



SORBONNE  
UNIVERSITÉ

# Model Checking Contest results for 2024

Fabrice Kordon — LIP6, Sorbonne Université, France



# MCC in a nutshell

Every year since 2011  
2024 is the 14<sup>th</sup> edition



- Qualification phase
- Execution phase
- Consolidation phase
- Construction of results

Home Calls Models Results Publications Committees Rules Previous Editions

## Model Checking Contest 2024

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

Complete Results for the 2024 Edition of the Model Checking Contest

Last Updated June 16, 2024

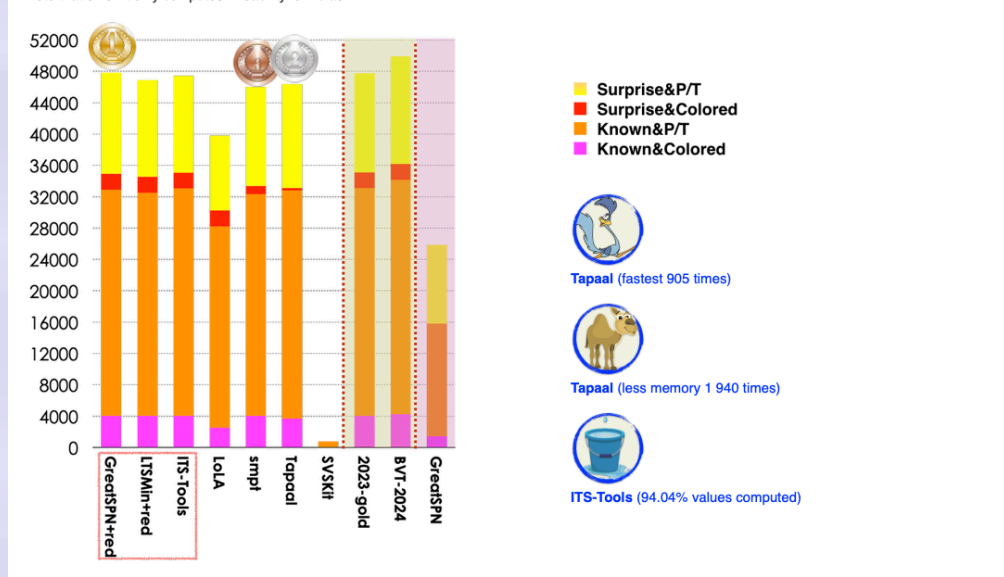
### 5.4. Winners in the Reachability Formulas Category

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

- **LTSMin+red** ranked first (47 826 pts, 93.80% of computed values),
- **Tapaa!** ranked second (46 358 pts, 93.19% of computed values),
- **emp** ranked third (46 015 pts, 92.15% of computed values).

Then ITS-Tools got 47 441 pts (94.04% of computed values), LTSMin+red got 46 871 pts (92.49% of computed values), LoLA got 46 871 pts (92.49% of computed values), SVSKit got 8 011 pts (4.80% of computed values), and GreatSPN got 25 880 pts (45.61% of computed values). The total number of values computed was 51 100. The total number of values collected was 47 766 pts (94.56% of computed values). BVT-2024 (Best Virtual Tool) collected 49 906 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%	53736	53736
LTSMin+red	100,000%	53097	53097
ITS-Tools	100,000%	53986	53986
LoLA	99,785%	45949	46048



# What is performed by tools

## Answering 6 classes of problems...

<b>Examination</b>	<b>Family</b>
StateSpace	<b>StateSpace</b>
ReachabilityDeadlock	<b>GlobalProperties</b>
QuasiLiveness	
StableMarking	
Liveness	
OneSafe	
UpperBounds	<b>UpperBounds</b>
ReachabilityCardinality	<b>Reachability</b>
ReachabilityFireability	
CTLCardinality	<b>CTL</b>
CTLFireability	
LTLCardinality	<b>LTL</b>
LTLFireability	

... on 137 models with a total of 1802 instances

**23 426 queries (1 to 16 values to be returned per query)**




# 5 surprise models in 2024

 **BloksWorld** (20 instances colored + 20 instances P/T)

 Sebastian Lassen & Henrik Ginnerup

 **CO4** (21 instances P/T)


 Quentin Nivon & Hubert Garavel

 **FireWire** (20 instances P/T)

 Hubert Garavel

 **MedleyA** (23 instances P/T)

 Hubert Garavel & Quentin Nivon

 **Reslolation** (20 instances P/T)

 Wendelin Serwe & Hubert Garavel






124 total instances to be processed by tools

We need new models for the next editions

P/T, Symmetric Nets  
Possibly parameterised



# Participating tools in 2024

tool	PT	Col	who	place	Country
GreatSPN+red	✓	✓	Y. Thierry-Mieg	Univ. Torino & Sorbonne Université	
ITS-Tools	✓	✓	Y. Thierry-Mieg	Sorbonne Université	
LoLA	✓	✓	K. Wolf	Univ. Rostock	
LTSMin+red	✓	✓	Y. Thierry-Mieg	Univ. Twente & Sorbonne Université	
NoHD	✓	✓	B. Smith	No affiliation	
smpt	✓	✓	N. Amat	LAAS-CNRS & IMDEA	
SVSKit	✓	✓	D. Morard	Univ. Geneva	
Tapaal	✓	✓	J. Srba	Aalborg University	
TINA.tedd	✓	✓	B. Berthomieu	LAAS-CNRS	
GreatSPN	✓	✓	Y. Thierry-Mieg	Univ. Torino (+ wrapping)	
LTSMin	✓	✓	Y. Thierry-Mieg	Univ. Twente (+ wrapping)	
Gold-2023	✓	✓	F. Kordon (assembling)	ITS-Tools/LoLa/Tapaal/tedd	
BVT	✓	✓	F. Kordon (assembling)	artificial tool	-



## 🌐 Generic ratio surprise models / known models

- 📌 Surprise models score should weight approx 1/2 of known models
- 📌 ~ ×10 (exactly 10.7 but rounded to 10)

## 🌐 Reuse of «difficult formulas» between 2023 and 2024

- 📌 Difficulté = less than 3 tools could cope with it
- 📌 😞 no time for detailed analysis yet
  - ▶ check if tools could cope with these formulas
  - ▶ check if this increased the overall difficulty

## 🌐 A new set of formulas

- 📌 201824 formulas (UpperBound, reachability, CTL, LTL)
  - ▶ including the some reused from 2023 (specifically tagged)
  - ▶ CTL, LTL, Reachability only



# Execution in 2024

 **Thank you!!!**

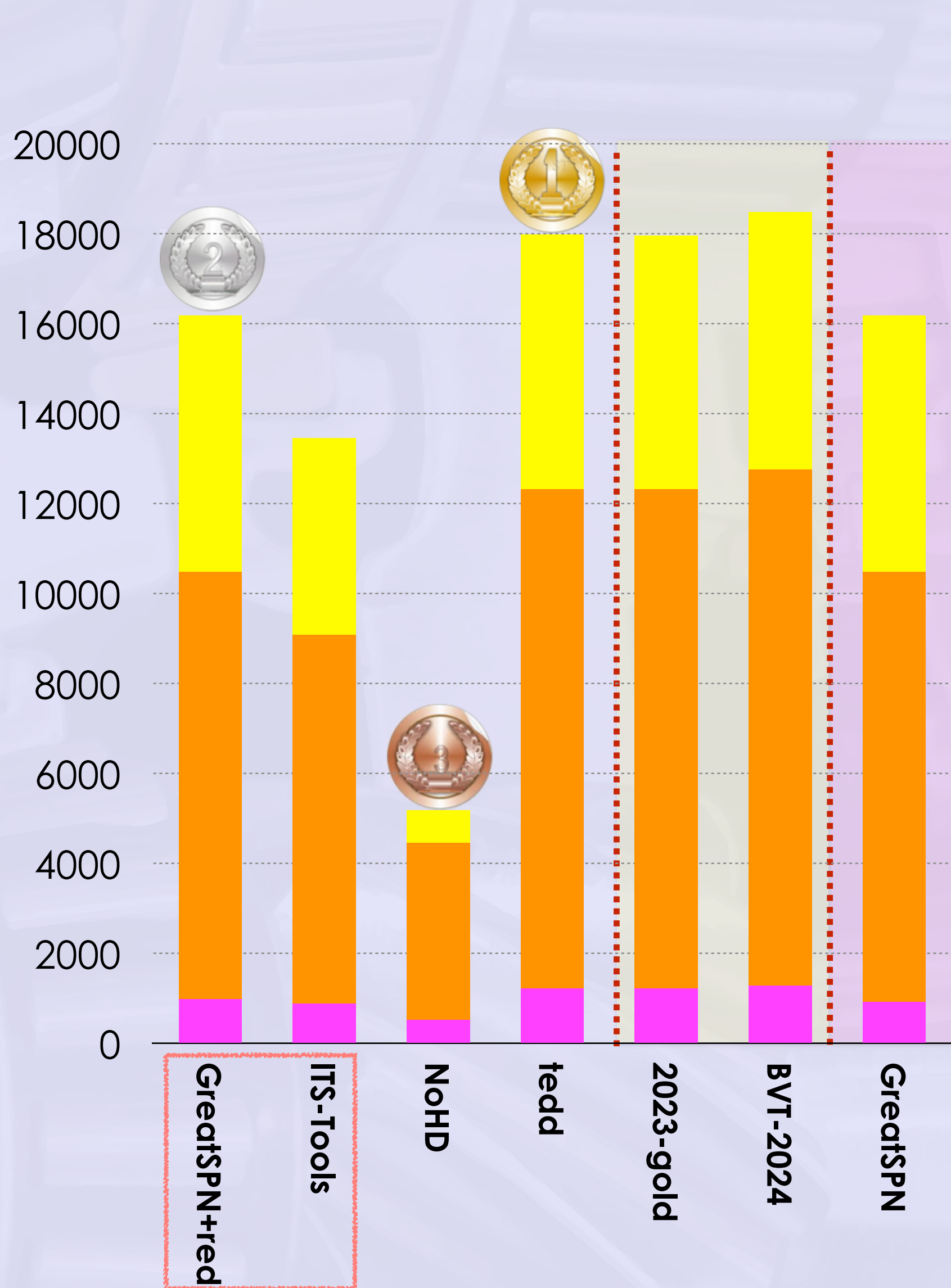
 Sorbonne Université & University of Rostock

## Involved Machines and Execution of the Benchmarks

	<b>tall</b>	<b>Small</b>	<b>Tajo</b>	<b>Total</b>
Physical Cores	15x32 @ 2.1GHz	23x12 @ 2.4GHz	96 @ 2.4GHz	—
Memory (GB)	15x384	23x64	2048	—
Used Cores (sequential tools)	15x31, 12x31 VM in //	23x3, 9x3 VM in //	95, 95 VM in //	—
Used Cores (parallel tools)	15x28 (4 per VM), 11x7 VM in //	23x8 (4 per VM), 9x2 VM in //	92 (4 per VM), 23 VM in //	—
Number of runs	169 572	89 273	22 152	<b>281 112</b>
Total CPU consumed	2 203d, 20h, 30m, 23s	1 163d, 6h, 11m, 15s	474d, 1h, 48m, 33s	<b>3 841d, 4h, 34m, 10s</b>
Total CPU	<b><i>about 10 years, 6 months and 9 days</i></b>			—
Time spent to complete benchmarks	<b>about 20 days</b>			—
Estimated boot time of VMs + management (overhead)	<b>about 12d (Included in total CPU) so <math>\approx</math> 3.0 % overhead</b>			—



# StateSpace Examination



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored



**tedd**  
273 times



**tedd-c**  
68.98%

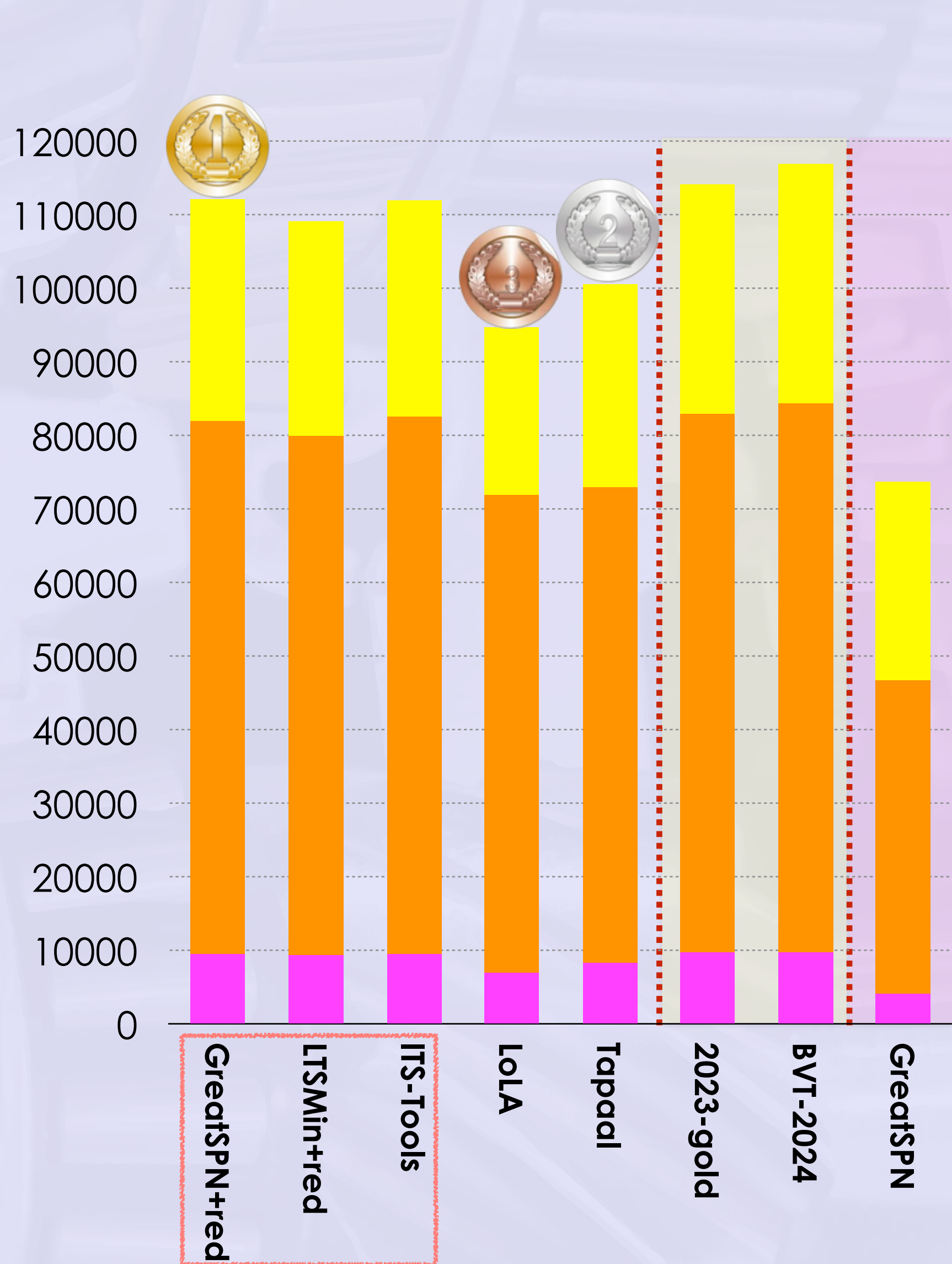


**GreatSPN+red**  
221 times

	confidence	success	computed
GreatSPN+red	99.904%	4184	4188
ITS-Tools	100.000%	3010	3010
NoHD	100.000%	1666	1666
tedd	100.000%	4378	4378
2023-gold	100.000%	4380	4380
BVT-2024	100.000%	4381	4381
GreatSPN	99.904%	4182	4186



# Global Properties Examination



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored



**LoLA**  
6 647 times



**GreatSPN+red**  
98.57%

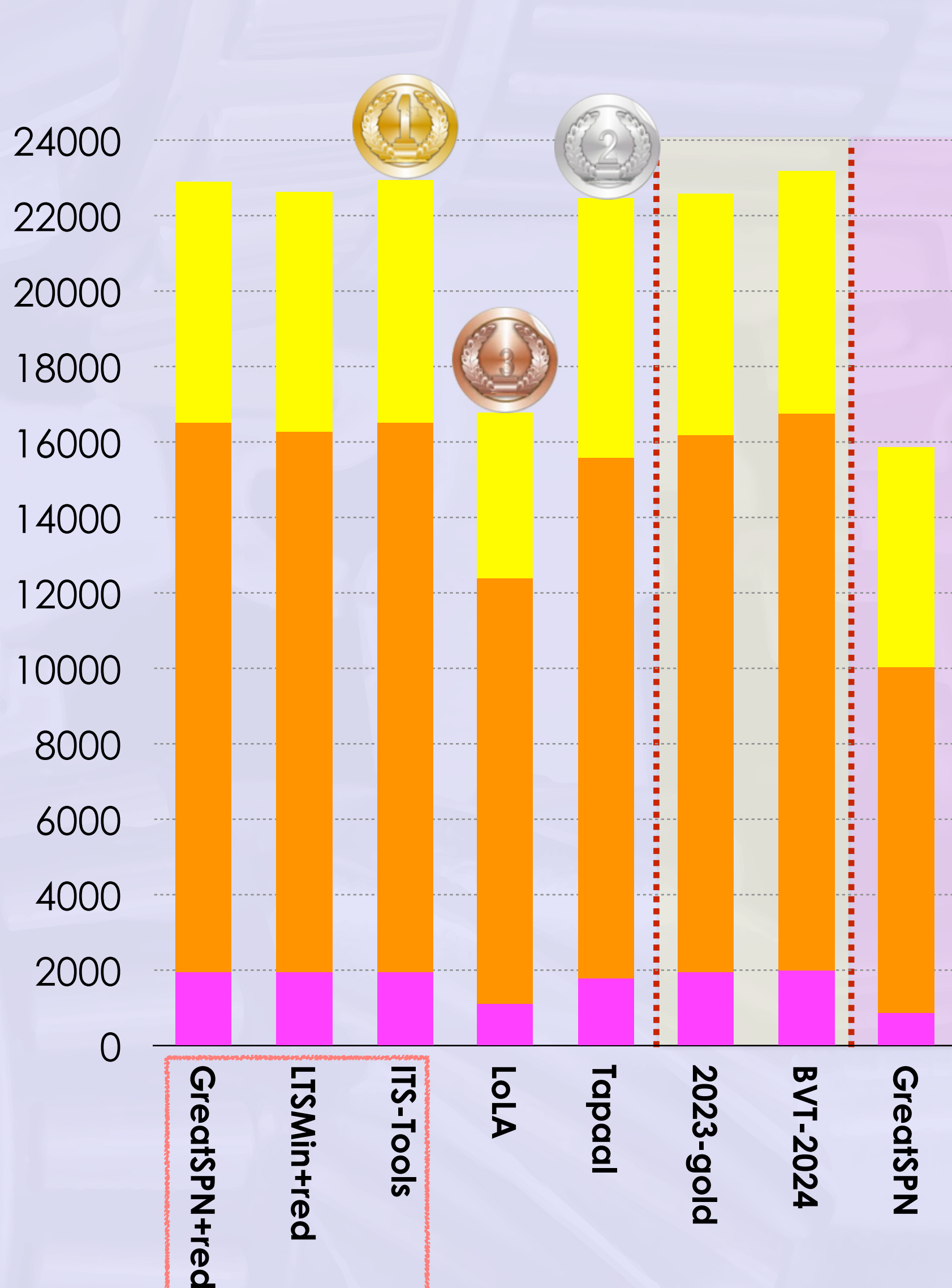


**LoLA**  
4 515 times

	confidence	success	computed
<b>GreatSPN+red</b>	<b>100,000 %</b>	8302	8302
<b>LTSMIn+red</b>	<b>100,000 %</b>	8123	8123
<b>ITS-Tools</b>	<b>100,000 %</b>	8342	8342
<b>LoLA</b>	<b>99,807 %</b>	7239	7253
<b>Tappaal</b>	<b>99,986 %</b>	7367	7368
<b>2023-gold</b>	<b>100,000 %</b>	8355	8355
<b>BVT-2024</b>	<b>100,000 %</b>	8366	8366
<b>GreatSPN</b>	<b>100,000 %</b>	4698	4698



# UpperBounds Examination



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored



**LoLA**  
720 times



**GreatSPN+red**  
95.26%

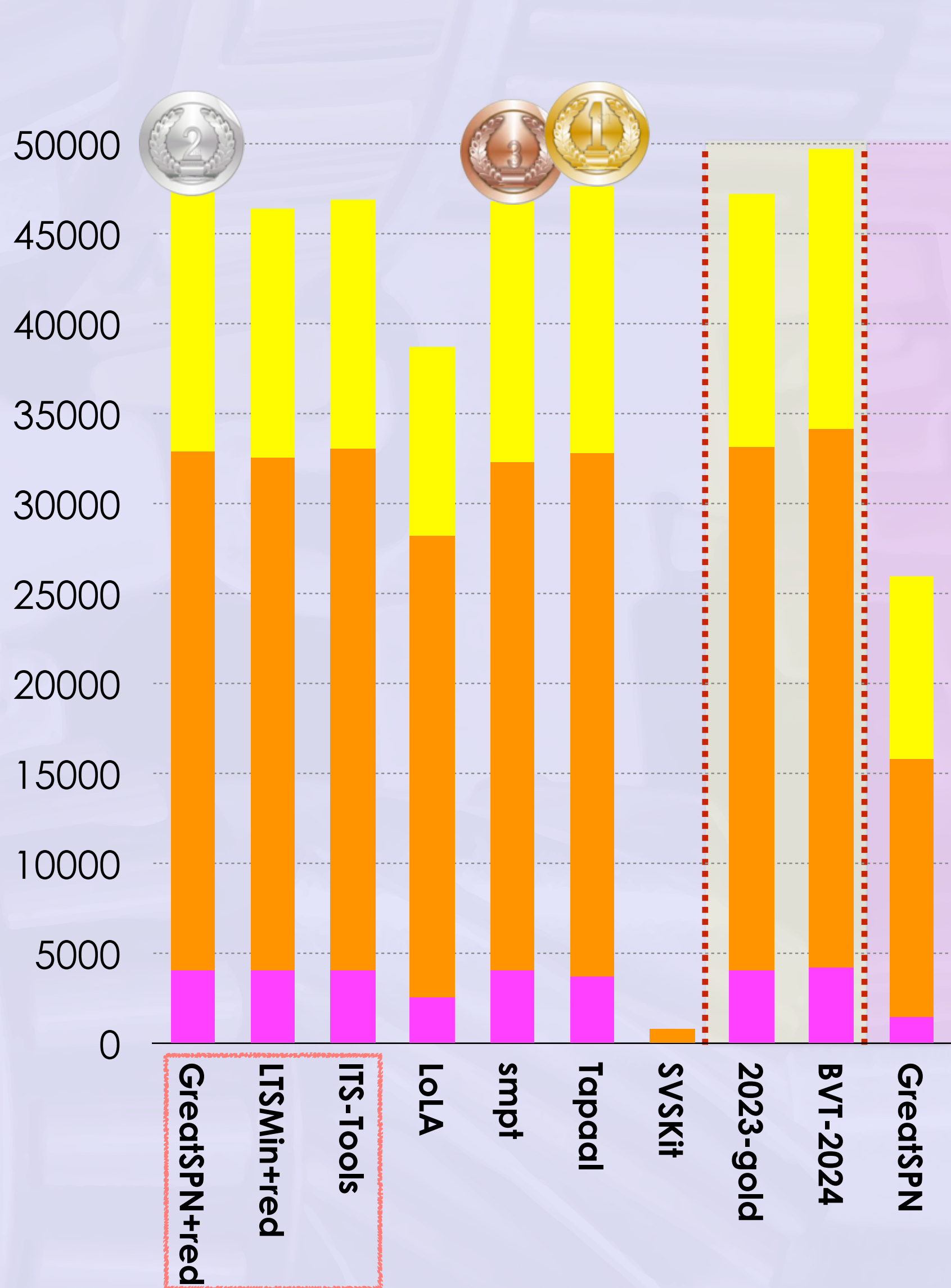


**Tappaal**  
483 times

	confidence	success	computed
<b>GreatSPN+red</b>	<b>100.000%</b>	26456	26456
<b>LTSMIn+red</b>	<b>100.000%</b>	26097	26097
<b>ITS-Tools</b>	<b>100.000%</b>	26435	26435
<b>LoLA</b>	<b>98.005%</b>	20824	21248
<b>Tappaal</b>	<b>100.000%</b>	25394	25394
<b>2023-gold</b>	<b>100.000%</b>	26386	26386
<b>BVT-2024</b>	<b>100.000%</b>	26468	26468
<b>GreatSPN</b>	<b>100.000%</b>	16251	16251



# Reachability Examination



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored



**Tapaal**  
905 times



**ITS-Tools**  
94.04%

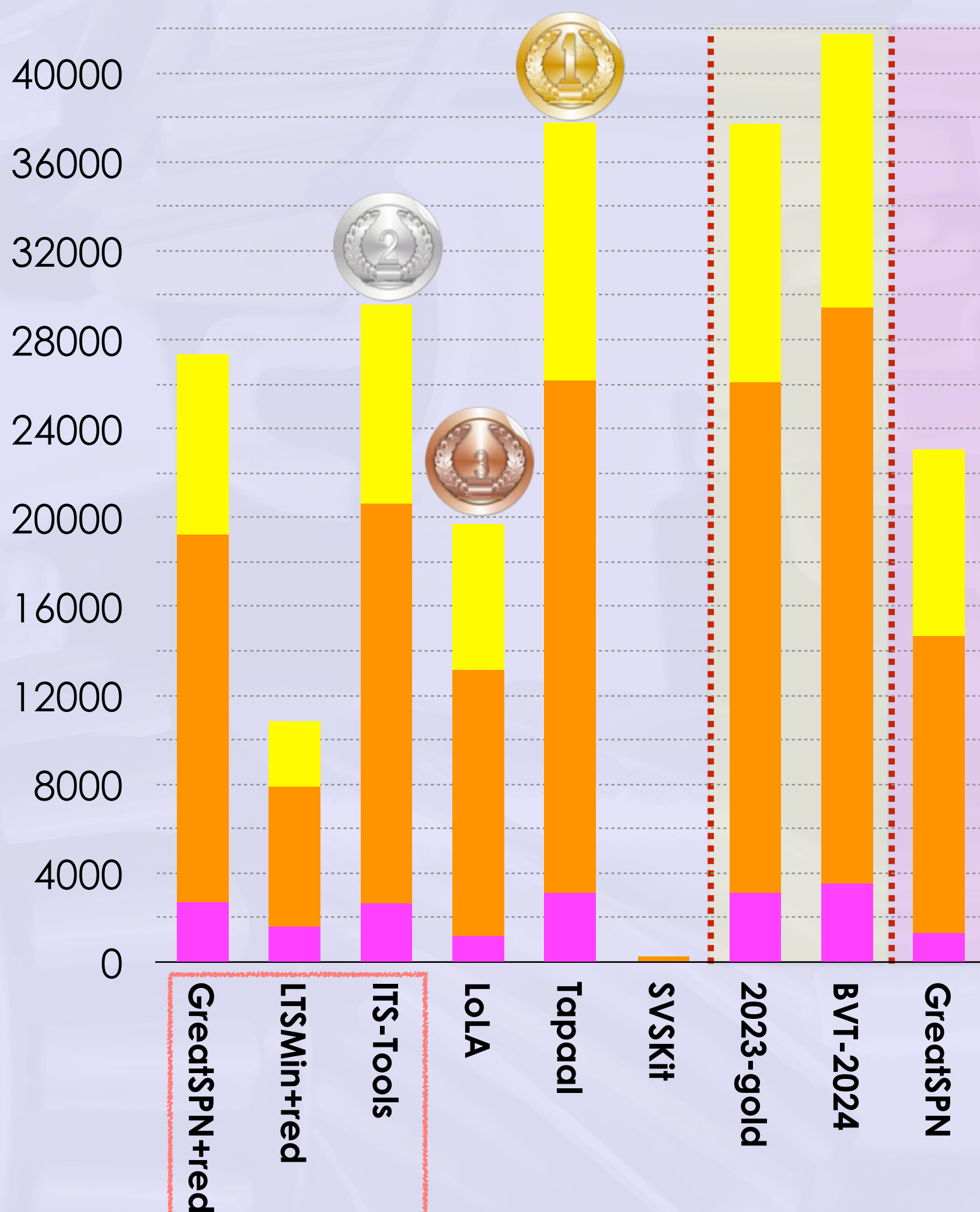


**Tapaal**  
1940 times

	confidence	success	computed
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	99,989 %	53624	53630
BVT-2024	100,000 %	54201	54201
GreatSPN	100,000 %	25161	25161



# CTL Examination



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored



**LoLA**  
259 times



**Tapaal**  
74.16%

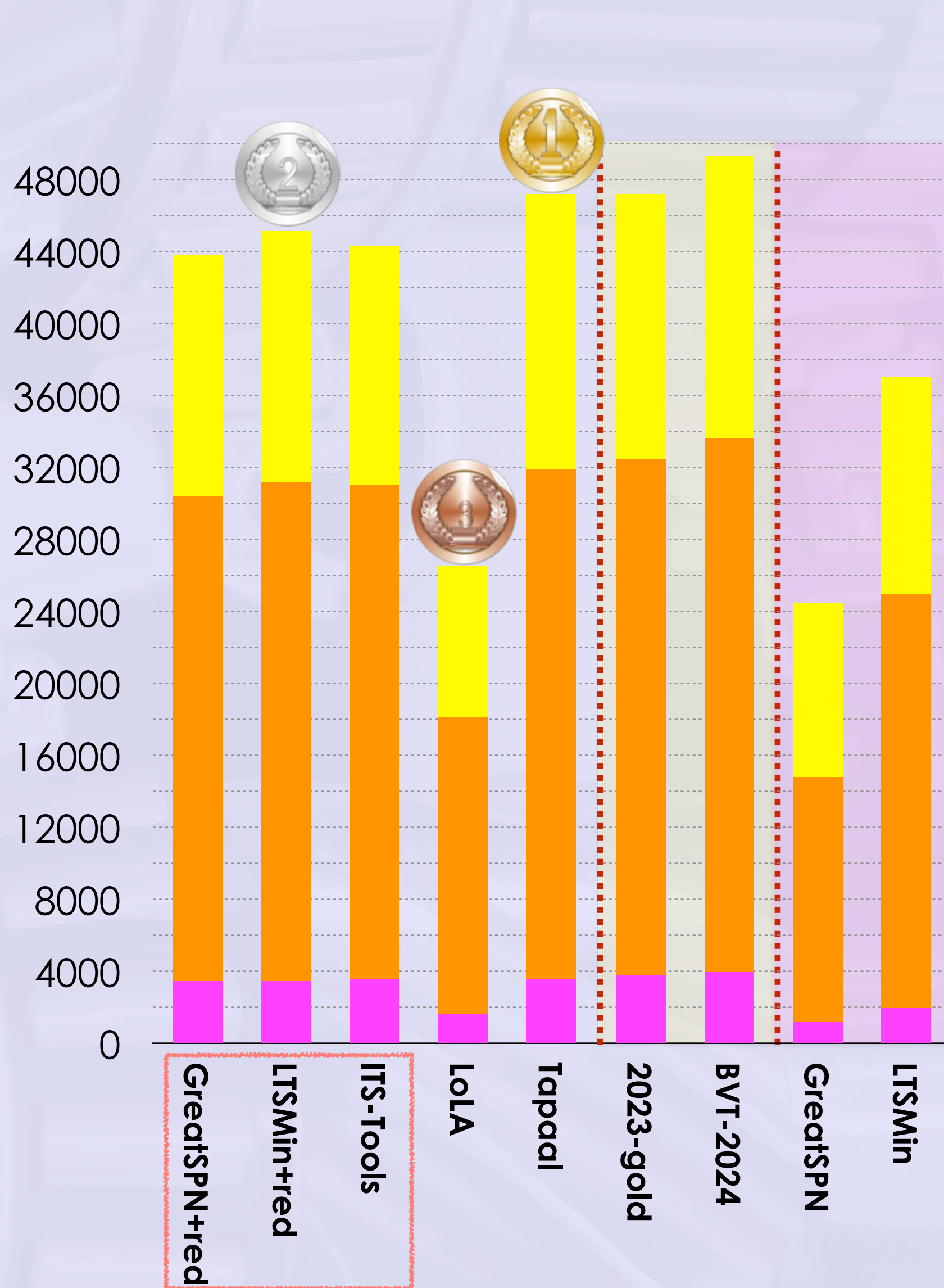


**GreatSPN+red**  
93 times

	confidence	success	computed
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	99,994 %	32104	32106
BVT-2024	100,000 %	32428	32428
GreatSPN	100,000 %	20189	20189



# LTL Examination



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored



**Tapaal**  
690 times



**Tapaal**  
90.71%



**Tapaal**  
621 times

	confidence	success	computed
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	99,986 %	51280	51287
BVT-2024	100,000 %	51416	51416
GreatSPN	98,057 %	24230	24710
LTSMIn	98,588 %	41480	42074



# Concluding remarks

## Details soon available online

- <https://mcc.lip6.fr>
- 177 950 charts (execution, comparisons, etc.)
- 195 582 web pages (exécution report, tool comparaison, etc.)

They are displayed in the table below, together with scores (total, sum per category of model and sum per model).

Summary of Results for ReachabilityCardinality									
	GreatSPN+red	LTSMIn+red	ITS-Tools	LoLA	smpt	Tapaal	2023-gold	BVT-2024	GreatSPN
Total Points	24264.48	23729.38	24003.73	20737.01	23195.09	23916.01	24146.98	25026.98	12949.22

All «Surprise» models									
	GreatSPN+red	LTSMIn+red	ITS-Tools	LoLA	smpt	Tapaal	2023-gold	BVT-2024	GreatSPN
Total Points	7642.20	7234.50	7321.70	6079.40	6933.70	7194.40	7433.30	7932.40	4975.30
fastest	3	7	20	18	4	43	3	16	4
less memory	2	1	0	20	2	67	1	21	4

BlocksWorld — Colored (1010.00 pts max)									
	GreatSPN+red	LTSMIn+red	ITS-Tools	LoLA	smpt	Tapaal	2023-gold	BVT-2024	GreatSPN
Score	1010.00	1010.00	1010.00	1010.00	514.59	410.59	1010.00	1010.00	0.00
fastest	0	3	0	17	0	0	0	0	0
smallest memory	0	0	0	20	0	0	0	0	0

01	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	DNF
02	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	DNF
03	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	DNF
04	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	DNF
05	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	DNF
06	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	DNF
07	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	DNF
08	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	DNF
09	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	TTTTFFTTFTTTF	DNF



# Concluding remarks

## Details soon available online

- <https://mcc.lip6.fr>
- 177 950 charts (execution, comparisons, etc.)
- 195 582 web pages (exécution report, tool comparaison, etc.)

They are displayed in the table below, together with scores (total, sum per category of model and sum per model).

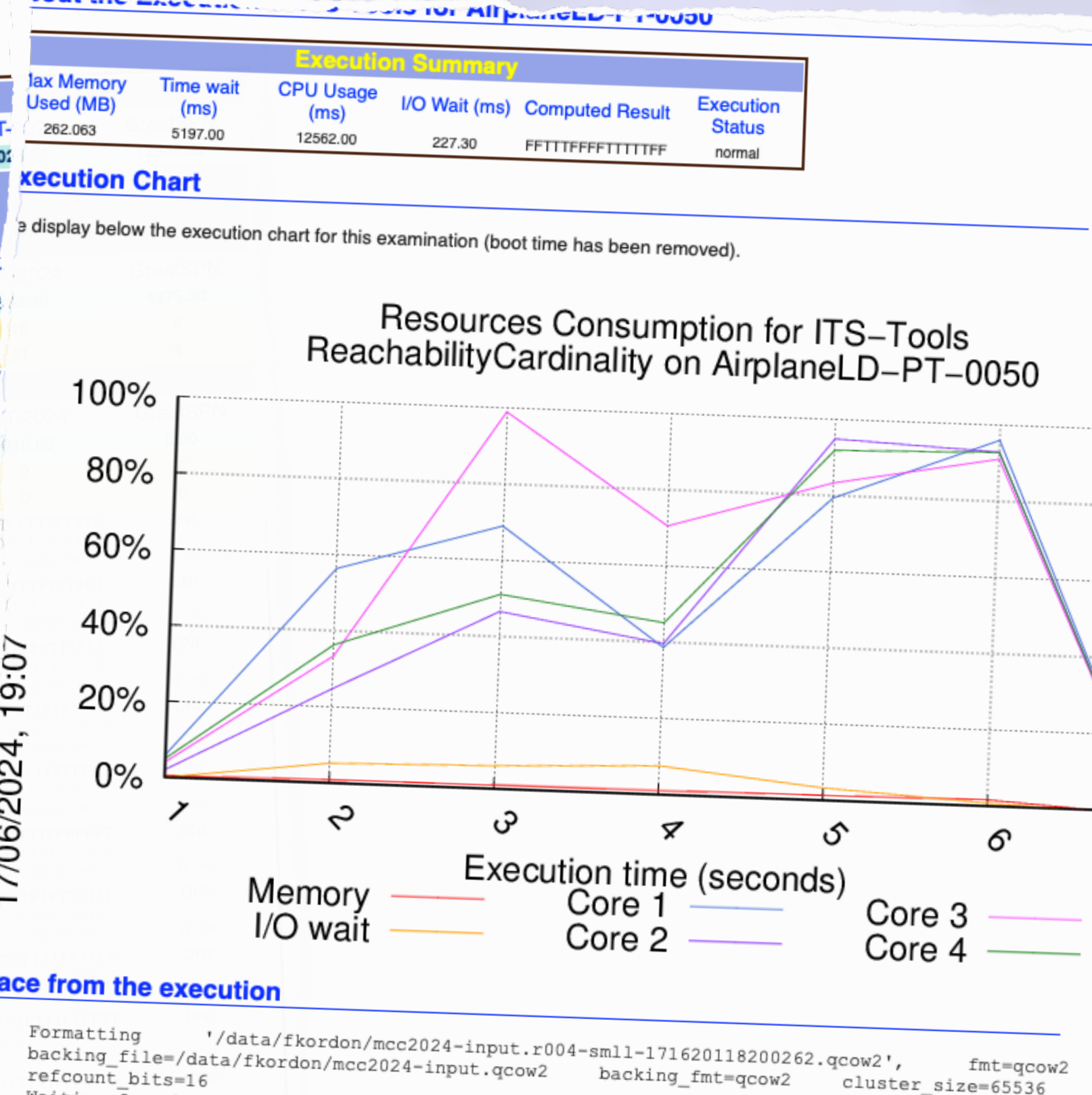
Summary of Results for ReachabilityCardinality							
	GreatSPN+red	LTSMin+red	ITS-Tools	LoLA	smpt	Tapaal	2023-gold
Total Points	24264.48	23729.38	24003.73	20737.01	23195.09	23916.01	24146.98

All «Surprise» models							
	GreatSPN+red	LTSMin+red	ITS-Tools	LoLA	smpt	Tapaal	2023-gold
Total Points	7642.20	7234.50	7321.70	6079.40	6933.70	7194.40	7433.30
fastest	3	7	20	18	4	43	3
less memory	2	1	0	20	2	67	1

BlocksWorld — Colored (1010.00 pts max)							
	GreatSPN+red	LTSMin+red	ITS-Tools	LoLA	smpt	Tapaal	2023-gold
Score	1010.00	1010.00	1010.00	1010.00	514.59	410.59	1010.00
fastest	0	3	0	17	0	0	0
smallest memory	0	0	0	20	0	0	0





# Concluding remarks

## Details soon available online

- <https://mcc.lip6.fr>
- 177 950 charts (execution, comparisons, etc.)
- 195 582 web pages (exécution report, tool comparaison, etc.)

They are displayed in the table below, together with scores (total, sum per category of model and sum per model).

Summary of Results for ReachabilityCardinality			
	GreatSPN+red	LTSMIn+red	ITS-Tools
Total Points	24264.48	23729.38	24003.73

Statistics on the executions			
	Tapaal	GreatSPN+red	Both tools
All computed OK	3	21	1397
Tapaal = GreatSPN+red	—	—	47
Tapaal > GreatSPN+red	—	—	218
Tapaal < GreatSPN+red	—	—	116
Do not compete	0	0	0
Error detected	20	0	0
Cannot Compute + Time-out	1	3	0

Execution Summary			
Max Memory Used (MB)	Time wait (ms)	CPU Usage (ms)	I/O Wait (ms)
262.063	5197.00	12562.00	227.30

On the chart below, ● denote cases where the two tools did computed all results without error, ◆ denote cases where the two tool did computed the same number of values (but not al values in the examination), ▲ denote cases where Tapaal computed more values than GreatSPN+red, ▼ denote cases where Tapaal computed less values than GreatSPN+red, ◆ denote the cases where at least one tool did not competed, ○ denote the cases where at least one tool computed a bad value and □ denote the cases where at least one tool stated it could not compute a result or timed-out.

Tapaal wins when points are below the diagonal, GreatSPN+red wins when points are above the diagonal.

Max memory, Tapaal versus GreatSPN+red for ReachabilityCardinality (All models)

Execution time, Tapaal versus GreatSPN+red for ReachabilityCardinality (All models)

Description for ITS-Tools on AirplaneLD-PT-0050



# Concluding remarks

## Details soon available online

- <https://mcc.lip6.fr>
- 177 950 charts (execution, comparisons, etc.)
- 195 582 web pages (exécution report, tool comparaison, etc.)

The collage displays various performance and execution data:

- Summary of Results for Reachability Cardinality:** A table comparing tools: GreatSPN+red (24264.48), LTSMin+red (23729.38), ITS-Tools (24003.73), LoLA, smpt, Tapaal (24146.98), 2023-gold, and BVT- (250).
- Execution Summary:** A table with columns: Max Memory Used (MB) [262.063], Time wait (ms) [5197.00], CPU Usage (ms) [12562.00], I/O Wait (ms) [227.30], Computed Result [FFTTTTTTTTTTTT], and Execution Status [normal].
- Statistics on the executions:** A table comparing Tapaal and GreatSPN+red for metrics like Smallest Memory Footprint (612 vs 190) and Shortest Execution Time (772 vs 730).
- Evolution of Memory:** A line chart titled "Memory Usage to Process Surprise Models (StateSpace)" showing memory usage from 0.25GB to 16.00GB over time for multiple models.
- Surprise models:** A section with a legend for model types and a note: "124 instances of models had to be processed. When the curves become vertical, it means the corresponding tool cannot compute anymore model."
- Comparison charts:** A scatter plot titled "Tapaal versus GreatSPN+red Reachability Cardinality (All models)" and a line chart titled "Description for ITS-Tools on AirplaneLD-PT-0050" showing execution time in seconds for Core 3 and Core 4.



# Concluding remarks

## Details soon available online

- <https://mcc.lip6.fr>
- 177 950 charts (execution, comparisons, etc.)
- 195 582 web pages (exécution report, tool comparaison, etc.)

The screenshot displays a web interface with several components:

- Summary of Results for Reachability Cardinality:** A table with columns for tool names and total points. Visible data includes:

Tool	Total Points
GreatSPN+red	24264.48
LTSMIn+red	23729.38
ITS-Tools	24003.73
LoLA	24146.98
smpt	25000.00
2023-gold	24146.98
BVT-2024	25000.00
- Execution Summary:** A table with columns: Max Memory Used (MB), Time wait (ms), CPU Usage (ms), I/O Wait (ms), Computed Result, Execution Status. Visible data:

Max Memory Used (MB)	Time wait (ms)	CPU Usage (ms)	I/O Wait (ms)	Computed Result	Execution Status
262.063	5197.00	12562.00	227.30	FFTTT...	
- «Surprise» models:** A section with a text box stating: "The charts are below, click on the chart to enlarge it. 124 instances of models have corresponding tool cannot compute anymore model."
- Evolution of Memory:** A line chart titled "Memory Usage to Process Surprise Models (StateSpace)". The y-axis ranges from 0.25GB to 16.00GB. It shows multiple lines representing different tools, with memory usage generally increasing over time for most tools.
- Memory Consumption:** A line chart titled "FireWire, P/T for StateSpace, Memory consumption". The y-axis is "Megabytes" (0 to 16384) and the x-axis is "model instances" (01 to 20). It shows several lines, with a notable spike in memory usage for one instance (around instance 18).



# Concluding remarks

## Details soon available online

- <https://mcc.lip6.fr>
- 177 950 charts (execution, comparisons, etc.)
- 195 582 web pages (exécution report, tool comparaison, etc.)

The screenshot shows a website interface for a model checking contest. At the top, there is a table titled "Summary of Results for Reachability Cardinality" with columns for various tools and their scores. Below this, there is an "Execution Summary" table with columns for Max Memory Used (MB), Time wait (ms), CPU Usage (ms), I/O Wait (ms), Computed Result, and Execution Status. A yellow paper overlay with a red pushpin is placed over the middle of the page, containing the text: "Paper under printing «Behind the Scene of the Model Checking Contest, Analysis of Results from 2018 to 2023»". Below the paper, there are two line charts. The left chart is titled "Memory Usage to Process Surprise Models (StateSpace)" and shows memory usage in GB for various tools over time. The right chart is titled "Evolution of Memory" and shows memory usage in Megabytes for various tools across 20 model instances.

Tool	Score
GreatSPN+red	24264.48
LTSMIn+red	23729.38
ITS-Tools	24003.73
LoLA	24146.98
smpt	25000.00
Tapaal	24146.98
2023-gold	24146.98
BVT-2024	25000.00

Max Memory Used (MB)	Time wait (ms)	CPU Usage (ms)	I/O Wait (ms)	Computed Result	Execution Status
262.063	5197.00	12562.00	227.30	FETTER	

**Surprise» models**

**Evolution of Memory**

**Memory Usage to Process Surprise Models (StateSpace)**

**Evolution of Memory**

**Paper under printing**  
«Behind the Scene of the Model Checking Contest, Analysis of Results from 2018 to 2023»



**Hubert Garavel**  
(Inria)



**Quentin Nivon**  
(Inria)



**Pierre Bouvier**  
(Kalray)

**Managing  
Models**

**Managing  
Execution +  
analysis**

**Fabrice Kordon**  
(sorbonne Univ.)



**Francis Hulin-Hubard**  
(CNRS)



**Loïg Jezequel**  
(U. Nantes)



**Emmanuel Paviot-Adet**  
(Univ. Paris Clt )



**Managing  
Formulas**



# Time for discussion!

Send us your benchmarks  
and models !

See you next year!

