Hardware Model Checking Competition 2025

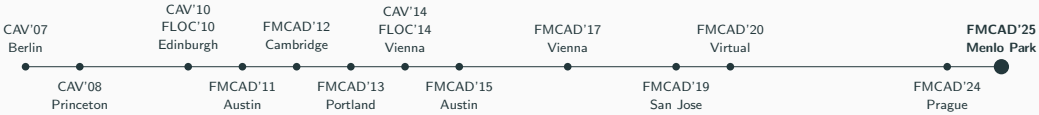
(13th Edition)

Mathias Preiner Nils Froleyks Armin Biere

https://hwmcc.github.io/2025

FMCAD, October 6-10, 2025,
Menlo Park, CA, USA

HWMCC Editions



Goals of HWMCC

- Collect large set of publicly available bit-level and word-level benchmarks
- Encourage researchers to work on novel model checking engines
- Provide a platform for comparison

Previous Years

- **Bit-level** tracks
  - AIGER format (`https://github.com/arminbiere/aiger`)
  - **SINGLE** safety (bad state) property track
  - How **DEEP** model checkers go on unsolved SINGLE instances
  - **LIVENESS** track (single "justice" property)
- **Word-level** tracks
  - BTOR2 format (`https://github.com/boolector/btor2tools`)
  - Introduced in **HWMCC'19** for the first time
  - **SINGLE** safety property tracks
    - **Bit-vectors**
    - **Bit-vectors+arrays**

This Year

Word-level Tracks (BTOR2 format)

- *Single* safety property tracks:
  bit-vectors, bit-vectors+arrays
- **Sat**: BTOR2 witness mandatory (new)

Bit-level Tracks (AIGER 1.9 format)

- *Single* safety property track
  - Benchmarks from word-level BV track
  - **Sat:** AIGER witness mandatory
  - **Unsat:** Certifaiger certificates mandatory
- *Liveness* property track (new)
  - HWMCC'17 liveness track benchmarks
  - **Sat**: AIGER witness mandatory

Competition Setup

- 3600s wall-clock limit
- 120GB memory limit
- 16 core/32 threads
- One machine per job
- Stanford CENTAUR cluster

Witness/Certificate Validation

- 36000s wall-clock limit
- 16GB memory limit
- 2 core/4 threads
- Answer only counted if validation succeeds

Benchmarks

2025 Submissions

- 13 array benchmarks
  - **236** safety properties
  - submitted by Jannis Harder (YosysHQ)
- 611 bit-vector benchmarks
  - **611** safety properties
  - submitted by Guangyu Hu, Xiaofeng zhou, Hongce Zhang, Wei Zhang (The Hong Kong University of Science and Technology)
- 7168 bit-vector, 1291 array benchmarks
  - **8459** safety properties
  - submitted by Po-Chun Chien (LMU Munich)

From Previous Years

- **5978** (2024), **35** (2020), **4802** (2019), **264** (2017 liveness)

Benchmark Selection

- **20532** BTOR2 benchmarks in total
  - 15378 bit-vector benchmarks, 5154 bit-vector+array benchmarks
- **Removed** "easy" benchmarks
  - Solved within 10s wall-clock by all 2025 participants
  - 3491 bit-vector, 1921 array benchmarks
- Divided all benchmarks into **12 families**
  - Grouped by submitter and submission year
- **Random selection** of ∼300 benchmarks from remaining *unique* benchmarks
  - Picked at most N **benchmarks per family**
    bit-vectors: $N = 35$, arrays: $N = 60$
- **Selected**:
  - Safety tracks: **330 bit-vector** and **310 bit-vector+array** benchmarks
    - 330 bit-vector benchmarks translated to AIGER with btor2aiger
  - Liveness track: **264** benchmarks

Model Checker Submissions

Submissions from 11 Teams (-2 from 2024)

- Bit-level safety: **6** competitive (-2 from 2024), 2 non-competitive
- Bit-level liveness: **1** competitive (-2 from 2017), 1 non-competitive
- Word-level BV: **6** competitive (-1 from 2024), 1 non-competitive
- Word-level Arrays: **3** competitive (-1 from 2024), 1 non-competitive

Non-Competitive Model Checkers (submitted by organizers)

- **voiraig**: Bit-level safety track
  - Reference model checker for AIGER + certificates by Nils Froleyks (JKU)
- **BtorMC**: Word-level safety tracks
  - Reference model checker for BTOR2 by Aina Niemetz, Mathias Preiner, Armin Biere (Stanford, Freiburg)
- **ABC Superprove**: Bit-level safety and liveness tracks
  - HWMCC winner of previous editions by Robert K. Brayton, Baruch Sterin, Alan Mishchenko (Berkely), HWMCC'20 version from OSS CAD Suite (release 2025-09-10)

Teams

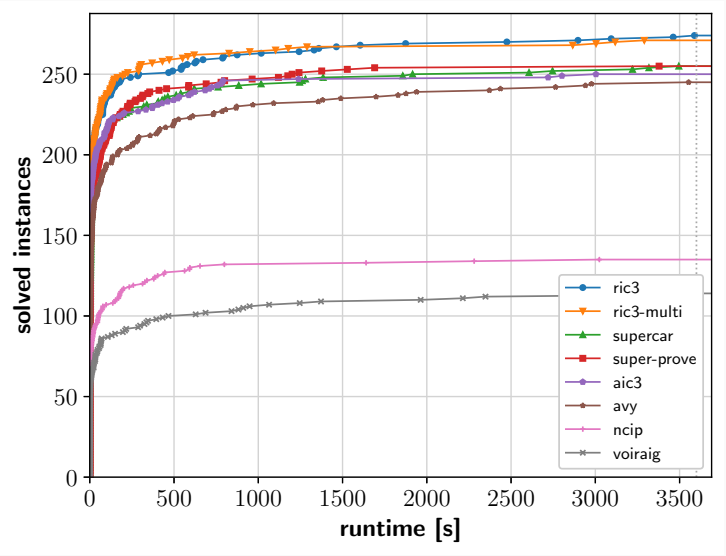
- **aIC3**: Xiaofeng Zhou, Guangyu Hu (HKUST), Hongce Zhang (HKUST & HKUST(GZ)), Wei Zhang (HKUST) (new)
- **AVR**: Aman Goel (AWS), Karem Sakallah (University of Michigan)
- **AVY**: Yakir Vizel, Basel Khouri, Andrew Luka (Technion), Arie Gurfinkel (UWaterloo)
- **Btor2-Select**: John-Lu (UWaterloo), Po-Chun Chien (LMU Munich), Nian-Ze Lee (NTU), Li-Zhe Liu (NTU), Vijay Ganesh (Georgia Tech)
- **IC3-ng**: Guangyu Hu (HKUST), Hongce Zhang (HKUST & HKUST(GZ)) (new)
- **NCIP**: Tobias Faller (University of Freiburg)
- **nuXmv**: Alberto Griggio (FBK)
- **Pono**: Áron Ricardo Perez-Lopez, Clark Barrett (Stanford University), Ahmed Irfan (SRI), Po-Chun Chien (LMU Munich)
- **rIC3**: Yuheng Su, Qiusong Yang, Yiwei Ci (Chinese Academy of Sciences)
- **rIC3-multi**: Chen Wish (NTU) (new)
- **Supercar**: Yibo Dong, Yechuan Xia, Hongtai Zhu, Jianwen Li, Geguang Pu (East China Normal University)

HWMCC'25 Ranking

- **4 Tracks**
  - Bit-level safety
  - Bit-level liveness
  - Word-level bit-vectors
  - Word-level bit-vectors+arrays
- **Ranked** by number of solved benchmarks (sat+unsat)
  - Gold: 1st place
  - Silver: 2nd place
  - Bronze: 3rd place
- 3 medals per track: **12 medals** in total

Results

Bit-Level Safety Track: Solved



Bit-Level Safety Track: Solved

	solved	sat	unsat	real [s]	cpu [s]	mem [mb]	best	uniq
ric3	274	99	175	207908	3272359	4146711	44	3
ric3-multi	271	101	170	238854	3254708	3286020	50	2
super-prove	255	88	167	283251	3375253	1920039	35	2
supercar	255	97	158	307317	4100273	3074068	4	1
aic3	250	81	169	308768	3826429	2433956	130	1
avy	245	83	162	352625	2454704	2202759	4	0
ncip	135	58	77	327572	3920514	24762956	5	1
voiraig	114	33	81	804417	804622	362791	12	0

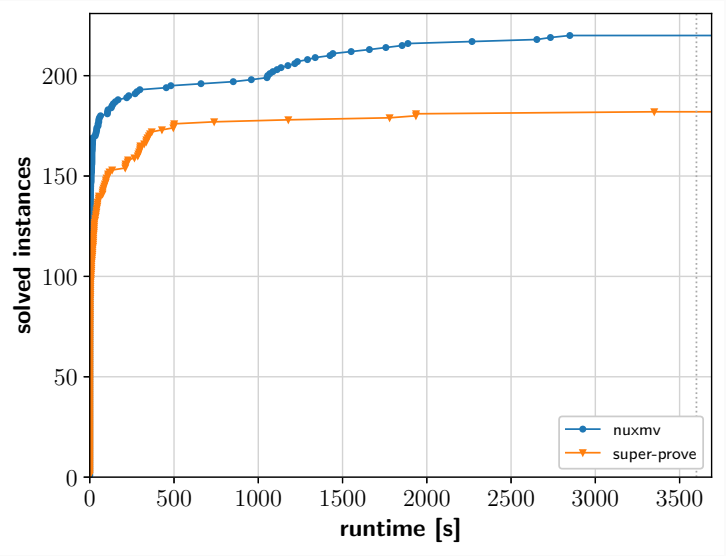
330 benchmarks, 1h wall-clock time limit, 120GB memory limit

Bit-Level Safety Track: Certified

		certified	sat		unsat		uniq	timeout
1	ric3	274	99		175		4	0
2	ric3-multi	266	101		165	(-4)	2	1
3	supercar	255	97		158		1	0
	aic3	248	81		167	(-2)	0	0
	avy	226	68	(-15)	158	(-3)	0	1
	ncip	135	58		77		1	0
	voiraig	114	33		81		0	0
	super-prove	80	80	(-8)	167	(-167)	0	0

Negative numbers in orange (gray) are invalid (missing) certificates.

Bit-Level Liveness Track: Solved



Bit-Level Liveness Track: Solved

	solved	sat	unsat	real [s]	cpu [s]	mem [mb]	best	uniq
nuxmv	220	107	113	199944	2360722	2345505	174	42
super-prove	182	86	96	209945	631775	4967024	50	4

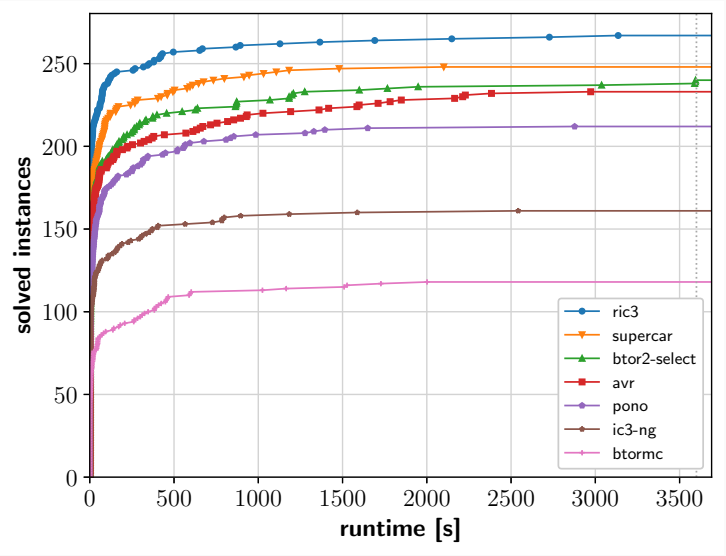
264 benchmarks, 1h wall-clock time limit, 120GB memory limit

Bit-Level Liveness Track: Certified

		certified	sat		unsat	uniq
1	nuxmv	220	107		113	127
	super-prove	96	0	(-86)	96	3

Negative numbers in gray are missing certificates.

Word-Level Bit-Vectors Track: Solved



Word-Level Bit-Vectors Track: Solved

	solved	sat	unsat	real [s]	cpu [s]	mem [mb]	best	uniq
ric3	267	97	170	196539	3001965	4799963	158	15
supercar	248	97	151	316318	4222838	3102287	8	1
btor2-select	240	86	154	350761	1687279	2518731	14	1
avr	233	83	150	380987	5440794	8551739	57	6
pono	212	80	132	346052	5503963	13212915	6	0
ic3-ng	161	47	114	621360	1643612	3220367	24	0
btormc	118	70	48	781675	781898	1474747	26	0

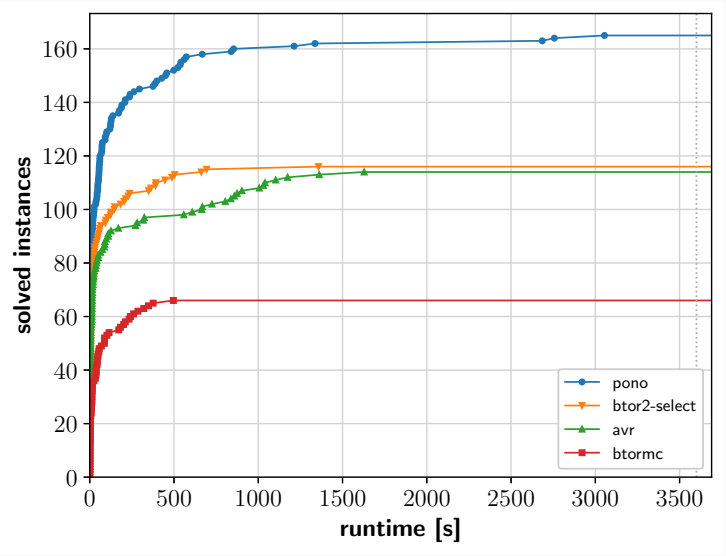
330 benchmarks, 1h wall-clock time limit, 120GB memory limit

Word-Level Bit-Vectors Track: Certified

		certified	sat		unsat	uniq	timeout
1	ric3	267	97		170	15	
2	supercar	248	97		151	1	
3	btor2-select	240	86		154	1	
	avr	229	79	(-4)	150	6	
	pono	212	80		132	0	
	ic3-ng	161	47		114	0	
	btormc	118	70		48	0	

Negative numbers in orange are invalid certificates.

Word-Level Arrays Track: Solved



Word-Level Arrays Track: Solved

	solved	sat	unsat	real [s]	cpu [s]	mem [mb]	best	uniq
pono	165	55	110	507922	5829196	11337189	92	45
btor2-select	116	58	58	460647	1579191	12021545	28	3
avr	114	45	69	367836	4322341	19333338	33	7
btormc	66	53	13	838396	838600	1622844	25	0

310 benchmarks, 1h wall-clock time limit, 120GB memory limit

Word-Level Arrays Track: Certified

		certified	sat		unsat	uniq	timeout
1	pono	165	55		110	45	0
2	avr	114	45		69	7	0
3	btor2-select	112	54	(-3)	58	2	1
	btormc	66	53		13	0	0

Negative numbers in orange are invalid certificates.

Results Summary

	gold	silver	bronze
ric3	2		
nuxmv	1		
pono	1		
supercar		1	1
ric3-multi		1	
avr		1	
btor2-select			2

Congratulations to the winners!

Competition Statistics

Solving

- Bit-Level
  - 900 hours wall-clock
  - 7778 hours CPU time
  - 47TB memory usage
  - 86 out of 594 unsolved
- Word-Level
  - 1436 hours wall-clock
  - 9681 hours CPU time
  - 77TB memory usage
  - 171 out of 640 unsolved

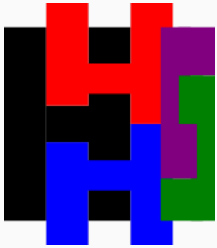
Certification

- Bit-Level
  - 33 hours wall-clock
  - 124GB memory usage
  - 2 timeouts
  - 833 sat witnesses (23 incorrect) checked with aigsim
  - 992 unsat certificates (9 incorrect) checked with certfaiger
- Word-Level
  - 19 hours wall-clock
  - 11GB memory usage
  - 1 timeout
  - 1180 sat witnesses (7 incorrect) checked with btorsim

Summary

HWMCC'25

- **11 teams**
- **16 competitive** entries in 4 tracks
- Bit-level **liveness** track



HWMCC'??

- Mandatory unsat certificates for word-level tracks
- Word-level liveness track

Thank you to all teams and benchmark submitters!