Time for Networks: Mutation Testing for Timed Automata Networks

Gilles Perrouin
gilles.perrouin@unamur.be
@GPerrouin

FORMALISE 2024, Lisbon, Portugal, April 14th

PReCISE
Namur Digital Institute
NADI

fnrs
FREEDOM TO RESEARCH

UNIVERSITÉ DE NAMUR

Time for Networks: Mutation Testing for Timed Automata Networks

David Cortés
Jesus Aranda
Universidad del Valle

James Ortiz
TELECOM Paris
IP PARIS

Davide Basile
Consiglio Nazionale delle Ricerche

Pierre-Yves Schobbens
UNIVERSITÉ DE NAMUR

Timed Safety-critical Systems

TAP Air Portugal Website, 2024

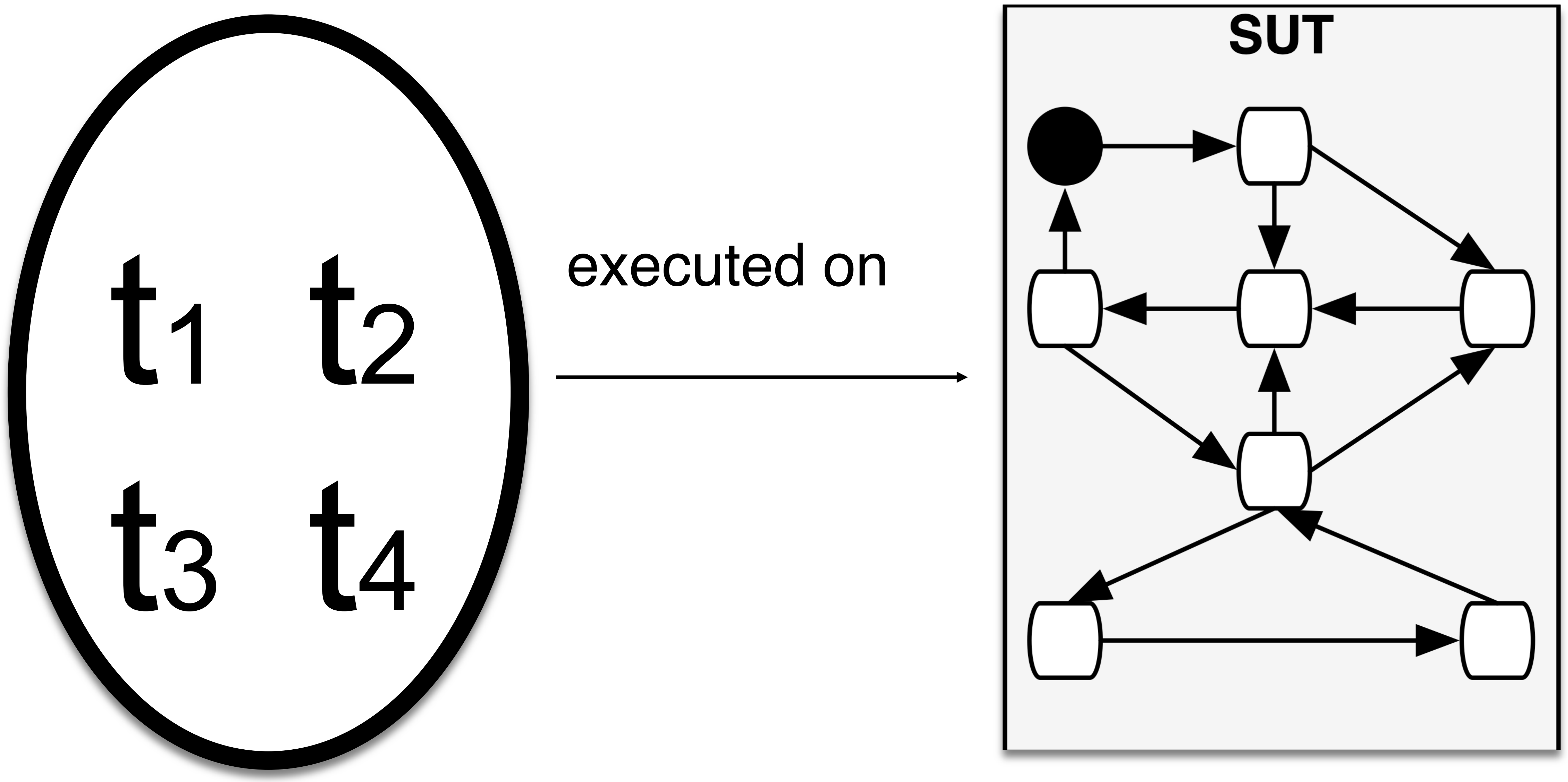
By Jcornelius - Own work, CC BY 2.5, https://commons.wikimedia.org/w/index.php?curid=1475282

Source:
https://en.wikipedia.org/wiki/Mechanical_ventilation#/media/File:Servo_I_Ventilator.jpg

Testing Timed Systems

Standards like DO-178 and ISO26262 require thorough testing of TS

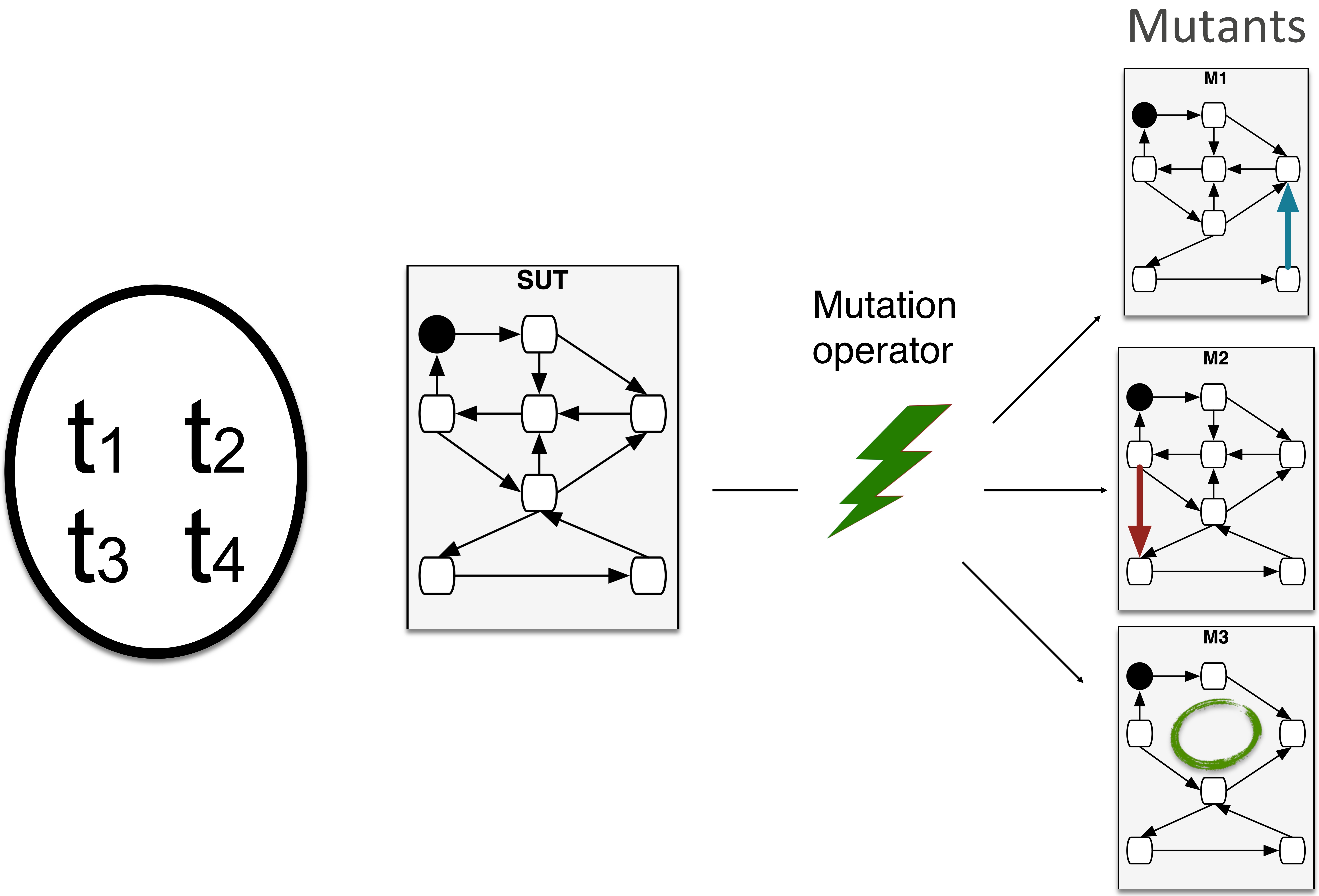
To ensure test quality, ***mutation testing*** can be used



Test Cases

System Under Test

Mutation Testing



Mutation Testing

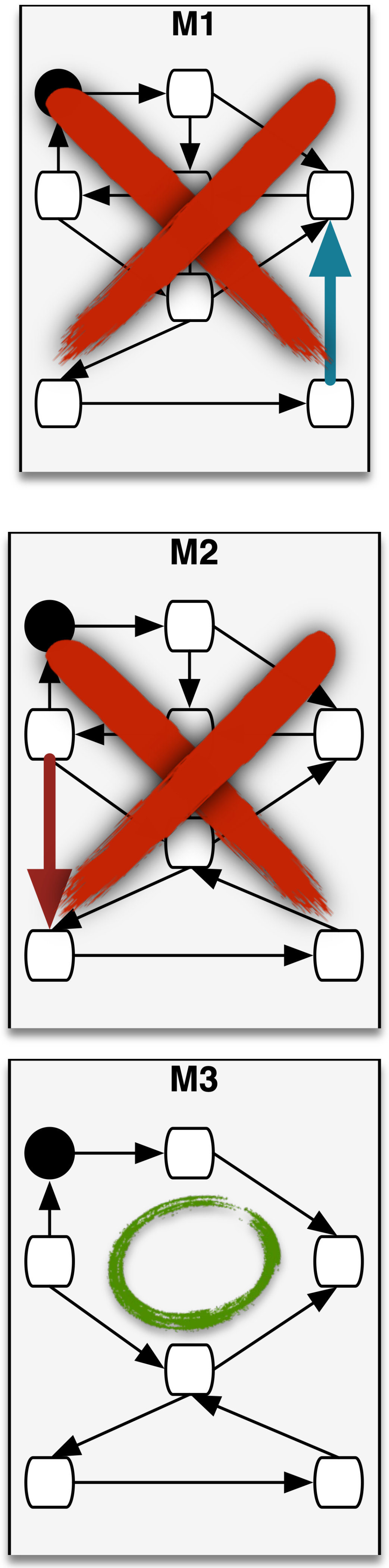
$$Score = \frac{Killed\ mutants}{Total\ mutants}$$

t_1 t_2
t_3 t_4

executed

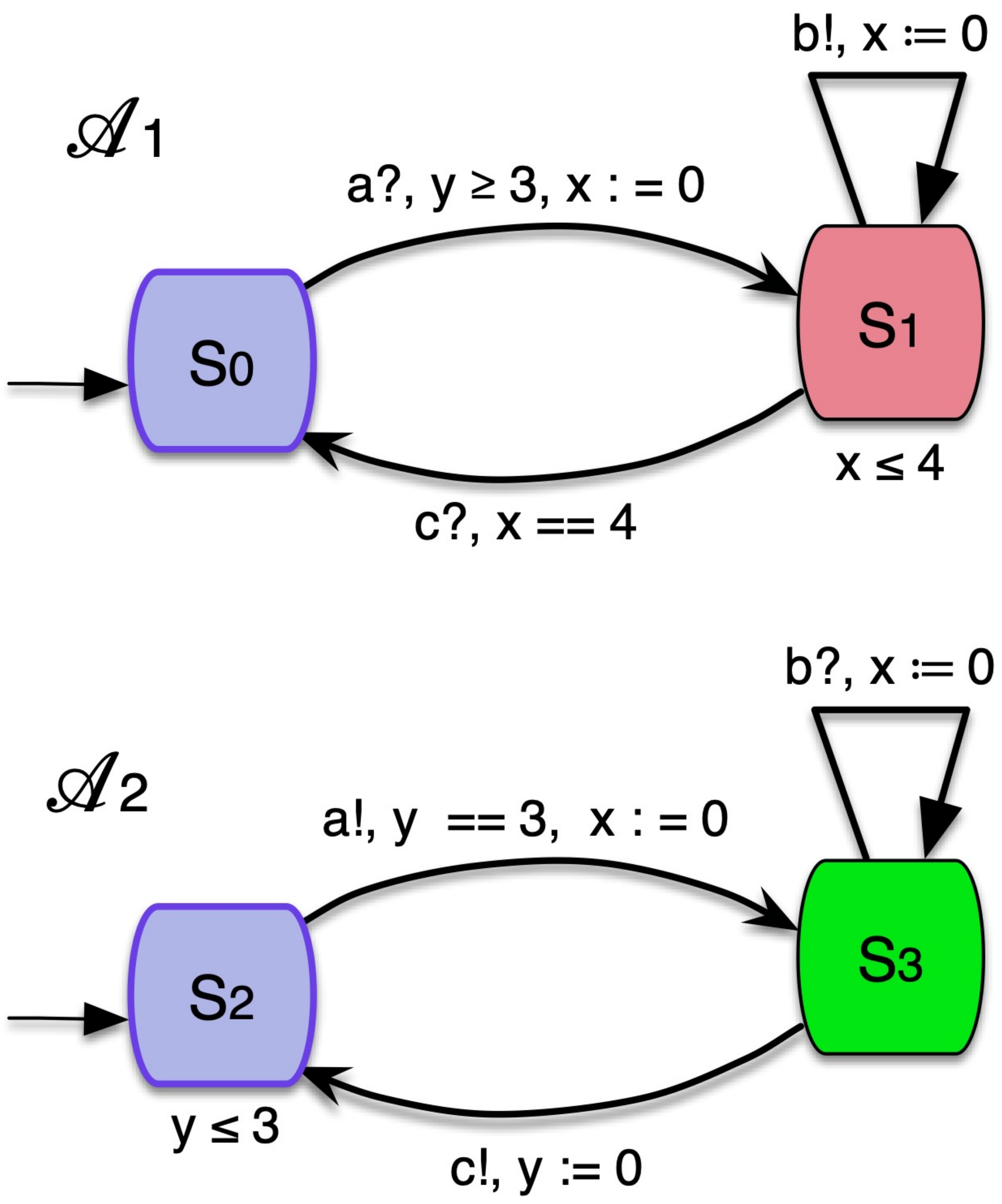
on

Mutants



Mutation score = 2/3

Modelling (Distributed) Timed Systems



Timed Automata with Input and Outputs: *TAIO*[1]

Network of Timed Automata (NTA): Parallel product of TAIO where input is blocking and synchronization is broadcast (one sender, zero or more receivers)

1. Kaynar, D., Lynch, N., Segala, R., & Vaandrager, F. (2010). *The theory of timed I/O automata*. Morgan & Claypool Publishers.

Mutation Operators for Timed Automata

Nilsson *et al.* [38]		Aichernig *et al.* [2]		Basile *et al.* [7]	
Op	**Description**	**Op**	**Description**	**Op**	**Description**
ET	Execution time	**CA**	Change action	**TMI**	Transition missing
IAT	Inter-arrival time	**CT**	Change target	**TAD**	Transition ADd
PO	Pattern offset	**CS**	Change source	**SMI**	State missing
LT	Lock time	**CG**	Change guard	**CXL**	Constant exchange L
UT	Unlock time	**NG**	Negate guard	**CXS**	Constant exchange S
HTS	Hold time shift	**CI**	Change invariant	**CCN**	Constraint negation
PC	Precedence constraints	**SL**	Sink location	-	-
-	-	**IR**	Invert reset	-	-

Mutation-based Testing Criteria for Timeliness

Robert Nilsson[†] — [†]Department of Computer Science, University of Skövde, Box 408SE, 541 28 Skövde, Sweden, {robert,sten}@ida.his.se
Jeff Offutt[*] — [*]Department of Information and Software Engineering, George Mason University, Fairfax, VA 22030 USA, ofut@ise.gmu.edu
Sten F. Andler[†]

Time for Mutants —Model-Based Mutation Testing with Timed Automata

Bernhard K. Aichernig , Florian Lorber , and Dejan Ničković

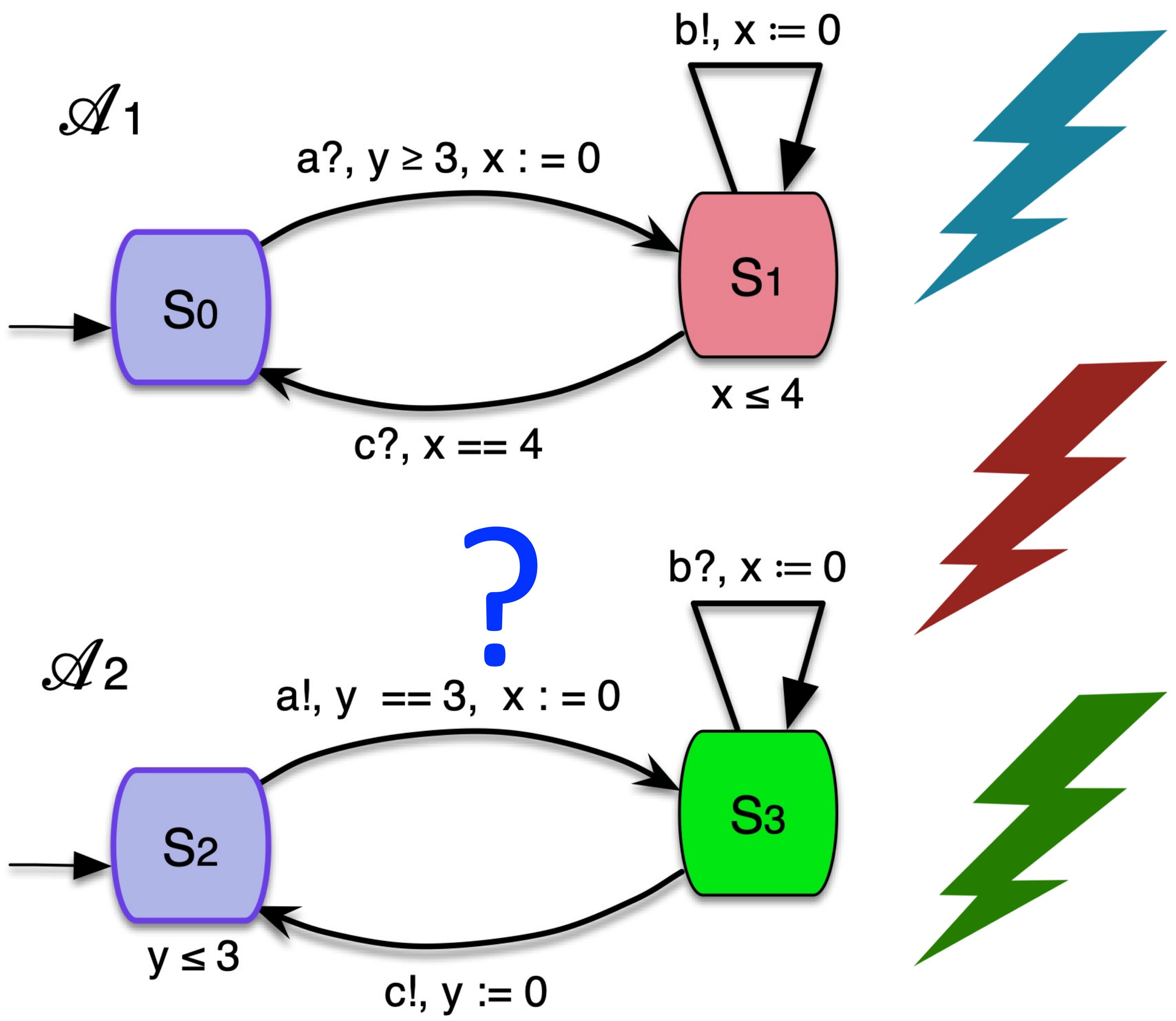
Tackling the Equivalent Mutant Problem in Real-Time Systems: The 12 Commandments of Model-Based Mutation Testing

Davide Basile, Maurice H. ter Beek, Maxime Cordy, Axel Legay

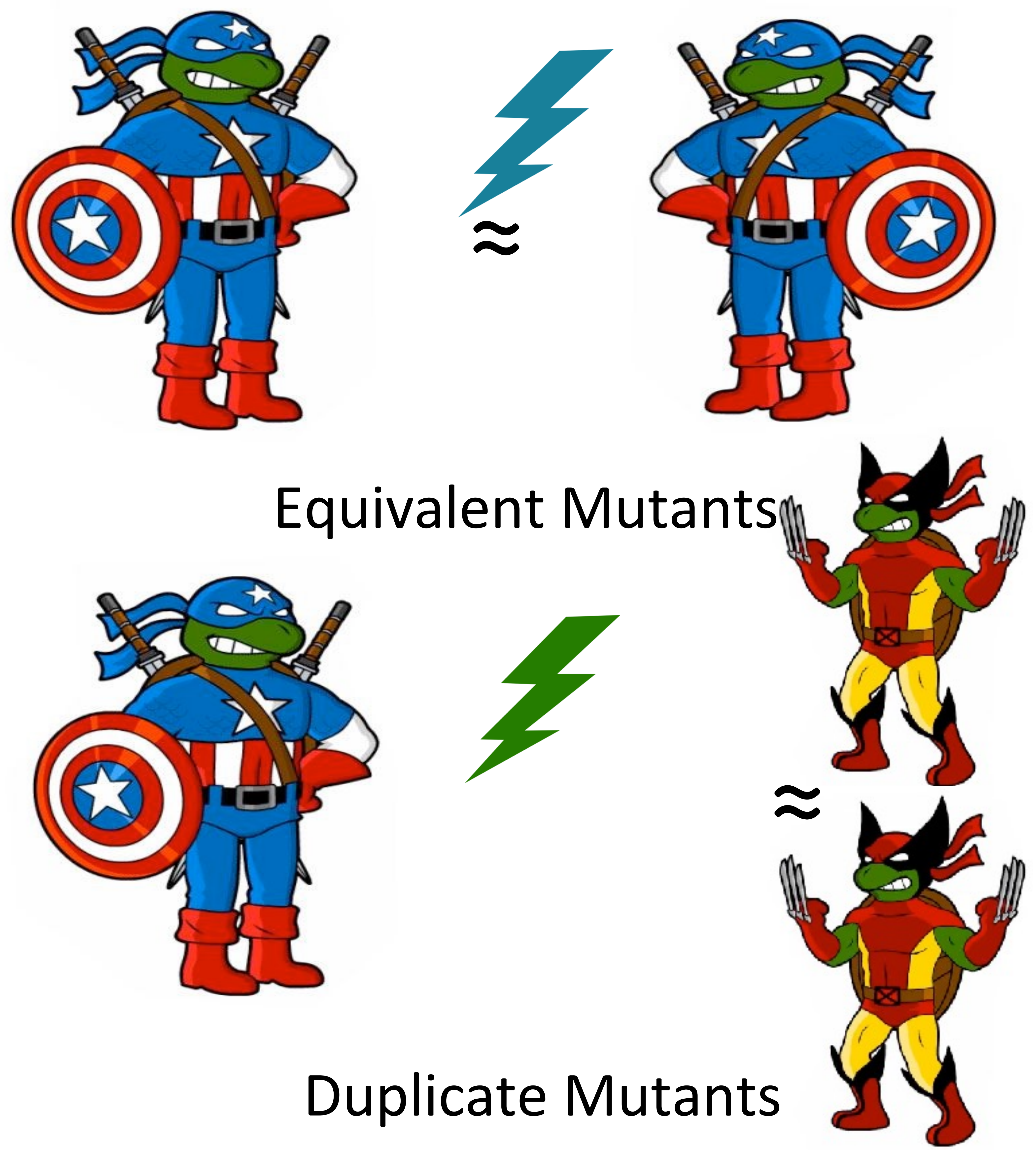
Timed Automata with Task model

Timed Automata with Input and Output

Motivation



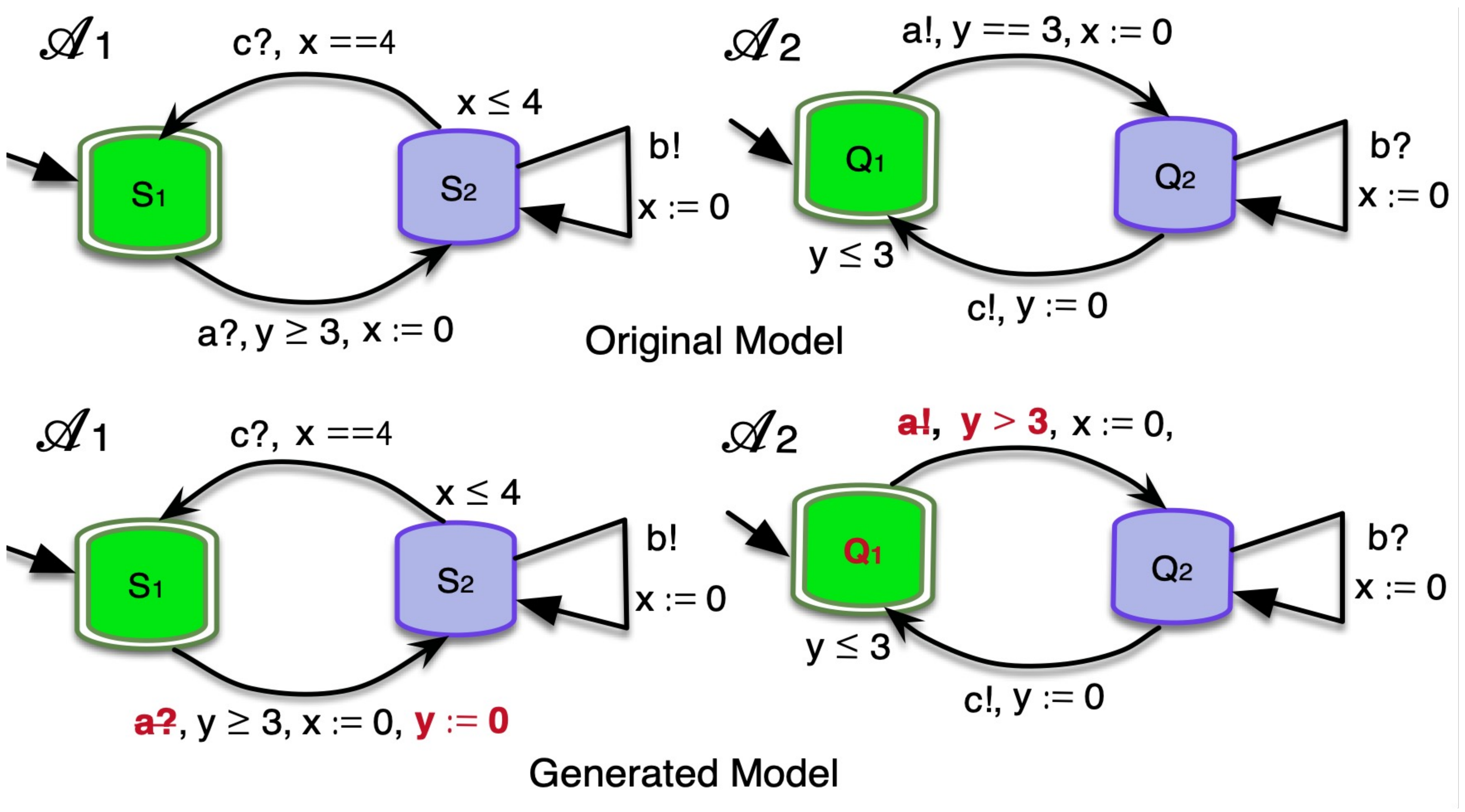
Existing operators do not focus on NTA

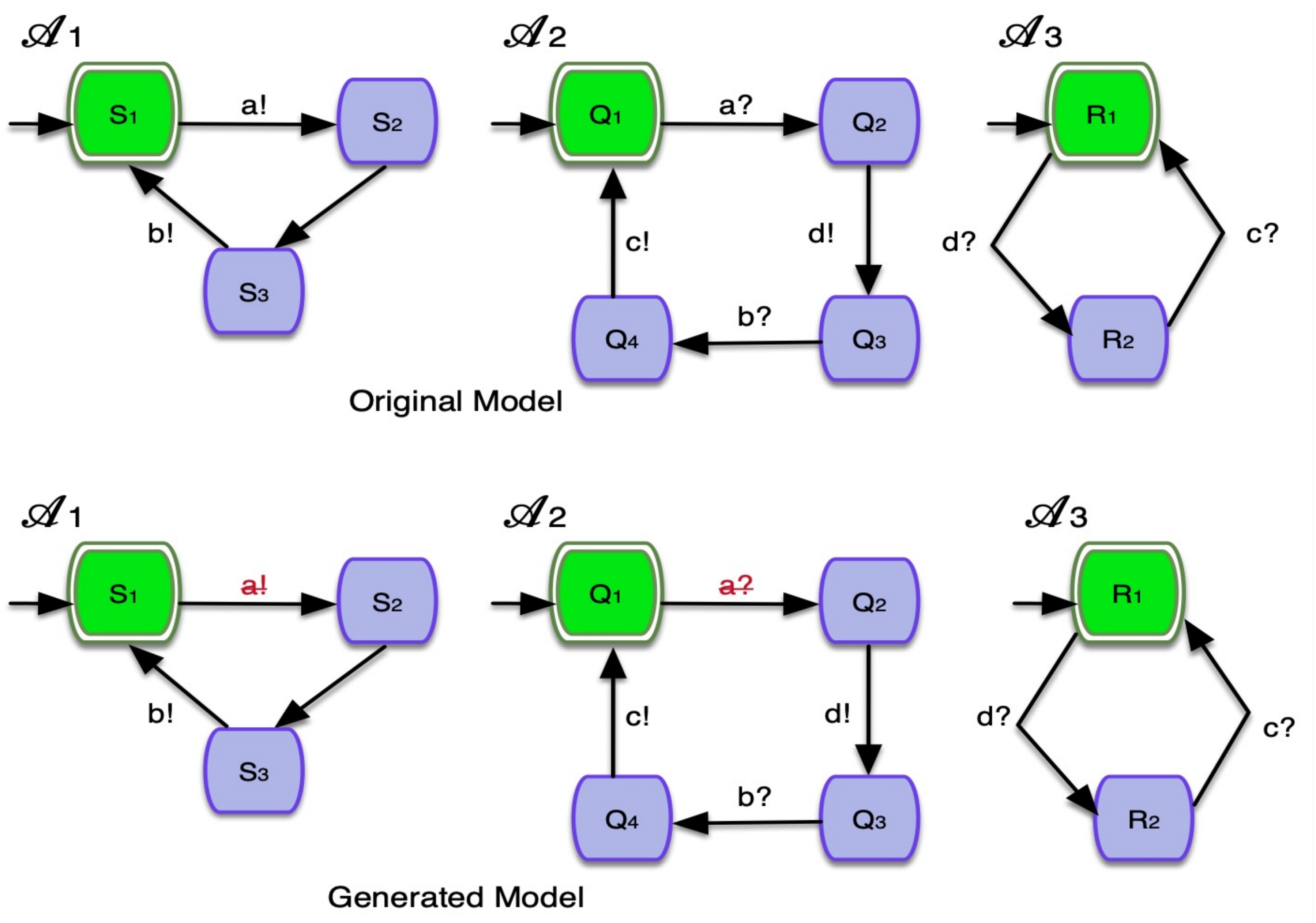


Mutation Operators for NTA

	Name	Acronym	Type
NTAIO	*syncSeq*	SS	Sequential
	delSync	DS	Interleave
	maskVarh	MVCh	Shadow global channel
	maskVarc	MVCc	Shadow global clock
	urgChan	UC	Urgent Channel
	BroadChan	BC	Broadcast Channel
TAIO	*urgLoc*	UL	Urgent Location
	commLoc	CL	Committed Location

Synchronisation Operators

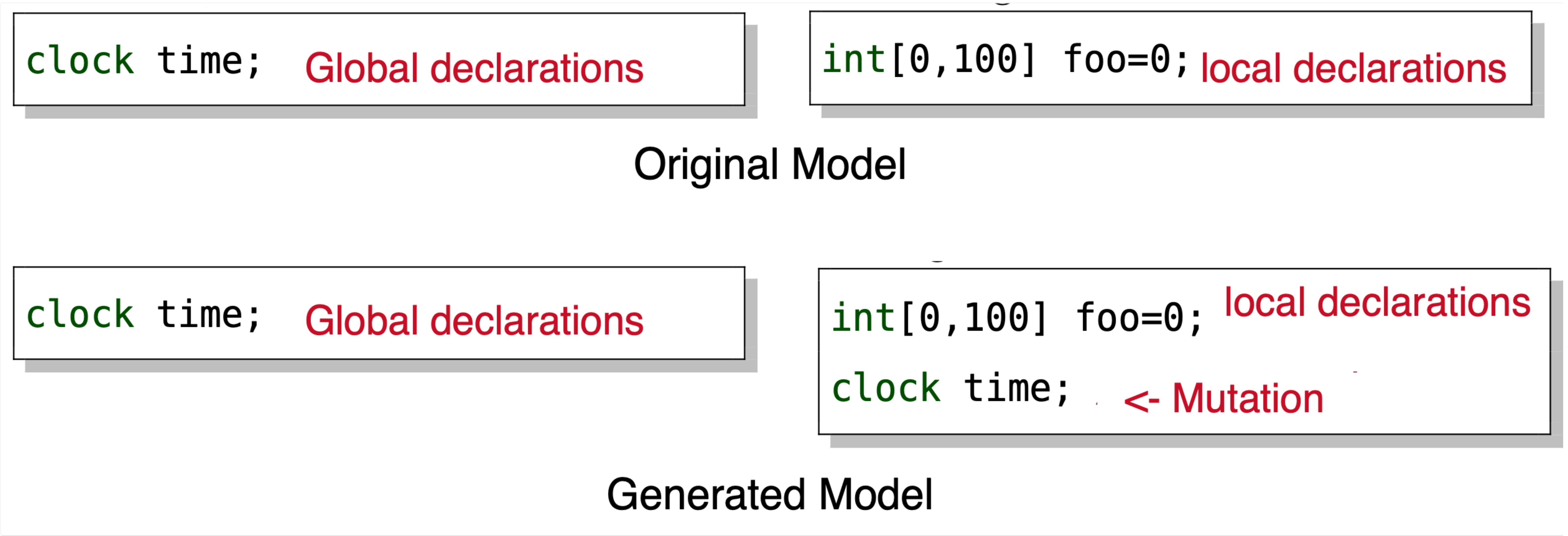




SyncSeq. Forces sequential behaviour by removing synchronisation events and commits a source location.

DelSeq. Removes synchronisation

Masking Operators



MVCc (or MVCl). Masks a global clock by introducing a local clock with the same name.

MVCh. Similar mutation but for channels (not shown here)

Channel Operators

```
chan go;
clock x;
```

Global declarations

```
broadcast chan go;
clock x;
```

Global declarations

Original Models

```
broadcast chan go;
clock x;
```

Global declarations

```
broadcast urgent chan go;
clock x;
```

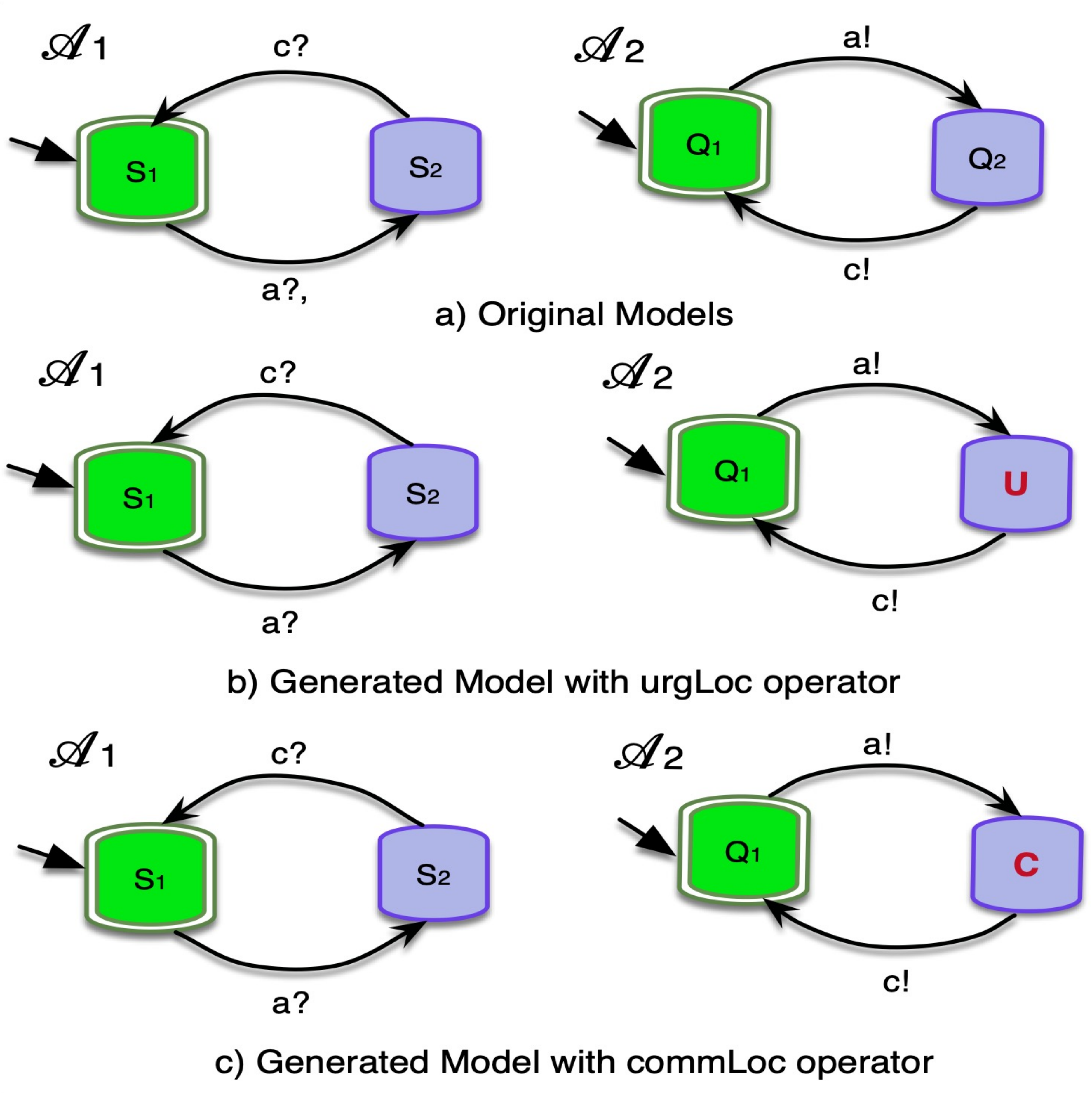
Global declarations

Generated Models

BroadChan. Transforms an unicast channel into a broadcast one

UrgChan. Make the channel urgent (prevent delays and encourage synchronisation)

Location Operators



UrgLoc. Make location as Urgent (no time can be waited for)

CommLoc. Make location as committed (freezes time)

Tackling Timed Mutants' Uselessness

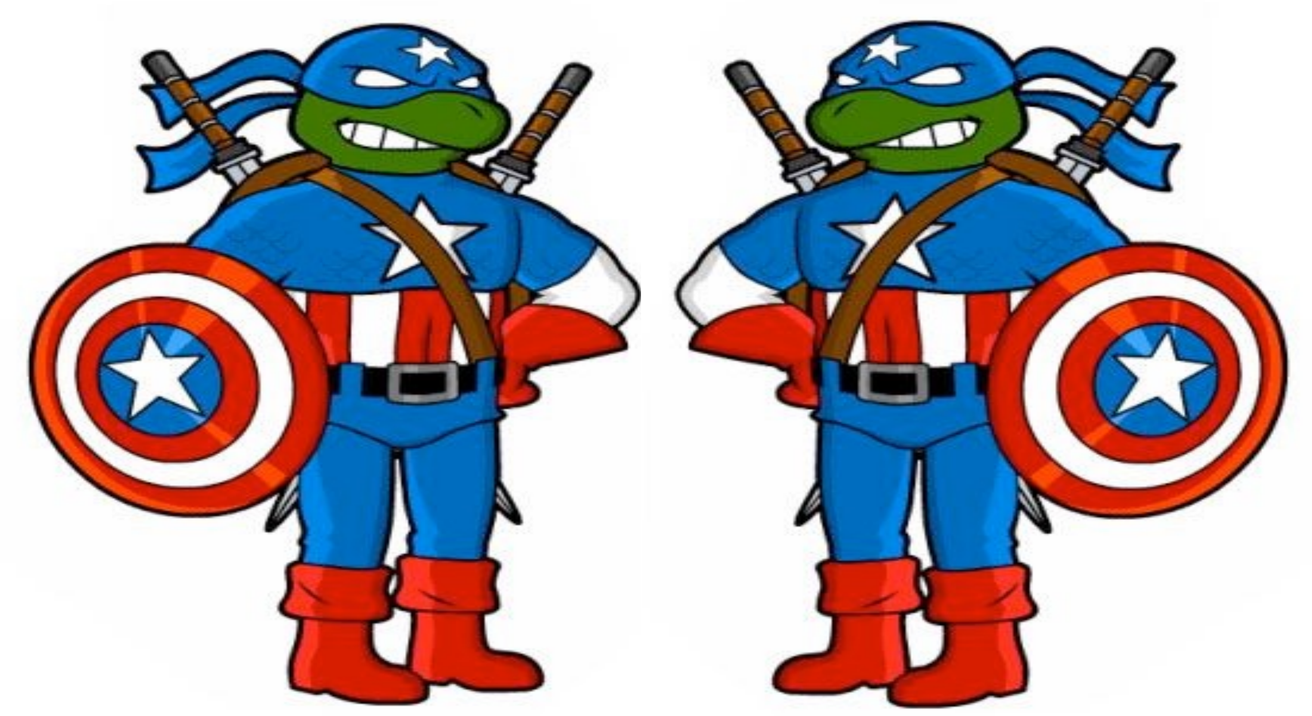
Davide Basile *et al.* ***Tackling the equivalent mutant problem in real-time systems: the 12 commandments of model-based mutation testing***. In SPLC'20: 24th ACM International Systems and Software Product Line Conference

Davide Basile *et al.* ***Static detection of equivalent mutants in real-time model-based mutation testing: An Empirical Evaluation***. Empirical Software Engineering 27, 7 (2022), 160.

Jaime Cuartas *et al.* ***MUPPAAL: Reducing and Removing Equivalent and Duplicate Mutants in UPPAAL***. In 2023 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW). IEEE, 52–61

Research Questions

RQ1. How many mutants?



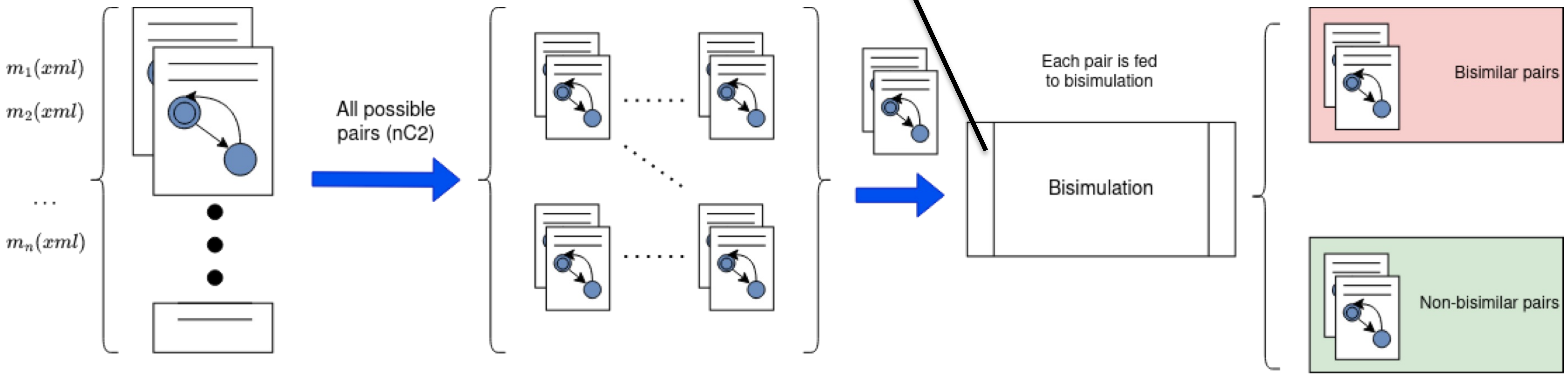
RQ2. TAIO vs NTAIO Operators (Equivalent Mutants)

RQ3. TAIO vs NTAIO Operators (Duplicate Mutants)

RQ4. Bisimulation Costs

Evaluation

James Jerson Ortiz, *et al*. 2017. ***Multi-timed Bisimulation for Distributed Timed Automata. In NASA Formal Methods*** - 9th International Symposium, NFM 2017, pp. 52–67.

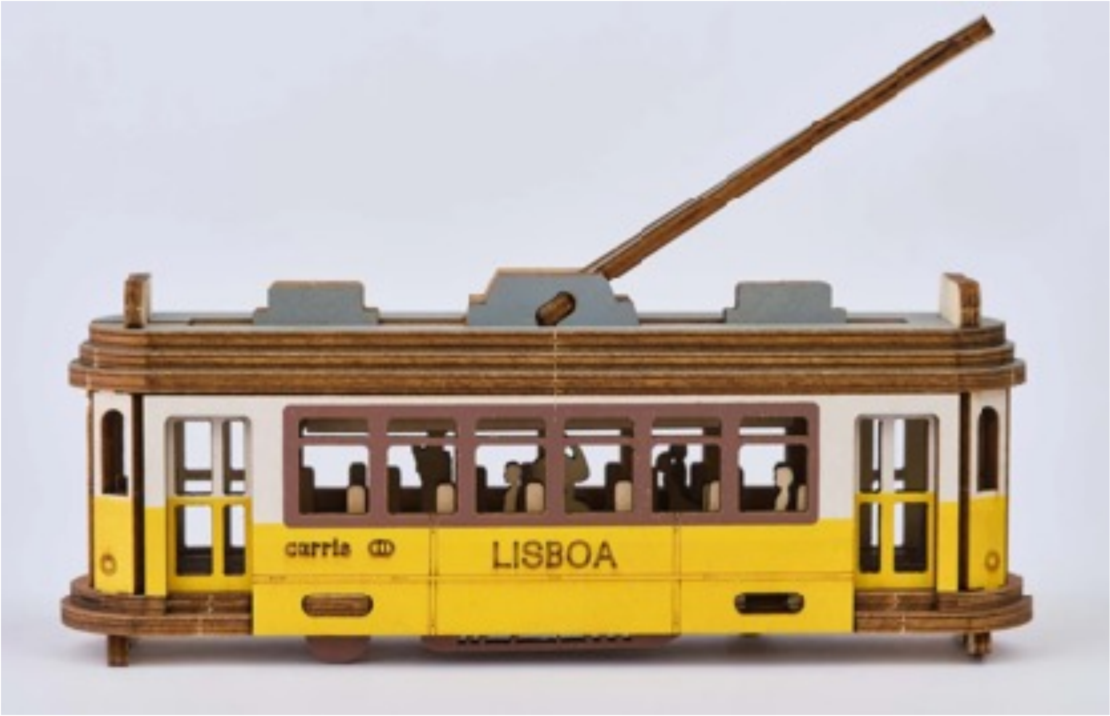


Jaime Cuartas *et al.* ***MUPPAAL: Reducing and Removing Equivalent and Duplicate Mutants in UPPAAL***. In 2023 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW). IEEE, 52–61

Case Studies

Collision Avoidance (CA)
Locations: 12
Transitions: 26

Train Gate Controller (TGC)
Locations: 17
Transitions: 16



Tram Door (TD)
Locations: 32
Transitions: 36

Gear Controller (GC)
Locations: 48
Transitions: 30

Mechanical Ventilator (MV)
Locations: 24
Transitions: 29

RQ1. How Many Mutants?

	CA	GC	TGC	MV	TD
DS	25	56	36	17	36
MVCh	22	40	16	60	55
MVCl	2	2	7	9	8
UC	11	20	8	12	11
UL	12	46	28	17	36
CL	12	46	28	20	31
SS	9	18	7	7	10
BC	0	0	0	7	0
Total	93	228	130	149	179

The Delete Synchronisation operator ***is the most applicable to our cases*** (170/779 ≈ 21%)

The Broadcast Channel i***s only applicable to MV*** since all the other cases already have a broadcast channel (7/779 ≈ 0.8%)

A diverse set of mutant operators applicable to various (N)TAIO constructs

RQ2. Mutant Equivalence

Case	Collision Avoidance	Gear Control	Train Gate Controller	Mechanical Ventilator	Tram Door	Average
DS	4/25 (16%)	9/56 (16%)	3/36 (8%)	15/17 (88%)	1/36 (3%)	26%
MVCh	3/22 (14%)	38/40 (95%)	14/16 (87%)	56/60 (93%)	36/55 (65%)	71%
MVCc	0/2 (0%)	1/2 (50%)	4/7 (57%)	6/9 (66%)	3/8 (37%)	42%
UC	5/11 (45 %)	19/20 (95%)	5/8 (62%)	8/12 (66%)	6/11 (54%)	64%
UL	4/12 (33%)	43/46 (93%)	21/28 (75%)	6/17 (35%)	24/36 (66%)	60%
CL	3/12 (25%)	43/46 (93%)	21/28 (75%)	9/20 (45%)	20/31 (64%)	60%
SS	0/9 (0%)	1/18 (5%)	0/7 (0%)	4/7 (57%)	0/10 (0%)	12%
BC	0 (0%)	0 (0%)	0 (0%)	2/7 (28%)	0 (0%)	28%
Total	19/93 (20%)	154/228 (67%)	75/130 (58%)	106/149 (71%)	90/187 (48%)	52%

The masking channel operator is the ***highest contributor*** to equivalent mutants

This ***is not true*** for the other masking operator (MVCc)

NTAIO mutation operators ***either generate the lowest or the highest numbers*** of equivalent mutants

RQ3. Mutant Duplicates

Case	Most Duplicated Operator per pair type
Collision Avoidance	MVCh (5/7)
Gear Controller	CL (2/7), UL (2/7), CL & UL (3/7)
Train Gate Controller	CL (2/7), UL (2/7), CL & UL (3/7)
Mechanical Ventilator	MVCh (6/7)
Tram Door	MVCh (4/6) (MVCc did yield 0 duplicate)

MVCh is the largest contributor to duplicate pairs

« Location » (UL and CL) operators contribute ***equally***

RQ4. Computational Costs

Case study	Mutant generation	Bisimulation of all pairs	Bisimulation per pair
CA	1 s	24 min	660 ± 41 ms
GC	3 s	172 min	1 s ± 126 ms
TGC	2 s	153 min	748 ± 52 ms
MV	2.5 s	165 min	1 s ± 177 ms
TD	3 s	195 min	726 ± 67 ms

Mutant generation is fast (as expected)

Overall comparisons can take ***more than 3 hours*** for Tram Door…

…but on average ***timed bisimulation takes less than 1s per comparison***

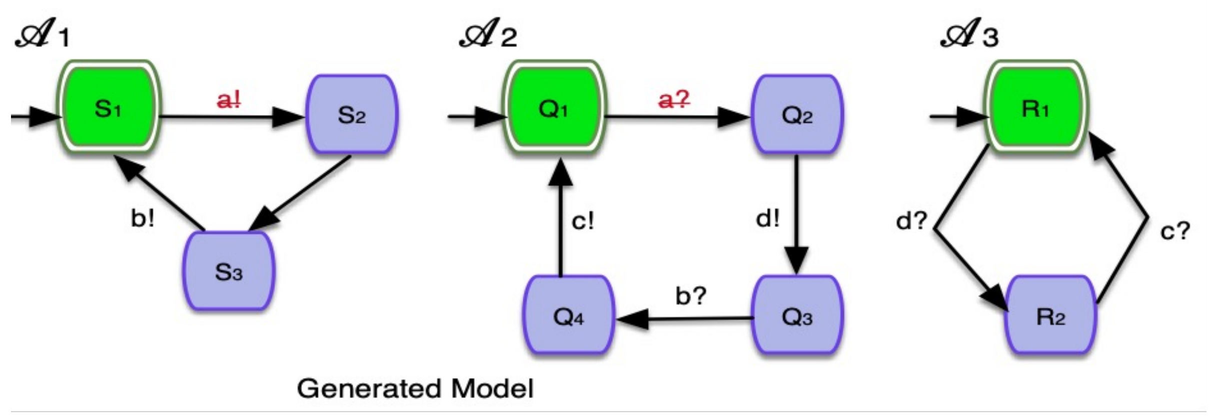
Discussion

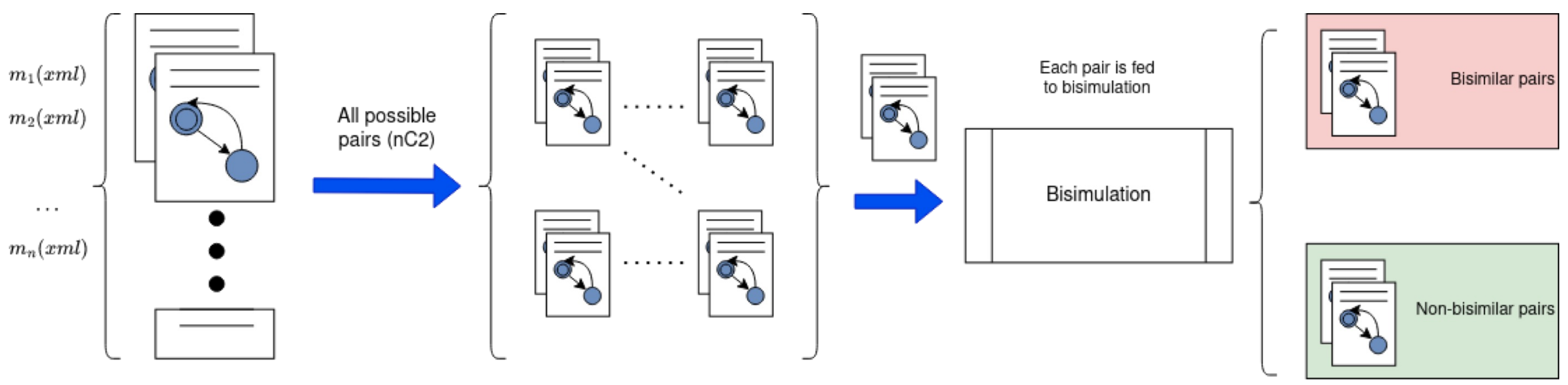
Number of equivalent mutants. Quite high, influenced by resets for MVCh operator.

Static detection of mutants. Static detection of equivalent mutants could save time for some mutant operators. The influence of resets is hard to detect statically. Precision/speed tradeoff to investigate.

Timed Bisimulation for mutant equivalence. Strongest relation that allows us to distinguish subtle behavioural faults. We would like to see the influence of other relations (TIOCO, …) in the future.

Stubbornness. Studies are required to see how useful are mutants and how hard they are to kill in practice.





8 mutation operators for ***NTAIO***

Mutation framework based on MUPPAAL

5 cases studies

Diversity of behaviours regarding equivalents/duplicates

Timed bisimulation ***scales***

GitLab

Time for Networks: Mutation Testing for Timed Automata Networks

Backup Slides

FORMALISE 2024, Lisbon, Portugal, April 14th

Case Studies Details

Case Studies	Instances	Locations	Transitions	Clocks	Channels	Broad. Channels	Urg. Channels
CA	2	12	26	1	0	11	0
GC	2	48	60	1	0	20	0
TGC	3	17	18	1	0	9	0
MV	5	24	29	6	7	12	0
TD	5	32	36	2	0	11	0

Jaime Cuartas, *et al.*. ***Formal Verification of a Mechanical Ventilator using UPPAAL***. In Proceedings of the 9th ACM SIGPLAN International Workshop on Formal Techniques for Safety-Critical Systems (FTSCS 2023). Association for Computing Machinery, New York, NY, USA, 2–13. https://doi.org/10.1145/3623503.3623536

Mutant Duplicates (per case study)

Table 6: Average redundant mutants for the CA case study.

Operator	No. duplicated pairs	Most duplicated operator
DS	56/8556 (0.65%)	MVCH 10 (27%)
MVCh	1055 /8556 (12.3%)	CL & UL 528 (50%)
MVCl	116 /8556 (1.3%)	MVCH 44 (38%)
UC	591 /8556 (7%)	MVCH 242 (40%)
UL	639 /8556 (7.5%)	MVCH 264 (41%)
CL	636 /8556 (7.5%)	MVCH 264 (42%)
SS	0/8556(0%)	N/A

Table 7: Average duplicate mutants for the TGC case study.

Operator	No. duplicated pairs	Most duplicated operator
DS	643 / 16770 (4%)	UL 153 (23%)
MVCh	1292 / 16770 (8%)	CL & UL 896 (69%)
MVCl	354 / 16770 (2%)	CL & UL 224 (63%)
UC	680 / 16770 (4%)	CL & UL 448 (65%)
UL	2107 / 16770 (12.5%)	CL 784 (37%)
CL	2108 / 16770 (12.5%)	UL 784 (37%)
SS	28 / 16770 (0.1%)	CL 12 (42%)

Table 8: average duplicate pairs of mutants for each operator, with the TD case study.

Operator	No. duplicated pairs	Most duplicated operator
DS	44 / 31,862 (0.1%)	MVCH 14 (31%)
MVCh	5569 / 31,862 (17.5%)	UL 1705 (30%)
UC	1356 / 31,862 (4.2%)	MVCH 605 (44%)
UL	3509 / 31,862 (11%)	MVCH 1705 (48%)
CL	3511 / 31,862 (11%)	MVCH 1704 (48%)
SS	3 / 31,862 (0.009%)	CL 1 (33%)

Table 9: Average duplicate pairs of mutants for each operator, with the GC case study.

Operator	No. duplicated pairs	Most duplicated operator
DS	939/ 51,756 (1.8%)	CL 167 (17%)
MVCh	5515 / 51,756 (10.6%)	CL & UL 3680 (66%)
MVCl	316 / 51,756 (0.6%)	CL & UL 184 (58%)
UC	2971 / 51,756 (5.7%)	CL & UL 1840 (62%)
UL	6221 / 51,756 (12%)	CL 2116 (34%)
CL	6221 / 51,756 (12%)	UL 2116 (34%)
SS	193 / 51,756 (0.3%)	UL 58 (30%)

Mutant Duplicates (MV)

Table 10: Average duplicate pairs of mutants for each operator, with the *MV* case study.

Operator	No. duplicated pairs	Most duplicated operator
DS	2448 / 22,052 (11%)	CL 167 (17%)
MVCh	4837 / 22,052 2 (22%)	MVCH 1546 (31%)
MVCl	972 / 22,052 (4.4%)	MVCH 504 (51%)
UC	1278 /22,052 (5.7%)	MVCH 672 (52%)
UL	909 / 22,052 (4.1%)	MVCH 379 (41%)
CL	1225 / 22,052 (5.5%)	MVCH 548 (44%)
BC	128 / 22,052 (0.5%)	MVCH 55 (42%)
SS	685 / 22,052 (3.1%)	MVCH 340 (49%)