TABLE 10
Compiler bug isolation effectiveness comparison with two state-of-the-art approaches (under Setting-1 in RQ1)-Details

Subject	Approach	Num. Top-1	⇑$Top-1$ (%)	Num. Top-5	⇑$Top-5$ (%)	Num. Top-10	⇑$Top-10$ (%)	Num. Top-20	⇑$Top-20$ (%)	MFR	⇑MFR (%)	MAR	⇑MAR (%)
GCC	DiWi [3]	7	57.14	19	36.84	32	28.13	40	25.00	23.27	35.67	23.31	32.26
		5	120.00	19	36.84	35	17.14	44	13.64	23.70	36.84	24.13	34.56
		7	57.14	22	18.18	30	36.67	40	25.00	23.43	36.11	23.88	33.88
		7	57.14	19	36.84	32	28.13	40	25.00	23.43	36.11	23.88	33.88
	RecBi [4]	8	37.50	24	8.33	35	17.14	45	11.11	19.33	22.56	19.90	20.65
		8	50.00	23	13.04	35	17.14	42	19.05	18.28	18.11	18.90	16.46
		8	37.50	21	23.81	33	24.24	40	25.00	24.68	39.34	25.25	37.47
		8	37.50	23	13.04	35	17.14	42	19.05	19.33	22.56	19.90	20.65
	LLM4CBI	11	-	24	-	37	-	50	-	14.97	-	15.79	-
		12	-	26	-	41	-	46	-	14.47	-	14.79	-
		11	-	26	-	41	-	50	-	15.10	-	15.93	-
		11	-	26	-	41	-	50	-	14.97	-	15.79	-
LLVM	DiWi [3]	2	400.00	17	41.18	27	22.22	40	20.00	27.17	45.53	27.19	45.38
		5	80.00	18	33.33	27	22.22	42	14.29	26.08	43.25	26.20	43.32
		4	150.00	20	20.00	29	13.79	40	20.00	26.83	44.84	26.90	44.80
		4	150.00	18	33.33	27	22.22	40	20.00	26.83	44.84	26.90	44.80
	RecBi [4]	6	66.67	16	50.00	30	10.00	41	17.07	25.32	41.55	25.65	42.11
		6	66.67	21	14.29	28	17.86	44	9.09	24.52	39.64	24.58	39.59
		6	66.67	21	14.29	29	13.79	44	9.09	25.53	42.03	25.63	42.06
		6	66.67	21	14.29	29	13.79	44	9.09	25.32	41.55	25.63	42.06
	LLM4CBI	10	-	24	-	33	-	46	-	14.80	-	14.85	-
		9	-	22	-	33	-	48	-	15.08	-	15.16	-
		10	-	24	-	39	-	49	-	12.77	-	12.91	-
		10	-	24	-	33	-	48	-	14.80	-	14.85	-

Note: Columns "⇑ ∗" present the improvement rates (%) of LLM4CBI over the compared approaches regarding various metrics. The numbers in a row highlighted with cyan refer to the median results of three times running.

TABLE 11
Compiler bug isolation effectiveness comparison with two state-of-the-art approaches (under Setting-2 in RQ1)-Details

Subject	Approach	Num. Top-1	⇑$Top-1$ (%)	Num. Top-5	⇑$Top-5$ (%)	Num. Top-10	⇑$Top-10$ (%)	Num. Top-20	⇑$Top-20$ (%)	MFR	⇑MFR (%)	MAR	⇑MAR (%)
GCC	DiWi [3]	6	33.33	20	20.00	32	9.38	42	4.76	22.15	28.13	22.18	23.67
		6	33.33	20	20.00	33	6.06	41	7.32	21.18	24.83	21.68	21.91
		4	100.00	19	26.32	33	6.06	41	7.32	23.13	31.17	23.65	28.41
		6	33.33	20	20.00	33	6.06	41	7.32	22.15	28.13	22.18	23.67
	RecBi [4]	6	33.33	22	9.09	34	2.94	43	2.33	18.88	15.68	19.21	11.87
		7	14.29	21	14.29	33	6.06	42	4.76	18.53	14.09	19.07	11.22
		7	14.29	22	9.09	34	2.94	41	7.32	19.75	19.39	20.44	17.17
		7	14.29	22	9.09	34	2.94	42	4.76	18.88	15.68	19.21	11.87
	LLM4CBI	7	-	24	-	34	-	44	-	16.97	-	17.13	-
		8	-	25	-	35	-	42	-	15.46	-	15.83	-
		8	-	24	-	36	-	44	-	15.92	-	16.33	-
		8	-	24	-	35	-	44	-	15.92	-	16.33	-
LLVM	DiWi [3]	5	60.00	21	14.29	28	25.00	39	10.26	25.62	35.79	25.74	35.59
		5	60.00	19	26.32	32	9.38	39	10.26	26.37	37.62	26.44	37.29
		4	100.00	19	26.32	29	20.69	39	10.26	26.37	37.62	26.43	37.27
		5	60.00	19	26.32	29	20.69	39	10.26	26.37	37.62	26.43	37.27
	RecBi [4]	7	14.29	21	14.29	34	2.94	40	7.50	19.92	17.42	20.02	17.18
		5	60.00	20	20.00	31	12.90	42	2.38	19.35	14.99	19.40	14.54
		7	14.29	21	14.29	34	2.94	40	7.50	19.43	15.34	19.76	16.09
		7	14.29	21	14.29	34	2.94	40	7.50	19.43	15.34	19.76	7.43
	LLM4CBI	7	-	22	-	34	-	45	-	16.20	-	16.35	-
		8	-	22	-	34	-	43	-	17.18	-	17.36	-
		8	-	24	-	36	-	42	-	16.45	-	16.58	-
		8	-	24	-	35	-	43	-	16.45	-	16.58	-

Note: Columns "⇑ ∗" present the improvement rates (%) of LLM4CBI over the compared approaches regarding various metrics. The numbers in a row highlighted with cyan refer to the median results of three times running.

TABLE 12
Compiler bug isolation effectiveness comparison with variant approaches -Details

Subject	Approach	Num. Top-1	$\Uparrow_{Top-1}$ (%)	Num. Top-5	$\Uparrow_{Top-5}$ (%)	Num. Top-10	$\Uparrow_{Top-10}$ (%)	Num. Top-20	$\Uparrow_{Top-20}$ (%)	MFR	$\Uparrow_{MFR}$ (%)	MAR	$\Uparrow_{MAR}$ (%)
GCC	LLM4CBI$_{ep}$	6	71.43	24	8.33	34	20.59	42	19.05	17.13	12.61	17.45	9.51
		7	57.14	22	18.18	35	17.14	41	21.95	17.62	15.04	18.28	13.62
		7	57.14	24	8.33	33	24.24	43	16.28	18.01	16.88	18.45	14.42
		7	57.14	24	8.33	34	20.59	42	19.05	17.62	15.04	18.28	13.62
	LLM4CBI$_{sp}$	7	57.14	23	13.04	33	24.24	40	25.00	19.12	21.71	19.67	19.73
		6	83.33	23	13.04	32	28.13	41	21.95	18.88	20.71	19.13	17.46
		6	83.33	22	18.18	33	24.24	40	25.00	19.78	24.32	19.82	20.33
		6	83.33	23	13.04	33	24.24	40	25.00	19.12	21.71	19.67	19.73
	LLM4CBI$_{rand}$	6	83.33	22	18.18	33	24.24	41	21.95	18.23	17.88	18.90	16.46
		6	83.33	23	13.04	33	24.24	43	16.28	19.71	24.05	19.94	20.81
		7	57.14	23	13.04	34	20.59	42	19.05	17.89	16.32	18.24	13.43
		6	83.33	23	13.04	33	24.24	42	19.05	18.23	17.88	18.73	15.70
	LLM4CBI$_{selnov}$	8	37.50	22	18.18	35	17.14	39	28.21	19.38	22.76	19.64	19.60
		8	37.50	22	18.18	35	17.14	41	21.95	19.00	21.21	19.67	19.73
		8	37.50	23	13.04	31	32.26	40	25.00	19.53	23.35	19.58	19.36
		8	37.50	22	4.55	35	17.14	40	25.00	19.38	22.76	19.64	19.60
	LLM4CBI	11	-	24	-	37	-	50	-	14.97	-	15.79	-
		12	-	26	-	41	-	46	-	14.47	-	14.79	-
		11	-	26	-	41	-	50	-	15.10	-	15.93	-
		11	-	26	-	41	-	50	-	14.97	-	15.79	-
LLVM	LLM4CBI$_{ep}$	8	25.00	21	14.29	32	3.13	43	11.63	17.27	14.30	17.49	15.09
		8	25.00	20	20.00	31	6.45	42	14.29	18.98	22.02	19.08	22.17
		8	25.00	20	20.00	31	6.45	43	11.63	19.02	22.19	19.78	24.92
		8	25.00	20	20.00	31	6.45	43	11.63	18.98	22.02	19.08	22.17
	LLM4CBI$_{sp}$	7	42.86	23	4.35	32	3.13	39	23.08	17.32	14.55	17.69	16.05
		6	66.67	21	14.29	30	10.00	43	11.63	17.75	16.62	17.88	16.95
		7	42.86	23	4.35	32	3.13	39	23.08	18.14	18.41	18.86	21.26
		7	42.86	23	4.35	32	3.13	39	23.08	17.75	16.62	17.88	16.95
	LLM4CBI$_{rand}$	8	25.00	23	4.35	31	6.45	41	17.07	16.10	8.07	16.46	9.78
		8	25.00	23	4.35	30	10.00	40	20.00	17.18	13.85	17.83	16.71
		9	11.11	22	9.09	32	3.13	43	11.63	16.40	9.76	16.52	10.11
		8	25.00	23	4.35	31	6.45	41	17.07	16.40	9.76	16.52	10.11
	LLM4CBI$_{selnov}$	8	25.00	22	9.09	32	3.13	44	9.09	15.57	4.95	15.61	4.87
		7	42.86	22	9.09	31	6.45	40	20.00	16.55	10.57	16.82	11.71
		6	66.67	16	50.00	31	6.45	40	20.00	17.80	16.85	17.85	16.81
		7	42.86	22	9.09	31	6.45	40	20.00	16.55	10.57	25.63	31.92
	LLM4CBI	10	-	24	-	33	-	46	-	14.80	-	14.85	-
		9	-	22	-	33	-	48	-	15.08	-	15.16	-
		10	-	24	-	39	-	49	-	12.77	-	12.91	-
		10	-	24	-	33	-	48	-	14.80	-	14.85	-

Note: Columns "$\Uparrow *$" present the improvement rates (%) of LLM4CBI over the compared approaches regarding various metrics. The numbers in a row highlighted with cyan refer to the median results of three times running.