $\pm$ 2	27 $\pm$ 0	37 $\pm$ 0	1,124 $\pm$ 2	677 $\pm$ 2	446 $\pm$ 1	1 $\pm$ 0	60.3 $\pm$ 0.1
<i>countries-and-timezones</i>	106 $\pm$ 0	297 $\pm$ 0	84 $\pm$ 0	2 $\pm$ 0	10 $\pm$ 0	211 $\pm$ 0	183 $\pm$ 0	28 $\pm$ 0	0 $\pm$ 0	86.73 $\pm$ 0
<i>crawler-url-parser</i>	176 $\pm$ 0	495 $\pm$ 0	215 $\pm$ 1	24 $\pm$ 1	19 $\pm$ 0	256 $\pm$ 0	140 $\pm$ 0	99 $\pm$ 0	0 $\pm$ 0	58.58 $\pm$ 0
<i>delta</i>	462 $\pm$ 0	1,356 $\pm$ 0	595 $\pm$ 2	28 $\pm$ 2	19 $\pm$ 0	733 $\pm$ 0	597 $\pm$ 1	109 $\pm$ 1	26 $\pm$ 1	85.08 $\pm$ 0.12
<i>image-downloader</i>	42 $\pm$ 0	124 $\pm$ 0	41 $\pm$ 0	4 $\pm$ 0	2 $\pm$ 0	79 $\pm$ 0	62 $\pm$ 0	15 $\pm$ 0	0 $\pm$ 0	80.52 $\pm$ 0
<i>node-dirty</i>	154 $\pm$ 0	448 $\pm$ 0	160 $\pm$ 0	30 $\pm$ 0	10 $\pm$ 0	258 $\pm$ 0	146 $\pm$ 0	99 $\pm$ 0	13 $\pm$ 0	61.63 $\pm$ 0
<i>node-geo-point</i>	140 $\pm$ 0	405 $\pm$ 0	104 $\pm$ 1	2 $\pm$ 0	9 $\pm$ 0	299 $\pm$ 1	215 $\pm$ 1	82 $\pm$ 0	0 $\pm$ 0	72.47 $\pm$ 0.19
<i>node-jsonfile</i>	68 $\pm$ 0	204 $\pm$ 0	48 $\pm$ 0	4 $\pm$ 0	0 $\pm$ 0	152 $\pm$ 0	54 $\pm$ 0	45 $\pm$ 0	53 $\pm$ 0	70.39 $\pm$ 0
<i>plural</i>	153 $\pm$ 0	433 $\pm$ 0	110 $\pm$ 0	50 $\pm$ 0	16 $\pm$ 0	273 $\pm$ 0	198 $\pm$ 0	74 $\pm$ 0	1 $\pm$ 0	72.89 $\pm$ 0
<i>pull-stream</i>	351 $\pm$ 0	1,026 $\pm$ 1	237 $\pm$ 1	16 $\pm$ 0	11 $\pm$ 1	774 $\pm$ 0	439 $\pm$ 1	279 $\pm$ 1	56 $\pm$ 0	63.99 $\pm$ 0.13
<i>q</i>	1,051 $\pm$ 0	3,084 $\pm$ 1	1,107 $\pm$ 1	122 $\pm$ 1	59 $\pm$ 1	1,855 $\pm$ 1	137 $\pm$ 1	1,634 $\pm$ 1	83 $\pm$ 1	11.9 $\pm$ 0.02
<i>spacl-core</i>	134 $\pm$ 0	390 $\pm$ 0	144 $\pm$ 1	18 $\pm$ 0	6 $\pm$ 0	228 $\pm$ 1	175 $\pm$ 1	33 $\pm$ 1	1 $\pm$ 0	84.24 $\pm$ 0.62
<i>zip-a-folder</i>	49 $\pm$ 0	142 $\pm$ 0	41 $\pm$ 0	3 $\pm$ 0	1 $\pm$ 0	98 $\pm$ 0	27 $\pm$ 1	3 $\pm$ 0	68 $\pm$ 1	96.94 $\pm$ 0
<i>Total</i>	3,376	9,832	3,162	329	199	6,341	3,051	2,946	303	69.67

TABLE CLVI

SUMMARY OF MUTANTS FOR CODELLAMA-34B-INSTRUCT-NOINSTRUCTIONS-0.0 (RUNS 378, 379, 380, 381, 382). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,408 $\pm$ 1	235 $\pm$ 0	36 $\pm$ 0	45 $\pm$ 0	1,137 $\pm$ 0	696 $\pm$ 0	440 $\pm$ 0	1 $\pm$ 0	61.29 $\pm$ 0.01
<i>countries-and-timezones</i>	106 $\pm$ 0	311 $\pm$ 0	85 $\pm$ 0	8 $\pm$ 0	4 $\pm$ 0	218 $\pm$ 0	174 $\pm$ 0	44 $\pm$ 0	0 $\pm$ 0	79.82 $\pm$ 0
<i>crawler-url-parser</i>	176 $\pm$ 0	498 $\pm$ 0	211 $\pm$ 0	23 $\pm$ 0	20 $\pm$ 0	264 $\pm$ 0	134 $\pm$ 0	112 $\pm$ 0	0 $\pm$ 0	54.47 $\pm$ 0
<i>delta</i>	462 $\pm$ 0	1,348 $\pm$ 0	563 $\pm$ 0	26 $\pm$ 0	18 $\pm$ 0	759 $\pm$ 0	611 $\pm$ 1	116 $\pm$ 1	32 $\pm$ 0	84.77 $\pm$ 0.12
<i>image-downloader</i>	42 $\pm$ 0	126 $\pm$ 0	39 $\pm$ 0	2 $\pm$ 0	0 $\pm$ 0	85 $\pm$ 0	69 $\pm$ 1	15 $\pm$ 1	0 $\pm$ 0	82.62 $\pm$ 0.65
<i>node-dirty</i>	154 $\pm$ 0	448 $\pm$ 0	161 $\pm$ 1	26 $\pm$ 0	9 $\pm$ 0	260 $\pm$ 1	147 $\pm$ 1	103 $\pm$ 1	11 $\pm$ 0	60.6 $\pm$ 0.31
<i>node-geo-point</i>	140 $\pm$ 0	400 $\pm$ 0	85 $\pm$ 0	7 $\pm$ 0	13 $\pm$ 0	308 $\pm$ 0	230 $\pm$ 0	76 $\pm$ 0	0 $\pm$ 0	75.16 $\pm$ 0
<i>node-jsonfile</i>	68 $\pm$ 0	199 $\pm$ 0	42 $\pm$ 0	9 $\pm$ 0	1 $\pm$ 0	148 $\pm$ 0	45 $\pm$ 0	51 $\pm$ 0	52 $\pm$ 0	65.54 $\pm$ 0
<i>plural</i>	153 $\pm$ 0	427 $\pm$ 1	109 $\pm$ 0	57 $\pm$ 0	17 $\pm$ 0	261 $\pm$ 0	189 $\pm$ 0	71 $\pm$ 0	1 $\pm$ 0	72.74 $\pm$ 0.13
<i>pull-stream</i>	351 $\pm$ 0	1,024 $\pm$ 0	225 $\pm$ 1	19 $\pm$ 0	16 $\pm$ 0	780 $\pm$ 1	466 $\pm$ 1	248 $\pm$ 1	66 $\pm$ 0	68.16 $\pm$ 0.09
<i>q</i>	1,051 $\pm$ 0	3,071 $\pm$ 1	1,019 $\pm$ 1	94 $\pm$ 1	60 $\pm$ 0	1,958 $\pm$ 1	137 $\pm$ 1	1,727 $\pm$ 1	94 $\pm$ 1	11.8 $\pm$ 0.05
<i>spacl-core</i>	134 $\pm$ 0	388 $\pm$ 0	158 $\pm$ 0	30 $\pm$ 0	7 $\pm$ 0	200 $\pm$ 0	155 $\pm$ 0	31 $\pm$ 0	1 $\pm$ 0	83.44 $\pm$ 0.04
<i>zip-a-folder</i>	49 $\pm$ 0	143 $\pm$ 0	41 $\pm$ 0	5 $\pm$ 0	1 $\pm$ 0	97 $\pm$ 0	26 $\pm$ 0	4 $\pm$ 0	67 $\pm$ 0	95.88 $\pm$ 0
<i>Total</i>	3,376	9,790	2,972	342	211	6,476	3,080	3,037	325	68.95

TABLE CLVII

SUMMARY OF MUTANTS FOR CODELLAMA-34B-INSTRUCT-ONEMUTATION-0.0 (RUNS 365, 366, 369, 370, 371). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	478 $\pm$ 2	67 $\pm$ 0	6 $\pm$ 0	11 $\pm$ 0	405 $\pm$ 2	244 $\pm$ 2	161 $\pm$ 1	0 $\pm$ 0	60.24 $\pm$ 0.13
<i>countries-and-timezones</i>	106 $\pm$ 0	104 $\pm$ 0	24 $\pm$ 0	1 $\pm$ 0	2 $\pm$ 0	79 $\pm$ 0	66 $\pm$ 0	13 $\pm$ 0	0 $\pm$ 0	83.29 $\pm$ 0.56
<i>crawler-url-parser</i>	176 $\pm$ 0	170 $\pm$ 0	70 $\pm$ 1	8 $\pm$ 0	5 $\pm$ 0	92 $\pm$ 1	50 $\pm$ 1	36 $\pm$ 0	0 $\pm$ 0	58.47 $\pm$ 0.31
<i>delta</i>	462 $\pm$ 0	452 $\pm$ 0	182 $\pm$ 1	4 $\pm$ 0	9 $\pm$ 0	266 $\pm$ 1	221 $\pm$ 1	37 $\pm$ 0	8 $\pm$ 0	86.11 $\pm$ 0.03
<i>image-downloader</i>	42 $\pm$ 0	42 $\pm$ 0	8 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	34 $\pm$ 0	26 $\pm$ 0	8 $\pm$ 0	0 $\pm$ 0	76.47 $\pm$ 0
<i>node-dirty</i>	154 $\pm$ 0	152 $\pm$ 0	50 $\pm$ 0	3 $\pm$ 1	3 $\pm$ 0	99 $\pm$ 1	55 $\pm$ 0	41 $\pm$ 1	3 $\pm$ 0	58.83 $\pm$ 0.32
<i>node-geo-point</i>	140 $\pm$ 0	137 $\pm$ 0	32 $\pm$ 0	0 $\pm$ 0	3 $\pm$ 0	105 $\pm$ 0	74 $\pm$ 0	30 $\pm$ 0	0 $\pm$ 0	71.15 $\pm$ 0
<i>node-jsonfile</i>	68 $\pm$ 0	68 $\pm$ 0	11 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	57 $\pm$ 0	18 $\pm$ 0	18 $\pm$ 0	21 $\pm$ 0	68.42 $\pm$ 0
<i>plural</i>	153 $\pm$ 0	147 $\pm$ 0	39 $\pm$ 0	7 $\pm$ 1	6 $\pm$ 0	101 $\pm$ 1	71 $\pm$ 1	30 $\pm$ 0	0 $\pm$ 0	70.18 $\pm$ 0.16
<i>pull-stream</i>	351 $\pm$ 0	350 $\pm$ 0	67 $\pm$ 0	3 $\pm$ 0	1 $\pm$ 0	280 $\pm$ 0	165 $\pm$ 0	95 $\pm$ 0	20 $\pm$ 0	66 $\pm$ 0.16
<i>q</i>	1,051 $\pm$ 0	1,035 $\pm$ 0	306 $\pm$ 0	26 $\pm$ 0	17 $\pm$ 0	703 $\pm$ 0	46 $\pm$ 0	630 $\pm$ 1	27 $\pm$ 0	10.35 $\pm$ 0.07
<i>spacl-core</i>	134 $\pm$ 0	132 $\pm$ 0	41 $\pm$ 0	3 $\pm$ 0	2 $\pm$ 0	88 $\pm$ 0	63 $\pm$ 0	17 $\pm$ 0	0 $\pm$ 0	78.75 $\pm$ 0
<i>zip-a-folder</i>	49 $\pm$ 0	48 $\pm$ 0	9 $\pm$ 0	0 $\pm$ 0	1 $\pm$ 0	39 $\pm$ 0	19 $\pm$ 0	17 $\pm$ 0	3 $\pm$ 0	56.41 $\pm$ 0
<i>Total</i>	3,376	3,315	905	62	60	2,348	1,118	1,133	82	64.97

TABLE CLVIII

SUMMARY OF MUTANTS FOR GPT-4O-MINI-FULL-0.0 (RUNS 58, 59, 60, 61, 63). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,433 $\pm$ 3	448 $\pm$ 6	0 $\pm$ 0	36 $\pm$ 2	986 $\pm$ 4	599 $\pm$ 4	387 $\pm$ 2	0 $\pm$ 0	60.72 $\pm$ 0.25
<i>countries-and-timezones</i>	106 $\pm$ 0	308 $\pm$ 2	101 $\pm$ 3	0 $\pm$ 0	10 $\pm$ 2	207 $\pm$ 2	171 $\pm$ 3	36 $\pm$ 2	0 $\pm$ 0	82.64 $\pm$ 0.98
<i>crawler-url-parser</i>	176 $\pm$ 0	517 $\pm$ 1	234 $\pm$ 4	0 $\pm$ 0	10 $\pm$ 1	283 $\pm$ 4	178 $\pm$ 3	91 $\pm$ 2	0 $\pm$ 0	66.1 $\pm$ 0.48
<i>delta</i>	462 $\pm$ 0	1,346 $\pm$ 3	643 $\pm$ 5	1 $\pm$ 1	39 $\pm$ 3	701 $\pm$ 4	562 $\pm$ 3	103 $\pm$ 4	35 $\pm$ 1	85.25 $\pm$ 0.52
<i>image-downloader</i>	42 $\pm$ 0	123 $\pm$ 1	57 $\pm$ 1	0 $\pm$ 0	3 $\pm$ 1	67 $\pm$ 1	45 $\pm$ 3	22 $\pm$ 3	0 $\pm$ 0	67.26 $\pm$ 3.9
<i>node-dirty</i>	154 $\pm$ 0	453 $\pm$ 1	191 $\pm$ 4	0 $\pm$ 0	9 $\pm$ 1	262 $\pm$ 4	152 $\pm$ 4	102 $\pm$ 3	7 $\pm$ 1	61.04 $\pm$ 1.08
<i>node-geo-point</i>	140 $\pm$ 0	395 $\pm$ 3	86 $\pm$ 2	0 $\pm$ 0	24 $\pm$ 3	310 $\pm$ 3	219 $\pm$ 4	88 $\pm$ 2	0 $\pm$ 0	71.33 $\pm$ 0.71
<i>node-jsonfile</i>	68 $\pm$ 0	197 $\pm$ 1	45 $\pm$ 1	0 $\pm$ 0	7 $\pm$ 1	153 $\pm$ 2	65 $\pm$ 3	26 $\pm$ 2	62 $\pm$ 2	83.09 $\pm$ 1.74
<i>plural</i>	153 $\pm$ 0	430 $\pm$ 2	112 $\pm$ 2	5 $\pm$ 1	25 $\pm$ 2	313 $\pm$ 1	258 $\pm$ 4	54 $\pm$ 3	1 $\pm$ 1	82.61 $\pm$ 0.84
<i>pull-stream</i>	351 $\pm$ 0	1,039 $\pm$ 2	300 $\pm$ 1	0 $\pm$ 0	14 $\pm$ 2	739 $\pm$ 3	430 $\pm$ 6	244 $\pm$ 3	65 $\pm$ 3	66.96 $\pm$ 0.43
<i>q</i>	1,051 $\pm$ 0	3,080 $\pm$ 5	1,292 $\pm$ 5	5 $\pm$ 2	73 $\pm$ 5	1,784 $\pm$ 9	130 $\pm$ 5	1,591 $\pm$ 7	63 $\pm$ 3	10.82 $\pm$ 0.32
<i>spack-core</i>	134 $\pm$ 0	390 $\pm$ 2	157 $\pm$ 3	0 $\pm$ 0	11 $\pm$ 1	233 $\pm$ 3	198 $\pm$ 3	18 $\pm$ 2	0 $\pm$ 0	91.66 $\pm$ 0.86
<i>zip-a-folder</i>	49 $\pm$ 0	144 $\pm$ 1	62 $\pm$ 2	0 $\pm$ 0	3 $\pm$ 1	82 $\pm$ 1	18 $\pm$ 3	5 $\pm$ 1	59 $\pm$ 2	94.37 $\pm$ 1.68
<i>Total</i>	3,376	9,857	3,727	11	263	6,120	3,026	2,768	292	71.06

TABLE CLIX

SUMMARY OF MUTANTS FOR LLAMA-3.3-70B-INSTRUCT-FULL-0.0 (RUNS 23, 24, 25, 26, 27). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,413 $\pm$ 3	282 $\pm$ 7	0 $\pm$ 0	59 $\pm$ 5	1,130 $\pm$ 6	682 $\pm$ 5	448 $\pm$ 3	1 $\pm$ 0	60.39 $\pm$ 0.19
<i>countries-and-timezones</i>	106 $\pm$ 0	305 $\pm$ 1	65 $\pm$ 2	0 $\pm$ 0	13 $\pm$ 1	240 $\pm$ 2	204 $\pm$ 4	36 $\pm$ 3	0 $\pm$ 0	84.86 $\pm$ 1.34
<i>crawler-url-parser</i>	176 $\pm$ 0	506 $\pm$ 1	179 $\pm$ 3	0 $\pm$ 0	22 $\pm$ 1	328 $\pm$ 2	203 $\pm$ 3	113 $\pm$ 2	0 $\pm$ 0	64.18 $\pm$ 0.65
<i>delta</i>	462 $\pm$ 0	1,338 $\pm$ 3	540 $\pm$ 3	0 $\pm$ 0	49 $\pm$ 3	798 $\pm$ 4	634 $\pm$ 8	128 $\pm$ 5	37 $\pm$ 3	84.01 $\pm$ 0.69
<i>image-downloader</i>	42 $\pm$ 0	122 $\pm$ 1	44 $\pm$ 2	0 $\pm$ 0	5 $\pm$ 1	79 $\pm$ 2	52 $\pm$ 6	27 $\pm$ 4	0 $\pm$ 0	65.52 $\pm$ 6.21
<i>node-dirty</i>	154 $\pm$ 0	453 $\pm$ 2	124 $\pm$ 3	1 $\pm$ 0	11 $\pm$ 2	327 $\pm$ 5	167 $\pm$ 7	142 $\pm$ 2	19 $\pm$ 2	56.68 $\pm$ 1.07
<i>node-geo-point</i>	140 $\pm$ 0	400 $\pm$ 4	39 $\pm$ 2	0 $\pm$ 0	21 $\pm$ 3	361 $\pm$ 4	258 $\pm$ 4	100 $\pm$ 5	0 $\pm$ 0	71.96 $\pm$ 1.09
<i>node-jsonfile</i>	68 $\pm$ 0	199 $\pm$ 1	27 $\pm$ 2	0 $\pm$ 0	5 $\pm$ 1	172 $\pm$ 2	64 $\pm$ 2	38 $\pm$ 2	70 $\pm$ 2	77.81 $\pm$ 1.31
<i>plural</i>	153 $\pm$ 0	429 $\pm$ 2	92 $\pm$ 3	0 $\pm$ 0	27 $\pm$ 2	337 $\pm$ 3	245 $\pm$ 4	92 $\pm$ 5	0 $\pm$ 0	72.79 $\pm$ 1.33
<i>pull-stream</i>	351 $\pm$ 0	1,044 $\pm$ 2	266 $\pm$ 3	0 $\pm$ 0	9 $\pm$ 2	777 $\pm$ 3	465 $\pm$ 2	262 $\pm$ 4	51 $\pm$ 2	66.35 $\pm$ 0.45
<i>q</i>	1,051 $\pm$ 0	3,073 $\pm$ 3	852 $\pm$ 9	0 $\pm$ 1	82 $\pm$ 3	2,221 $\pm$ 12	127 $\pm$ 3	2,011 $\pm$ 13	82 $\pm$ 3	9.43 $\pm$ 0.19
<i>spack-core</i>	134 $\pm$ 0	385 $\pm$ 1	130 $\pm$ 4	0 $\pm$ 0	16 $\pm$ 1	256 $\pm$ 5	208 $\pm$ 5	34 $\pm$ 2	1 $\pm$ 0	85.93 $\pm$ 0.74
<i>zip-a-folder</i>	49 $\pm$ 0	145 $\pm$ 0	26 $\pm$ 2	0 $\pm$ 0	2 $\pm$ 0	119 $\pm$ 2	45 $\pm$ 37	5 $\pm$ 0	69 $\pm$ 38	95.8 $\pm$ 0.06
<i>Total</i>	3,376	9,812	2,665	2	321	7,145	3,352	3,436	330	68.9

TABLE CLX

SUMMARY OF MUTANTS FOR MIXTRAL-8X7B-INSTRUCT-FULL-0.0 (RUNS 360, 361, 362, 363, 364). EACH COLUMN SHOWS THE AVERAGE NUMBER OF MUTANTS FROM ALL RUNS, $\pm$ THE STANDARD DEVIATION.

application	#prompts	Candidates	Invalid	Identical	Duplicate	#mutants	#killed	#survived	#timeout	mut. score
<i>Complex.js</i>	490 $\pm$ 0	1,284 $\pm$ 11	310 $\pm$ 10	0 $\pm$ 0	19 $\pm$ 4	974 $\pm$ 11	592 $\pm$ 10	382 $\pm$ 6	0 $\pm$ 0	60.75 $\pm$ 0.63
<i>countries-and-timezones</i>	106 $\pm$ 0	267 $\pm$ 3	65 $\pm$ 3	2 $\pm$ 0	4 $\pm$ 1	200 $\pm$ 4	159 $\pm$ 5	41 $\pm$ 2	0 $\pm$ 0	79.29 $\pm$ 1.2
<i>crawler-url-parser</i>	176 $\pm$ 0	404 $\pm$ 4	171 $\pm$ 6	0 $\pm$ 0	1 $\pm$ 1	233 $\pm$ 9	120 $\pm$ 6	100 $\pm$ 4	0 $\pm$ 0	54.54 $\pm$ 0.81
<i>delta</i>	462 $\pm$ 0	1,118 $\pm$ 14	450 $\pm$ 7	0 $\pm$ 0	21 $\pm$ 3	668 $\pm$ 9	509 $\pm$ 10	124 $\pm$ 5	35 $\pm$ 1	81.37 $\pm$ 0.89
<i>image-downloader</i>	42 $\pm$ 0	104 $\pm$ 2	38 $\pm$ 2	0 $\pm$ 0	1 $\pm$ 0	66 $\pm$ 2	42 $\pm$ 2	24 $\pm$ 1	0 $\pm$ 0	64.04 $\pm$ 1.97
<i>node-dirty</i>	154 $\pm$ 0	326 $\pm$ 15	121 $\pm$ 7	0 $\pm$ 0	10 $\pm$ 0	205 $\pm$ 8	121 $\pm$ 6	76 $\pm$ 4	8 $\pm$ 2	62.98 $\pm$ 1.27
<i>node-geo-point</i>	140 $\pm$ 0	343 $\pm$ 3	84 $\pm$ 4	0 $\pm$ 0	14 $\pm$ 2	260 $\pm$ 4	175 $\pm$ 5	79 $\pm$ 2	0 $\pm$ 0	68.83 $\pm$ 1.14
<i>node-jsonfile</i>	68 $\pm$ 0	152 $\pm$ 7	23 $\pm$ 2	0 $\pm$ 0	4 $\pm$ 1	129 $\pm$ 5	54 $\pm$ 3	28 $\pm$ 4	47 $\pm$ 2	78.28 $\pm$ 2.19
<i>plural</i>	153 $\pm$ 0	302 $\pm$ 3	72 $\pm$ 2	0 $\pm$ 0	10 $\pm$ 2	230 $\pm$ 3	171 $\pm$ 4	59 $\pm$ 2	0 $\pm$ 0	74.45 $\pm$ 1.19
<i>pull-stream</i>	351 $\pm$ 0	928 $\pm$ 10	245 $\pm$ 8	1 $\pm$ 0	6 $\pm$ 1	682 $\pm$ 9	393 $\pm$ 7	243 $\pm$ 4	46 $\pm$ 2	64.38 $\pm$ 0.61
<i>q</i>	1,051 $\pm$ 0	2,409 $\pm$ 34	763 $\pm$ 18	3 $\pm$ 0	49 $\pm$ 4	1,643 $\pm$ 20	115 $\pm$ 2	1,458 $\pm$ 22	70 $\pm$ 3	11.23 $\pm$ 0.29
<i>spack-core</i>	134 $\pm$ 0	330 $\pm$ 3	151 $\pm$ 3	0 $\pm$ 0	3 $\pm$ 1	179 $\pm$ 1	134 $\pm$ 0	22 $\pm$ 1	1 $\pm$ 0	85.9 $\pm$ 0.6
<i>zip-a-folder</i>	49 $\pm$ 0	118 $\pm$ 3	38 $\pm$ 3	0 $\pm$ 0	0 $\pm$ 0	80 $\pm$ 4	26 $\pm$ 4	45 $\pm$ 4	8 $\pm$ 1	43.46 $\pm$ 3.81
<i>Total</i>	3,376	8,086	2,532	7	142	5,547	2,610	2,682	215	63.81

P. StrykerJS standard mutation operators

Table CLXI shows the results of running the standard mutation operators of *StrykerJS* on the subject applications.

application	#mutants	#killed	#survived	#timeout	mut. score	time
<i>Complex.js</i>	1,302	763	539	0	58.60	405.08
<i>countries-and-timezones</i>	140	134	6	0	95.71	142.37
<i>crawler-url-parser</i>	226	143	83	0	63.27	433.53
<i>delta</i>	834	686	88	60	89.45	2,747.04
<i>image-downloader</i>	43	28	11	4	74.42	284.20
<i>node-dirty</i>	160	78	56	26	65.00	215.93
<i>node-geo-point</i>	158	98	60	0	62.03	357.70
<i>node-jsonfile</i>	61	31	5	25	91.80	188.91
<i>plural</i>	180	143	37	0	79.44	53.66
<i>pull-stream</i>	474	318	116	40	75.53	694.33
<i>q</i>	1,058	68	927	63	12.38	7,075.02
<i>spacel-core</i>	259	239	20	0	92.28	1,053.16
<i>zip-a-folder</i>	74	38	8	28	89.19	513.27
<i>Total</i>	4,969	2,767	1,956	246	—	14,164.20

TABLE CLXI
RESULTS OF APPLYING THE STANDARD MUTATION OPERATORS OF STRYKERJS.

Q. String edit distance measurements for different LLMs

Table CLXII shows average string similarity for each project, for each of the five LLMs under consideration.

Project	<i>codellama-34b-instruct</i>	<i>codellama-13b-instruct</i>	<i>llama-3.3-70b-instruct</i>	<i>gpt-4o-mini</i>	<i>mistral-8x7b-instruct</i>
<i>Complex.js</i>	4.27	4.63	6.90	6.35	8.95
<i>countries-and-timezones</i>	11.13	9.78	11.71	11.33	13.65
<i>crawler-url-parser</i>	9.50	8.82	10.71	10.05	12.65
<i>delta</i>	9.55	8.88	13.01	11.44	14.15
<i>image-downloader</i>	12.67	10.74	12.66	13.88	14.21
<i>node-dirty</i>	7.53	7.42	12.08	10.20	12.08
<i>node-geo-point</i>	8.86	8.15	8.97	9.53	15.27
<i>node-jsonfile</i>	9.73	10.07	11.10	11.00	10.81
<i>plural</i>	8.14	6.29	9.27	8.31	10.37
<i>pull-stream</i>	6.72	8.71	7.82	8.06	9.67
<i>q</i>	8.61	9.71	11.17	9.92	13.23
<i>spacel-core</i>	9.30	10.84	8.95	8.70	12.61
<i>zip-a-folder</i>	9.85	12.38	11.87	11.23	11.61

TABLE CLXII
AVERAGE STRING SIMILARITY TO THE ORIGINAL CODE FRAGMENTS THAT THEY REPLACE, FOR MUTANTS GENERATED USING FIVE LLMs AT TEMPERATURE 0.0.