Model Checking Contest 2024

14th edition, Geneva, Switzerland, June 25, 2024

Complete Results for the 2024 Edition of the Model Checking Contest

Last Updated
July 7, 2024

1. Introduction

Important Note : we lately discovered a problem with the colored instances of BlocksWorld leading to a problem with results for 2024. We thus updated data from the contest by not considering these instances for 2024 (we will deal with corrections to be done for 2025)

This page summarizes the results for the 2024 edition of the Model Checking Contest (MCC'2024). This page is divided in three sections:

- First, we **list the qualified tools** for the MCC'2024,
- Then, we provide some informations about **the experimental conditions of the MCC'2024**,
- Then, we present an access **to details about results**,
- Then, we provide the **list of winners of the MCC'2024**,
- Finally, we provide an attempt to evaluate **tool reliability** based on the comparison of the results provided in the contest.

IMPORTANT: all these documents can be reused in scientific material and papers but must respect the Creative Common license CC-BY-NC-SA.

2. List of Qualified Tools in 2024

15 tools where submitted this year. They all successfully went through a qualification process requiring about 1625 runs (each tool had to answer each examination for the first instance of each «known» model).

We introduce two classes of tools :

- ***Original tools:*** these are tools prepared for the current year ; it may also be a association of tools or some previously existing tools with additional features. It must be something original.
- ***Reference tools:*** these are existing reference tools that have been repackaged for the model checking contest so that they can compete as a identified version ; this may help to measurer the evolution of techniques or to make a comparison between a new version of a given tool and a previous one.

Data about these tools are summarized in the table below. For any tool, you can download the disk image that was provided with all its data. You may use these to reproduce measures locally and perform comparison with your own tool on the same benchmark. Please note that one tool (with two variants) was out of competition this year: this was agreed between the tool developer and the organizers and is part of an experiment with precomputed deep-learning.

IMPORTANT: all tool developers agreed to provide the original image disk embedding the tool they submitted his year (see links in the table below). You may operate these tools on your own. To do so, you need the second **disk image (mounted by the other one) that contains all models for 2024 together with the produced formulas**. This image is mounted with the default configuration, as well as in he default disk image provided in the tool submission kit (**see here**).

IMPORTANT: You also have access to the **archive containing all models and the corresponding formulas for 2024 here**.

IMPORTANT: Note that Gold2023 is an hybrid artificial tool made with the tools that won categories in 2023. It correspond usually to several virtual machines, so the corresponding archive is a bit larger. For the MCC'2024, Gold 2023 is composed as follows:

- ***tedd-c*** for the StateSpace Category
- ***ITS-Tools*** for the GlobalProperties, UpperBounds categories, LTL Formulas categories
- ***Lola+red*** for Reachability Formulas
- ***Tapaal*** for the CTL Formulas

The table below presents all participating tools for 2024.

Summary of the Participating Tools

Tool name	Supported Petri nets	Representative Author	Origin	Type of execution	Link to the submitted disk image
Tools competing in 2024					
Gold2023	P/T and colored	Fabrice Kordon	Aalborg/Paris/ Toulouse	Collateral Processing	
GreatSPN+red	P/T and colored	Yann Thierry-Mieg	Univ. Torino (Italy) & Sorbonne Université (France) (driver)	Collateral Processing	
ITS-Tools	P/T and colored	Yann Thierry-Mieg	Sorbonne Université (France)	Collateral Processing	
LoLA	P/T and colored	Karsten Wolf	Univ. Rostock (Germany)	Collateral Processing	
LTSMin+red	P/T and colored	Yann Thierry-Mieg	Univ. Twente (the Netherlands) & Sorbonne Université (France) (driver)	Collateral Processing	
NoHD	P/T and colored	Benjamin Smith	No affiliation (United States)	Collateral Processing	
smpt	P/T and colored	Nicolas Amat	LAAS-CNRS (France) & IMDEA (Spain)	Collateral Processing	
SVSKit	P/T and colored	Damien Morard	Univ. Geneva (Switzerland)	Sequential Processing	
Tapaal	P/T and colored	Jiri Srba	Aalborg University (Denmark)	Collateral Processing	
TINA.tedd	P/T and colored	Bernard Berthomieu	LAAS-CNRS (France)	Collateral Processing	
Reference Tools					
GreatSPN-meddly	P/T and colored	Elvio Amparore	Univ. Torino (Italy)	Collateral Processing	
LTSMin	P/T and colored (unfolding by ITS)	Jeroen Meijer and Tom van Dijk (2019) edited by Yann Thierry-Mieg	Univ. Twente (the Netherlands) & Sorbonne Université (France)	Collateral Processing	

The table below lists the techniques reported per examination (and for all the tool variants when applicable).

Techniques Reported by the Participating Tools (per examination)

Tool name	StateSpace	GlobalProperties	UpperBounds	Reachability	CTL	LTL	All together
Tools competing in 2024							
Gold2023	COLLATERAL_PROCESSING DECISION_DIAGRAMS EXPLICIT LATTICE_POINTS_COUNTING LINEAR_EQUATIONS STRUCTURAL_REDUCTION	BESTFIRST_WALK COLLATERAL_PROCESSING CPN_APPROX DEADLOCK_TEST DECISION_DIAGRAMS EXHAUSTIVE_WALK EXPLICIT INITIAL_STATE INVARIANTS LTSMIN	BESTFIRST_WALK CPN_APPROX DECISION_DIAGRAMS INITIAL_STATE PARIKH_WALK RANDOM_WALK SAT_SMT TOPOLOGICAL USE_NUPN	BESTFIRST_WALK COLLATERAL_PROCESSING CPN_APPROX EXHAUSTIVE_WALK EXPLICIT INITIAL_STATE OVER_APPROXIMATION PARIKH_WALK	COLLATERAL_PROCESSING CPN_APPROX CTL_CZERO EXPLICIT LP_APPROX QUERY_REDUCTION SAT_SMT STATE_COMPRESSION STRUCTURAL_REDUCTION STUBBORN_SETS	DECISION_DIAGRAMS EXPLICIT HOA INITIAL_STATE KNOWLEDGE LENGTHENING_INSENSITIVE LTSMIN PARTIAL_ORDER REACHABILITY_KNOWLEDGE	BESTFIRST_WALK COLLATERAL_PROCESSING CPN_APPROX CTL_CZERO DEADLOCK_TEST DECISION_DIAGRAMS EXHAUSTIVE_WALK EXPLICIT HOA

	TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	MARKED_SUFFIX_TEST PARIKH_WALK PARTIAL_ORDER PROBABILISTIC_WALK QUASILIVENESS_TEST RANDOM_WALK SAT_SMT SCC_TEST SIPHON_TEST SKELETON_TEST STRUCTURAL STRUCTURAL_REDUCTION TOPOLOGICAL TRIVIAL_UNMARKED_SCC_TEST USE_NUPN		PROBABILISTIC_WALK RANDOM_WALK SAT_SMT STATE_COMPRESSION STRUCTURAL_REDUCTION STUBBORN_SETS TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	TRACE_ABSTRACTION_REFINEMENT UNFOLDING_TO_PT	SAT_SMT SHORTENING_INSENSITIVE STACK_TEST STUTTER_TEST TOPOLOGICAL USE_NUPN	INITIAL_STATE INVARIANTS KNOWLEDGE LATTICE_POINTS_COUNTING LENGTHENING_INSENSITIVE LINEAR_EQUATIONS LP_APPROX LTSMIN MARKED_SUFFIX_TEST OVER_APPROXIMATION PARIKH_WALK PARTIAL_ORDER PROBABILISTIC_WALK QUASILIVENESS_TEST QUERY_REDUCTION RANDOM_WALK REACHABILITY_KNOWLEDGE SAT_SMT SCC_TEST SHORTENING_INSENSITIVE SIPHON_TEST SKELETON_TEST STACK_TEST STATE_COMPRESSION STRUCTURAL STRUCTURAL_REDUCTION STUBBORN_SETS STUTTER_TEST TOPOLOGICAL TRACE_ABSTRACTION_REFINEMENT TRIVIAL_UNMARKED_SCC_TEST UNFOLDING_TO_PT USE_NUPN
GreatSPN+red	DECISION_DIAGRAMS PARALLEL_PROCESSING TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK CONSTANT_TEST CPN_APPROX DEADLOCK_TEST DECISION_DIAGRAMS EXHAUSTIVE_WALK INITIAL_STATE INVARIANTS MARKED_SUFFIX_TEST PARALLEL_PROCESSING PARIKH_WALK PROBABILISTIC_WALK QUASILIVENESS_TEST QUASI_LIVE_REVERSIBLE RANDOM_WALK SAT_SMT SCC_TEST SIPHON_TEST SKELETON_TEST SMT_REFINEMENT STRUCTURAL STRUCTURAL_REDUCTION TOPOLOGICAL TRIVIAL_UNMARKED_SCC_TEST UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK COVER_WALK CPN_APPROX DECISION_DIAGRAMS INITIAL_STATE PARALLEL_PROCESSING PARIKH_WALK RANDOM_WALK REACHABILITY_MAX REACHABILITY_MIN SAT_SMT TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK CPN_APPROX DECISION_DIAGRAMS EXHAUSTIVE_WALK INITIAL_STATE OVER_APPROXIMATION PARALLEL_PROCESSING PARIKH_WALK PROBABILISTIC_WALK RANDOM_WALK SAT_SMT SMT_REFINEMENT STRUCTURAL_REDUCTION TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK DECISION_DIAGRAMS EXHAUSTIVE_WALK INITIAL_STATE PARALLEL_PROCESSING PARIKH_WALK PROBABILISTIC_WALK RANDOM_WALK SMT_REFINEMENT TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	DECISION_DIAGRAMS INITIAL_STATE KNOWLEDGE LENGTHENING_INSENSITIVE PARALLEL_PROCESSING REACHABILITY_KNOWLEDGE SHORTENING_INSENSITIVE STACK_TEST STUTTER_TEST TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK CONSTANT_TEST COVER_WALK CPN_APPROX DEADLOCK_TEST DECISION_DIAGRAMS EXHAUSTIVE_WALK INITIAL_STATE INVARIANTS KNOWLEDGE LENGTHENING_INSENSITIVE MARKED_SUFFIX_TEST OVER_APPROXIMATION PARALLEL_PROCESSING PARIKH_WALK PROBABILISTIC_WALK QUASILIVENESS_TEST QUASI_LIVE_REVERSIBLE RANDOM_WALK REACHABILITY_KNOWLEDGE REACHABILITY_MAX REACHABILITY_MIN SAT_SMT SCC_TEST SHORTENING_INSENSITIVE SIPHON_TEST SKELETON_TEST SMT_REFINEMENT STACK_TEST STRUCTURAL STRUCTURAL_REDUCTION STUTTER_TEST TOPOLOGICAL TRIVIAL_UNMARKED_SCC_TEST UNFOLDING_TO_PT USE_NUPN
ITS-Tools	DECISION_DIAGRAMS TOPOLOGICAL USE_NUPN	BESTFIRST_WALK COLLATERAL_PROCESSING CONSTANT_TEST CPN_APPROX DEADLOCK_TEST DECISION_DIAGRAMS EXHAUSTIVE_WALK EXPLICIT INITIAL_STATE INVARIANTS LTSMIN MARKED_SUFFIX_TEST PARIKH_WALK PARTIAL_ORDER PROBABILISTIC_WALK QUASILIVENESS_TEST QUASI_LIVE_REVERSIBLE RANDOM_WALK SAT_SMT SCC_TEST SIPHON_TEST SKELETON_TEST SMT_REFINEMENT STRUCTURAL STRUCTURAL_REDUCTION TOPOLOGICAL TRIVIAL_UNMARKED_SCC_TEST USE_NUPN	BESTFIRST_WALK COVER_WALK CPN_APPROX DECISION_DIAGRAMS INITIAL_STATE PARIKH_WALK RANDOM_WALK REACHABILITY_MAX REACHABILITY_MIN SAT_SMT TOPOLOGICAL USE_NUPN	BESTFIRST_WALK COLLATERAL_PROCESSING CPN_APPROX DECISION_DIAGRAMS EXHAUSTIVE_WALK EXPLICIT INITIAL_STATE LTSMIN OVER_APPROXIMATION PARIKH_WALK PARTIAL_ORDER PROBABILISTIC_WALK RANDOM_WALK SAT_SMT SMT_REFINEMENT STRUCTURAL_REDUCTION TOPOLOGICAL USE_NUPN	BESTFIRST_WALK COLLATERAL_PROCESSING DECISION_DIAGRAMS EXHAUSTIVE_WALK INITIAL_STATE PARIKH_WALK PROBABILISTIC_WALK RANDOM_WALK SMT_REFINEMENT TOPOLOGICAL USE_NUPN	DECISION_DIAGRAMS EXPLICIT HOA INITIAL_STATE KNOWLEDGE LENGTHENING_INSENSITIVE LTSMIN PARTIAL_ORDER REACHABILITY_KNOWLEDGE SAT_SMT SHORTENING_INSENSITIVE STACK_TEST STUTTER_TEST TOPOLOGICAL USE_NUPN	BESTFIRST_WALK COLLATERAL_PROCESSING CONSTANT_TEST COVER_WALK CPN_APPROX DEADLOCK_TEST DECISION_DIAGRAMS EXHAUSTIVE_WALK EXPLICIT HOA INITIAL_STATE INVARIANTS KNOWLEDGE LENGTHENING_INSENSITIVE LTSMIN MARKED_SUFFIX_TEST OVER_APPROXIMATION PARIKH_WALK PARTIAL_ORDER PROBABILISTIC_WALK QUASILIVENESS_TEST QUASI_LIVE_REVERSIBLE RANDOM_WALK REACHABILITY_KNOWLEDGE REACHABILITY_MAX REACHABILITY_MIN SAT_SMT SCC_TEST SHORTENING_INSENSITIVE SIPHON_TEST SKELETON_TEST SMT_REFINEMENT STACK_TEST STRUCTURAL STRUCTURAL_REDUCTION STUTTER_TEST TOPOLOGICAL TRIVIAL_UNMARKED_SCC_TEST USE_NUPN
LoLA	—	COLLATERAL_PROCESSING EXPLICIT STATE_COMPRESSION STUBBORN_SETS TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	EXPLICIT SEQUENTIAL_PROCESSING STATE_COMPRESSION STUBBORN_SETS TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	COLLATERAL_PROCESSING EXPLICIT STATE_COMPRESSION STUBBORN_SETS TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	COLLATERAL_PROCESSING EXPLICIT STATE_COMPRESSION STUBBORN_SETS TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	COLLATERAL_PROCESSING EXPLICIT STATE_COMPRESSION STUBBORN_SETS TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	COLLATERAL_PROCESSING EXPLICIT SEQUENTIAL_PROCESSING STATE_COMPRESSION STUBBORN_SETS TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN
LTSMin+red	—	BESTFIRST_WALK CONSTANT_TEST CPN_APPROX DEADLOCK_TEST EXHAUSTIVE_WALK INITIAL_STATE INVARIANTS MARKED_SUFFIX_TEST PARIKH_WALK PROBABILISTIC_WALK QUASILIVENESS_TEST QUASI_LIVE_REVERSIBLE RANDOM_WALK SAT_SMT SCC_TEST SIPHON_TEST SKELETON_TEST SMT_REFINEMENT STRUCTURAL STRUCTURAL_REDUCTION TOPOLOGICAL TRIVIAL_UNMARKED_SCC_TEST	BESTFIRST_WALK COVER_WALK CPN_APPROX INITIAL_STATE PARIKH_WALK RANDOM_WALK REACHABILITY_MAX REACHABILITY_MIN SAT_SMT TOPOLOGICAL	BESTFIRST_WALK CPN_APPROX EXHAUSTIVE_WALK INITIAL_STATE OVER_APPROXIMATION PARIKH_WALK PROBABILISTIC_WALK RANDOM_WALK SAT_SMT SMT_REFINEMENT STRUCTURAL_REDUCTION TOPOLOGICAL	BESTFIRST_WALK EXHAUSTIVE_WALK INITIAL_STATE PARIKH_WALK PROBABILISTIC_WALK RANDOM_WALK SMT_REFINEMENT TOPOLOGICAL	EXPLICIT INITIAL_STATE KNOWLEDGE LENGTHENING_INSENSITIVE PARALLEL_PROCESSING REACHABILITY_KNOWLEDGE SHORTENING_INSENSITIVE STACK_TEST STUTTER_TEST TOPOLOGICAL USE_NUPN	BESTFIRST_WALK CONSTANT_TEST COVER_WALK CPN_APPROX DEADLOCK_TEST EXHAUSTIVE_WALK EXPLICIT INITIAL_STATE INVARIANTS KNOWLEDGE LENGTHENING_INSENSITIVE MARKED_SUFFIX_TEST OVER_APPROXIMATION PARALLEL_PROCESSING PARIKH_WALK PROBABILISTIC_WALK QUASILIVENESS_TEST QUASI_LIVE_REVERSIBLE RANDOM_WALK REACHABILITY_KNOWLEDGE REACHABILITY_MAX REACHABILITY_MIN SAT_SMT SCC_TEST SHORTENING_INSENSITIVE SIPHON_TEST SKELETON_TEST SMT_REFINEMENT STACK_TEST STRUCTURAL STRUCTURAL_REDUCTION STUTTER_TEST TOPOLOGICAL TRIVIAL_UNMARKED_SCC_TEST USE_NUPN
NoHD	LINEAR_EQUATIONS PARALLEL_PROCESSING UNFOLDING_TO_PT	—	—	—	—	—	LINEAR_EQUATIONS PARALLEL_PROCESSING UNFOLDING_TO_PT
smpt	—	—	—	BMC BULK COLLATERAL_PROCESSING COLORED_WALK COMPOUND DUPLICATE IMPLICIT INITIAL_MARKING NET_UNFOLDING PARIKH PDR_REACH_SATURATED PROJECTION SAT_SMT SKELETON SLICING STATE_EQUATION STRUCTURAL_REDUCTION TAUTOLOGY TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN WALK	—	—	BMC BULK COLLATERAL_PROCESSING COLORED_WALK COMPOUND DUPLICATE IMPLICIT INITIAL_MARKING K-INDUCTION NET_UNFOLDING PARIKH PDR_REACH_SATURATED PROJECTION SAT_SMT SKELETON SLICING STATE_EQUATION STRUCTURAL-REDUCTION STRUCTURAL_REDUCTION TAUTOLOGY TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN WALK
SVSKit	—	—	—	IMPLICIT QUERY_REDUCTION SEQUENTIAL_PROCESSING STATE_COMPRESSION	IMPLICIT QUERY_REDUCTION SEQUENTIAL_PROCESSING STATE_COMPRESSION	—	IMPLICIT QUERY_REDUCTION SEQUENTIAL_PROCESSING STATE_COMPRESSION
Tapaal	—	COLLATERAL_PROCESSING CPN_APPROX CTL_CZERO EXPLICIT LP_APPROX QUERY_REDUCTION SAT_SMT STATE_COMPRESSION STRUCTURAL_REDUCTION STUBBORN_SETS TRACE_ABSTRACTION_REFINEMENT UNFOLDING_TO_PT	COLLATERAL_PROCESSING CPN_APPROX EXPLICIT QUERY_REDUCTION SAT_SMT STATE_COMPRESSION STRUCTURAL_REDUCTION STUBBORN_SETS TRACE_ABSTRACTION_REFINEMENT	COLLATERAL_PROCESSING CPN_APPROX EXPLICIT LP_APPROX QUERY_REDUCTION SAT_SMT STATE_COMPRESSION STRUCTURAL_REDUCTION STUBBORN_SETS TRACE_ABSTRACTION_REFINEMENT UNFOLDING_TO_PT	COLLATERAL_PROCESSING CPN_APPROX CTL_CZERO EXPLICIT LP_APPROX QUERY_REDUCTION SAT_SMT STATE_COMPRESSION STRUCTURAL_REDUCTION STUBBORN_SETS TRACE_ABSTRACTION_REFINEMENT UNFOLDING_TO_PT	AUTOMATON_HEUR AUT_STUB COLLATERAL_PROCESSING DIST_HEUR EXPLICIT HEURISTIC LOGFIRECOUNT_HEUR(5000) LP_APPROX NDFS QUERY_REDUCTION RANDOM_HEUR SAT_SMT STATE_COMPRESSION STRUCTURAL_REDUCTION STUBBORN STUBBORN_SETS TARJAN UNFOLDING_TO_PT WEAK_SKIP	AUTOMATON_HEUR AUT_STUB COLLATERAL_PROCESSING CPN_APPROX CTL_CZERO DIST_HEUR EXPLICIT HEURISTIC LOGFIRECOUNT_HEUR(5000) LP_APPROX NDFS OPTIM-1 QUERY_REDUCTION RANDOM_HEUR SAT_SMT STATE_COMPRESSION STRUCTURAL_REDUCTION STUBBORN STUBBORN_SETS TARJAN TRACE_ABSTRACTION_REFINEMENT UNFOLDING_TO_PT WEAK_SKIP

TINA.tedd	COLLATERAL_PROCESSING DECISION_DIAGRAMS EXPLICIT LATTICE_POINTS_COUNTING LINEAR_EQUATIONS STRUCTURAL_REDUCTION TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	—	—	—	—	—	COLLATERAL_PROCESSING DECISION_DIAGRAMS EXPLICIT LATTICE_POINTS_COUNTING LINEAR_EQUATIONS STRUCTURAL_REDUCTION TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN
Reference Tools							
GreatSPN (meddly)	DECISION_DIAGRAMS PARALLEL_PROCESSING TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK CONSTANT_TEST CPN_APPROX DEADLOCK_TEST DECISION_DIAGRAMS EXHAUSTIVE_WALK INITIAL_STATE INVARIANTS MARKED_SUFFIX_TEST PARALLEL_PROCESSING PARIKH_WALK PROBABILISTIC_WALK QUASILIVENESS_TEST QUASI_LIVE_REVERSIBLE RANDOM_WALK SAT_SMT SCC_TEST SIPHON_TEST SKELETON_TEST SMT_REFINEMENT STRUCTURAL STRUCTURAL_REDUCTION TOPOLOGICAL TRIVIAL_UNMARKED_SCC_TEST UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK COVER_WALK CPN_APPROX DECISION_DIAGRAMS INITIAL_STATE PARALLEL_PROCESSING PARIKH_WALK RANDOM_WALK REACHABILITY_MAX REACHABILITY_MIN SAT_SMT TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK CPN_APPROX DECISION_DIAGRAMS EXHAUSTIVE_WALK INITIAL_STATE OVER_APPROXIMATION PARALLEL_PROCESSING PARIKH_WALK PROBABILISTIC_WALK RANDOM_WALK SAT_SMT SMT_REFINEMENT STRUCTURAL_REDUCTION TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK DECISION_DIAGRAMS EXHAUSTIVE_WALK INITIAL_STATE PARALLEL_PROCESSING PARIKH_WALK PROBABILISTIC_WALK RANDOM_WALK SMT_REFINEMENT TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	DECISION_DIAGRAMS INITIAL_STATE KNOWLEDGE LENGTHENING_INSENSITIVE PARALLEL_PROCESSING REACHABILITY_KNOWLEDGE SHORTENING_INSENSITIVE STACK_TEST STUTTER_TEST TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN	BESTFIRST_WALK CONSTANT_TEST COVER_WALK CPN_APPROX DEADLOCK_TEST DECISION_DIAGRAMS EXHAUSTIVE_WALK INITIAL_STATE INVARIANTS KNOWLEDGE LENGTHENING_INSENSITIVE MARKED_SUFFIX_TEST OVER_APPROXIMATION PARALLEL_PROCESSING PARIKH_WALK PROBABILISTIC_WALK QUASILIVENESS_TEST QUASI_LIVE_REVERSIBLE RANDOM_WALK REACHABILITY_KNOWLEDGE REACHABILITY_MAX REACHABILITY_MIN SAT_SMT SCC_TEST SHORTENING_INSENSITIVE SIPHON_TEST SKELETON_TEST SMT_REFINEMENT STACK_TEST STRUCTURAL STRUCTURAL_REDUCTION STUTTER_TEST TOPOLOGICAL TRIVIAL_UNMARKED_SCC_TEST UNFOLDING_TO_PT USE_NUPN
LTSMin	—	—	—	—	—	EXPLICIT INITIAL_STATE KNOWLEDGE LENGTHENING_INSENSITIVE PARALLEL_PROCESSING REACHABILITY_KNOWLEDGE SHORTENING_INSENSITIVE STACK_TEST STUTTER_TEST TOPOLOGICAL USE_NUPN	EXPLICIT INITIAL_STATE KNOWLEDGE LENGTHENING_INSENSITIVE PARALLEL_PROCESSING REACHABILITY_KNOWLEDGE SHORTENING_INSENSITIVE STACK_TEST STUTTER_TEST TOPOLOGICAL USE_NUPN

3. Experimental Conditions of the MCC'2024

Each tool was submitted to 23 426 executions in various conditions (1 802 model/instances and 13 examinations per model/instance) for which it could report: DNC (do not compete), CC (cannot compute) or the result of the query. These executions were handled by **BenchKit**, that was developed in the context of the MCC for massive testing of software. Then, from the raw data provided by **BenchKit**, some post-analysis scripts consolidated these and computed a ranking.

16 GB of memory were allocated to each virtual machine (both parallel and sequential tools) and a confinement of one hour was considered (execution aborted after one hour). So, a total of 281 112 runs (execution of one examination by the virtual machine) generated 84 GB of raw data (essentially log files and CSV of sampled data).

The table below shows some data about the involved machines and their contribution to the computation of these results. This year, we affected only physical cores to the virtual machines (discarding logical cores obtained from hyper-threading) so the balance between the various machine we used is quite different from he one of past years.

Involved Machines and Execution of the Benchmarks

	tall	Small	Tajo	Total
Physical Cores	15×32 @ 2.1GHz	23×12 @ 2.4GHz	96 @ 2.4GHz	—
Memory (GB)	15×384	23×64	2048	—
Used Cores (sequential tools)	15×31, 12×31 VM in //	23×3, 9×3 VM in //	95, 95 VM in //	—
Used Cores (parallel tools)	15×28 (4 per VM), 11×7 VM in //	23×8 (4 per VM), 9×2 VM in //	92 (4 per VM), 23 VM in //	—
Number of runs	169 572	89 273	22 152	**281 112**
Total CPU consumed	2 203d, 20h, 30m, 23s	1 163d, 6h, 11m, 15s	474d, 1h, 48m, 33s	**3 841d, 4h, 34m, 10s**
Total CPU	***about 10 years, 6 months and 9 days***			—
Time spent to complete benchmarks	**about 20 days**			—
Estimated boot time of VMs + management (overhead)	**about 12d (Included in total CPU) so ≅ 3.0 ‰ overhead**			—

We are pleased to thanks those who helped in the execution of tools:

- Tajo was made available by colleagues at ***Rostock University*** ,
- Tall (we used 15 nodes) and small (we used 23 nodes) are clusters at LIP6 ***Sorbonne Université & CNRS*** .

4. The Results of the MCC'2024

This First table below presents detailed results about the MCC'2024.

Details about the Examinations in the MCC (part I): Details about Results and Scoring + Model Performance Charts

	Details about Results and Scoring	Model Performance Charts	Tool Resource consumption
StateSpace			
ReachabilityDeadlock (GlobalProperties)			
QuasiLiveness (GlobalProperties)			
StableMarking (GlobalProperties)			
Liveness (GlobalProperties)			
OneSafe (GlobalProperties)			
UpperBounds			
ReachabilityCardinality			
ReachabilityFireability			
CTLCardinality			
CTLFireability			
LTLCardinality			
LTLFireability			

This Second table below presents some performance analysis related to tools during the MCC'2024.

Details about the examinations in the MCC'2024 (part II) Tool Performance Charts

	All models	«Surprise» models only	«Known» models only
Original Tools (or combinations)			
2023-Gold			
GreatSPN+red			
ITS-Tools			

LoLA			
LTSMin+red			
NoHD			
smpt			
SVSKit			
Tapaal			
tedd			
Reference Tools			
GreatSPN			
LTSMin			

You can **download the full archive** (5.3 GB compressed and 84 GB uncompressed) of the 281 112 runs processed to compute the results of the MCC'2024. This archive contains execution traces, execution logs and sampling, as well as a large CSV files that summarizes all the executions. You may get separately the two mostly interesting CSV files:

- `GlobalSummary.csv` that summarizes all results from all runs in the contest (59 MB when expanded),
- `raw-result-analysis.csv` that contains the same data as the previous one but enriched with scoring information and the expected results (computed as a majority of tools pondered by their confidence rate, 70 MB when expanded).

Note that from the two CSV file, you can identify the unique run identifier that allows you to find the traces and any information in the archive (they are also available on the web site when the too did participated).

5. The Winners for the MCC'2024

This section presents the results for the main examinations that are:

- ***State Space*** generation,
- ***UpperBounds*** computation,
- ***GlobalProperties*** computation (ReachabilityDeadlock , QuasiLiveness, StableMarking, Liveness, OneSafe),
- ***Reachability*** Formulas (ReachabilityCardinality, ReachabilityFireability),
- ***CTL*** Formulas (CTLCardinality, CTLFireability),
- ***LTL*** Formulas (LTLCardinality, LTLFireability),

To avoid a too large disparity between models with numerous instances and those with only one, a normalization was applied so that the score, for an examination and a model, varies between 102 and 221 points. Therefore, providing a correct value may brings a different number of points according to the considered model. A multiplier was applied depending to the model category:

- ***×1*** for «Known» models,
- ***×10*** for «Surprise» models (computed from rule E-4.4 that states «the total score for all “surprise” models instances weight half the score for all the instances of “known” models»).

Let us introduce two «special» tools:

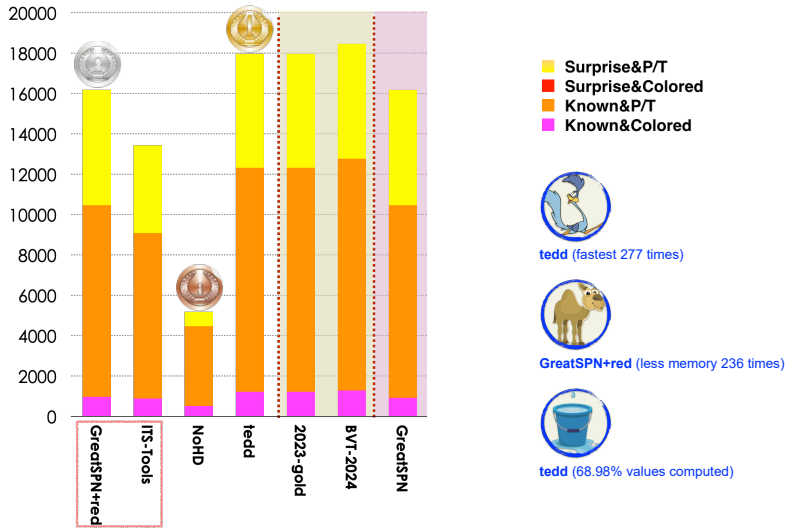
- ***2024-gold*** is an hybrid tool made of the gold-medal for the 2023 edition for each examination. It is a way to evaluate the progress of participants since the last edition of the MCC.
- ***BVT (Best Virtual Tool)*** computes the union of all the values computed by all other tools. It is very often the fastest and the tool having the smallest memory footprint, based on what the participating tool performed. It is a way to evaluate the complementarity between tools by comparing it to the gold medal

5.1. Winners in the StateSpace Category

5 tools out of 11 participated in this examination (plus 4 reference). Results based on the scoring shown below is:

- ***tedd*** ranked first (17 972 pts, 68.98% of computed values),
- ***GreatSPN+red*** ranked second (16 184 pts, 67.54% of computed values),
- ***NoHD*** ranked third (5 176 pts, 38.80% of computed values).

Then ITS-Tools got 13 439 pts (44.98% of computed values), and GreatSPN got 16 185 pts (66.98% of computed values). The the Gold-medal of 2023 collected 17 967 pts (68.92% of computed values). BVT-2024 (Best Virtual Tool) collected 18 473 pts and computed 72.38% of the total number of values in this category.



tedd (fastest 277 times)

GreatSPN+red (less memory 236 times)

tedd (68.98% values computed)

Estimated Tool Confidence rate for StateSpace (based on the «significant values» computed by tools) see section 6. for details

Tool name	Reliability	Correct Values	«significant values»
Tools competing in 2024			
GreatSPN+red	99.904%	4184	4188
ITS-Tools	100.000%	3010	3010
NoHD	100.000%	1666	1666
tedd	100.000%	4378	4378
2023-gold and BVT-2024			
2023-gold	100.000%	4380	4380
BVT-2024	100.000%	4381	4381
Reference tools			
GreatSPN	99.904%	4182	4186

Remarks about the StateSpace examination

Some detailed results state that marking ins infinite (use of the value «+Inf********»). There are some infinite models in our benchmark but our analysis tools have a constrains in the representation of very large state spaces which is the one of the Long_Long_Float Ada type (maximum value of 1.0E+4932). When

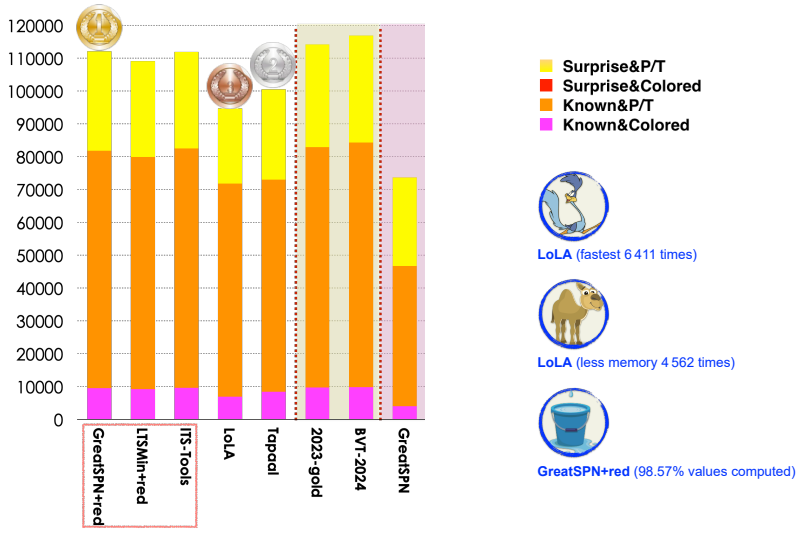
watching the execution report of some tools, you can check weather «+Inf*******» really means infinite or not.

5.2. Winners in the GlobalProperties Category

6 tools out of 11 participated in these examinations (ReachabilityDeadlock , QuasiLiveness, StableMarking, Liveness, OneSafe). Results based on the scoring shown below is:

- ***GreatSPN+red*** ranked first (112 090 pts, 98,57% of computed values),
- ***Tapaal*** ranked second (100 468 pts, 84.31% of computed values).
- ***LoLA*** ranked third (94 684 pts, 88.09% of computed values),

Then ITS-Tools got 111 888 pts (96,74% of computed values), LTSMin+red got 109 093 pts (93.64% of computed values), and GreatSPN got 73 727 pts (86.81% of computed values). The the Gold-medal of 2023 collected 114 169 pts (98,41% of computed values). BVT-2024 (Best Virtual Tool) collected 116 916 pts and computed 98,57% of the total number of values in this category.



Estimated Tool Confidence rate for GlobalPropertiesScores (based on the «significant values» computed by tools) see section 6. for details

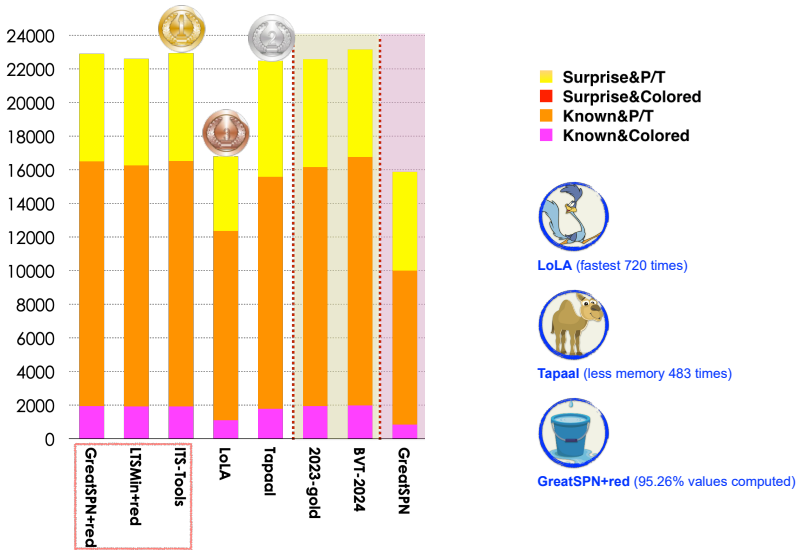
Tool name	Reliability	Correct Values	«significant values»
		Tools competing in 2024	
GreatSPN+red	100,000%	8302	8302
LTSMin+red	100,000%	8123	8123
ITS-Tools	100,000%	8342	8342
LoLA	99,807%	7239	7253
Tapaal	99,986%	7367	7368
		2023-gold and BVT-2024	
2023-gold	100,000%	8355	8355
BVT-2024	100,000%	8366	8366
		Reference tools	
GreatSPN	100,000%	4698	4698

5.3. Winners in the UpperBounds Category

6 tools out of 11 participated in this examination. Results based on the scoring shown below is:

- ***ITS-Tools*** ranked first (22 942 pts, 95.17% of computed values),
- ***Tapaal*** ranked second (22 471 pts, 89.62% of computed values),
- ***LoLA*** ranked third (16 792 pts, 75.49% of computed values).

Then GreatSPN+red got 22 910 pts (95.26% of computed values), LTSMin+red got 22 610 pts (93.69% of computed values), GreatSPN got 15 878 pts (58.20% of computed values), The the Gold-medal of 2023 collected 22 589 pts (68.92% of computed values). BVT-2024 (Best Virtual Tool) collected 23 169 pts and computed 72.38% of the total number of values in this category.



Estimated Tool Confidence rate for UpperBound (based on the «significant values» computed by tools) see section 6. for details

Tool name	Reliability	Correct Values	«significant values»

		Tools competing in 2024	
GreatSPN+red	100.000%	26456	26456
LTSMin+red	100.000%	26097	26097
ITS-Tools	100.000%	26435	26435
LoLA	98.005%	20824	21248
Tapaal	100.000%	25394	25394
		2023-gold and BVT-2024	
2023-gold	100.000%	26386	26386
BVT-2024	100.000%	26468	26468
		Reference tools	
GreatSPN	100.000%	16251	16251

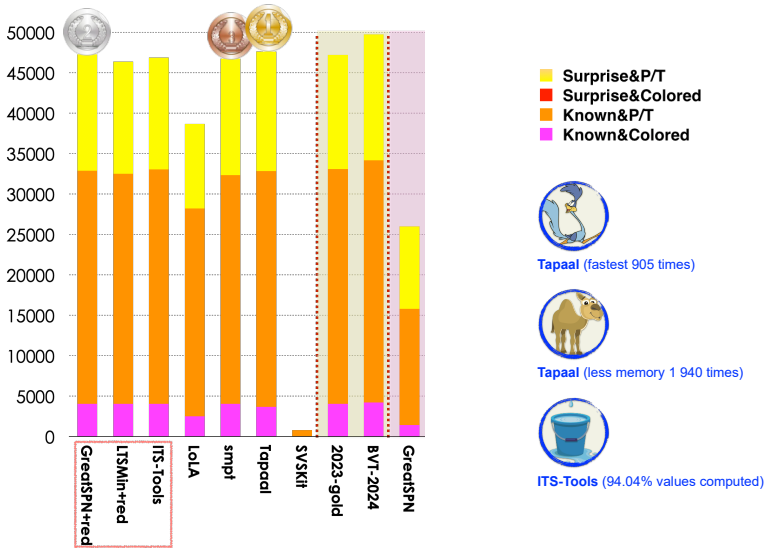
5.4. Winners in the Reachability Formulas Category

8 tools out of 11 participated in these examinations (ReachabilityCardinality and ReachabilityFireability). Results based on the scoring shown below is:

- ***Tapaal*** ranked first (47 629 pts, 93.19% of computed values),
- ***GreatSPN+red*** ranked seconf (47 277 pts, 93.80% of computed values),
- ***smtp*** ranked third (46 763 pts, 92.15% of computed values).

Then ITS-Tools got 46 877 pts (94.04% of computed values), LTSMin+red got 46 385 pts (92.49% of computed values), LoLA got 38 681 pts (80.20% of computed values), SVSKit got 801 pts (4.80% of computed values), and GreatSPN got 25 981 pts (45.61% of computed values). The the Gold-medal of 2023 collected 47 227 pts (94.56% of computed values). BVT-2024 (Best Virtual Tool) collected 49 746 pts and computed 97.76% of the total number of values in this category.

Note that SVSKit only computes Fireability formulas.



Estimated Tool Confidence rate for Reachability (based on the «significant values» computed by tools) see section 6. for details			
	Tools competing in 2024		
GreatSPN+red	100,000%	53096	53096
LTSMin+red	100,000%	52457	52457
ITS-Tools	100,000%	53346	53346
LoLA	99,782%	45309	45408
smpt	100,000%	52018	52018
Tapaal	100,000%	52504	52504
SVSKit	97,590%	1296	1328
	2023-gold and BVT-2024		
2023-gold	99,989%	53624	53630
BVT-2024	100,000%	54201	54201
	Reference tools		
GreatSPN	100,000%	25161	25161

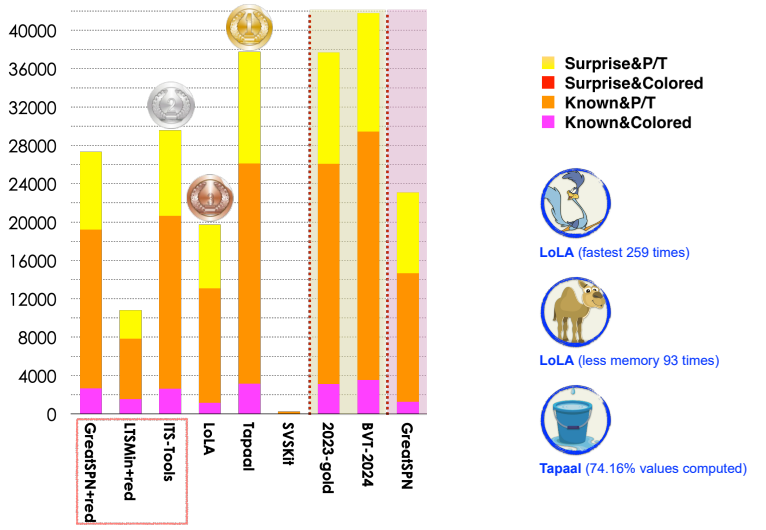
5.5. Winners in the CTL Formulas Category

7 tools out of 11 participated in these examinations (CTLCardinality and CTLFireability). Results based on the scoring shown below is:

- ***Tapaal*** ranked first (37 761 pts, 74.16% of computed values),
- ***ITS-Tools*** ranked second (29 573 pts, 57.37% of computed values),
- ***LoLA*** ranked third (19 733 pts, 38.19% of computed values).

Then GreatSPN+red got 27 353 pts (54.73% of computed values), LTSMin+red got 10 817 pts (23.03% of computed values), SVSKit got 257 pts (3.39% of computed values), and GreatSPN got 22 982 pts (43.00% of computed values). The the Gold-medal of 2023 collected 37 704 pts (74.02% of computed values). BVT-2024 (Best Virtual Tool) collected 41 785 pts and computed 83.08% of the total number of values in this category.

Note that SVSKit only computes Fireability formulas.



Estimated Tool Confidence rate for CTL (based on the «significant values» computed by tools) see section 6. for details

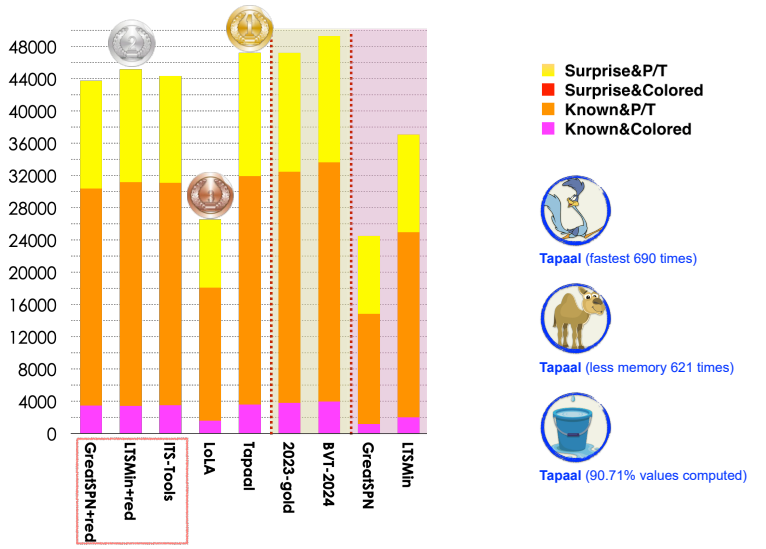
Tool name	Reliability	Correct Values	«significant values»
		Tools competing in 2024	
GreatSPN+red	99,996%	26711	26712
LTSMin+red	100,000%	12160	12160
ITS-Tools	99,996%	28213	28214
LoLA	99,942%	20818	20830
Tapaal	99,994%	32142	32144
SVSKit	98,465%	385	391
		2023-gold and BVT-2024	
2023-gold	99,994%	32104	32106
BVT-2024	100,000%	32428	32428
		Reference tools	
GreatSPN	100,000%	20189	20189

5.6. Winners in the LTL Formulas Category

7 tools out of 11 participated in these examinations (LTLCardinality and LTLFireability). Results based on the scoring shown below is:

- ***Tapaal*** ranked first (47 229 pts, 90.10% of computed values),
- ***LTSMin+red*** ranked second (45 147 pts, 90.71% of computed values),
- ***LoLA*** ranked third (26 551 pts, 52.22% of computed values).

Then ITS-Tools got 45 147 pts (88.37% of computed values), GreatSPN+red got 43 771 pts (86.84% of computed values), LTSMin got 37 077 pts (75.58% of computed values), and GreatSPN got 24 484 pts (45.04% of computed values). The the Gold-medal of 2023 collected 47 217 pts (92.54% of computed values). BVT-2024 (Best Virtual Tool) collected 49 305 pts and computed 95.67% of the total number of values in this category.



Estimated Tool Confidence rate for LTL (based on the «significant values» computed by tools) see section 6. for details

	Tools competing in 2024		
GreatSPN+red	99,990%	49124	49129
LTSMin+red	99,846%	50416	50494
ITS-Tools	99,998%	49857	49858
LoLA	99,949%	29139	29154
Tapaal	100,000%	49403	49403
	2023-gold and BVT-2024		
2023-gold	99,986%	51280	51287
BVT-2024	100,000%	51416	51416
	Reference tools		
GreatSPN	98,057%	24230	24710

LTSMin	98,588%	41480	42074

6. Estimation of the Global Tool Confidence

A confidence analysis enforces the computation of «right results» based on the answers of participating tools. To do so, we considered each value provided in the contest (a value is a partial result such as the result of a formula or a number provided for state space, bound computation, etc.). To do so, we processed as follows:

1. For each «line» (all tools for a given examination for a given instance), we selected all «significant values» where at least 3 tools do agree.
2. Based on this subset of values, we computed the ratio between the selected values for the tool and the number of good answers hey provide for such values. This ratio gave us a tool confidence rate that is provided in the table below.
3. This tool confidence rate rate was then applied to compute the scores presented in the dedicated section.

The table below provides, in first column, the computed confidence rates (that are naturally lower for tools where a bug was detected). Then, the table provides the number of correct results (column 2) out of the number of «significant values» selected for the tool (column 3). The last column shows the number of examinations (and their type) the tool was involved in.

Estimated Tool Confidence rate (based on the «significant values» computed by tools)

Tool name	Reliability	Correct Values	«significant values»	Involved Examinations
				Tools competing in 2024
GreatSPN+red	99.994%	167 873	167 883	**13** StateSpace, UpperBounds, CTLCardinality, CTLFireability, LTLCardinality, LTLFireability, ReachabilityCardinality, ReachabilityFireability, ReachabilityDeadlock, QuasiLiveness, StableMarking, Liveness, OneSafe
LTSMin+red	99.948%	149 253	149 331	**12** UpperBounds, CTLCardinality, CTLFireability, LTLCardinality, LTLFireability, ReachabilityCardinality, ReachabilityFireability, ReachabilityDeadlock, QuasiLiveness, StableMarking, Liveness, OneSafe
ITS-Tools	99.999%	169 203	169 205	**13** StateSpace, UpperBounds, CTLCardinality, CTLFireability, LTLCardinality, LTLFireability, ReachabilityCardinality, ReachabilityFireability, ReachabilityDeadlock, QuasiLiveness, StableMarking, Liveness, OneSafe
LoLA	99.545%	123 329	123 893	**12** UpperBounds, CTLCardinality, CTLFireability, LTLCardinality, LTLFireability, ReachabilityCardinality, ReachabilityFireability, ReachabilityDeadlock, QuasiLiveness, StableMarking, Liveness, OneSafe
NoHD	100.000%	1 666	1 666	**1** StateSpace
smpt	100.000%	52 018	52 018	**2** LTLCardinality, LTLFireability
Tapaal	99.998%	166 810	166 813	**12** UpperBounds, CTLCardinality, CTLFireability, LTLCardinality, LTLFireability, ReachabilityCardinality, ReachabilityFireability, ReachabilityDeadlock, QuasiLiveness, StableMarking, Liveness, OneSafe
tedd	100.000%	4 378	4 378	**1** StateSpace
SVSKit	97.789%	1 681	1 719	**2** CTLFireability, ReachabilityFireability
				2023-gold and BVT-2024
2023-gold	99.991%	176 129	176 144	**13** StateSpace, UpperBounds, CTLCardinality, CTLFireability, LTLCardinality, LTLFireability, ReachabilityCardinality, ReachabilityFireability, ReachabilityDeadlock, QuasiLiveness, StableMarking, Liveness, OneSafe
BVT-2024	100.000%	177 260	177 260	**13** StateSpace, UpperBounds, CTLCardinality, CTLFireability, LTLCardinality, LTLFireability, ReachabilityCardinality, ReachabilityFireability, ReachabilityDeadlock, QuasiLiveness, StableMarking, Liveness, OneSafe
				Reference tools
GreatSPN	99.492%	94 711	95 195	**13** StateSpace, UpperBounds, CTLCardinality, CTLFireability, LTLCardinality, LTLFireability, ReachabilityCardinality, ReachabilityFireability, ReachabilityDeadlock, QuasiLiveness, StableMarking, Liveness, OneSafe
LTSMin	98.588%	41 480	42 074	**2** LTLCardinality, LTLFireability