

Temporal Causality in Reactive Systems

Hadar Frenkel

January 5, 2025

BIU Formal Verification Seminar



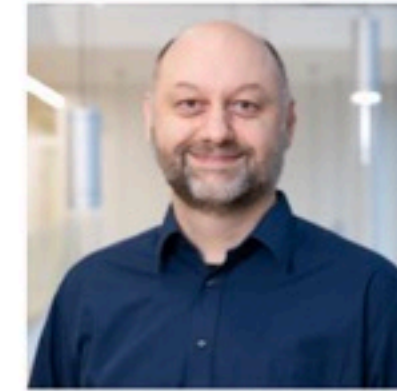
Raven Beutner



Norine Coenen



Raimund Dachzelt



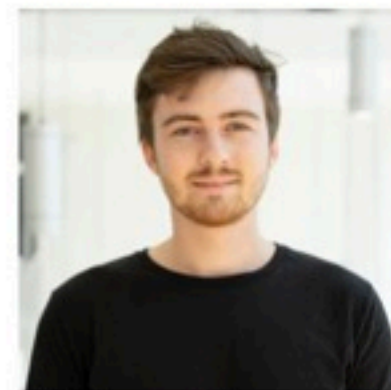
Bernd Finkbeiner



Christopher Hahn



Tom Horak



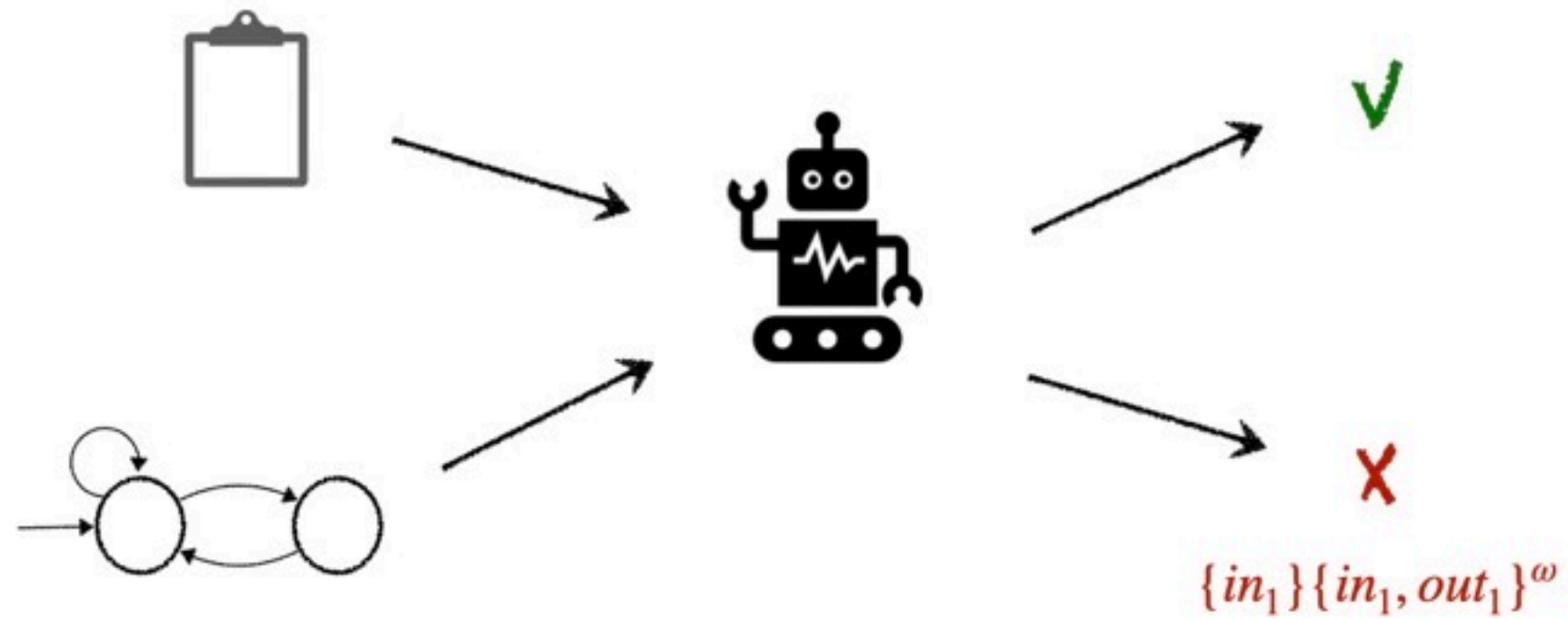
Niklas Metzger



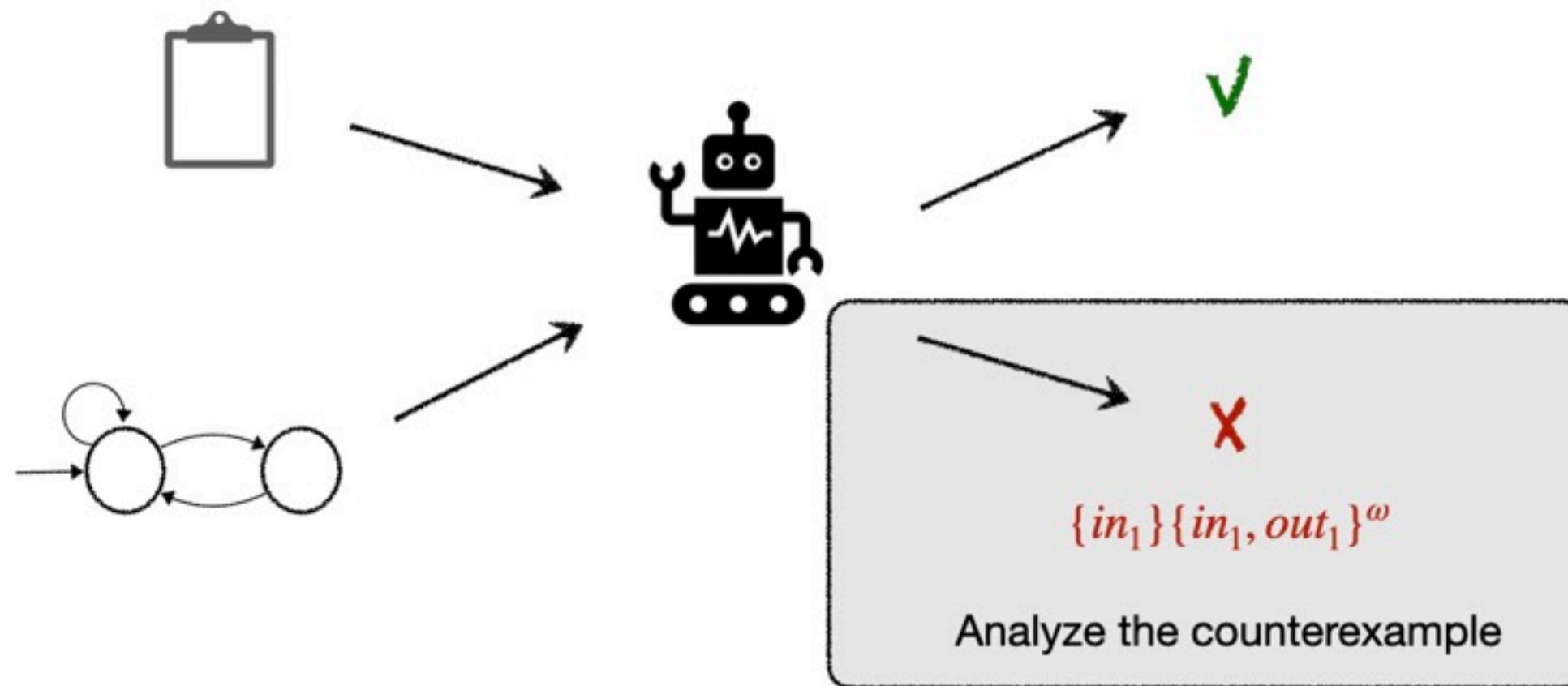
Julian Siber

Explaining Hyperproperty Violations. Coenen, Dachzelt, Finkbeiner, **F.**, Hahn, Horak, Metzger, Siber. (CAV 2022)
Temporal Causality in Reactive Systems. Coenen, Finkbeiner, **F.**, Hahn, Metzger, Siber. (ATVA 2022)
Checking and Sketching Causes on Temporal Sequences. Beutner, Finkbeiner, **F.**, Siber. (ATVA 2023)
Synthesis of Temporal Causality. Finkbeiner, **F.**, Metzger, Siber. (CAV 2024)

Motivation: Explaining Counterexamples



Motivation: Explaining Counterexamples

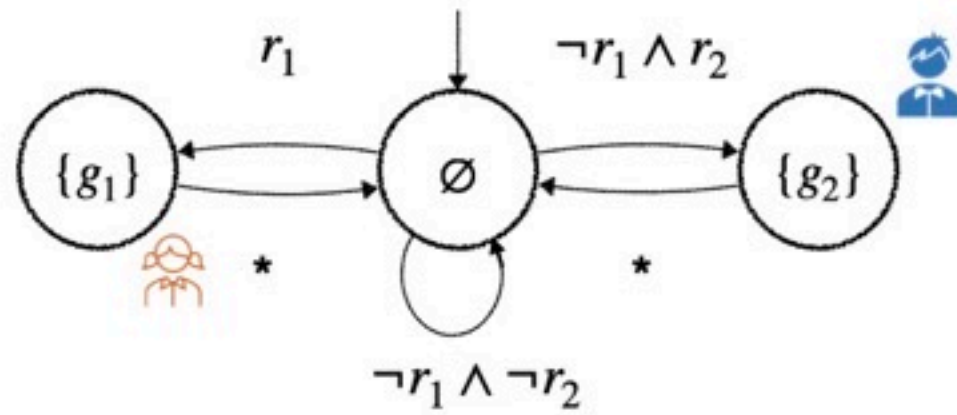


Setting

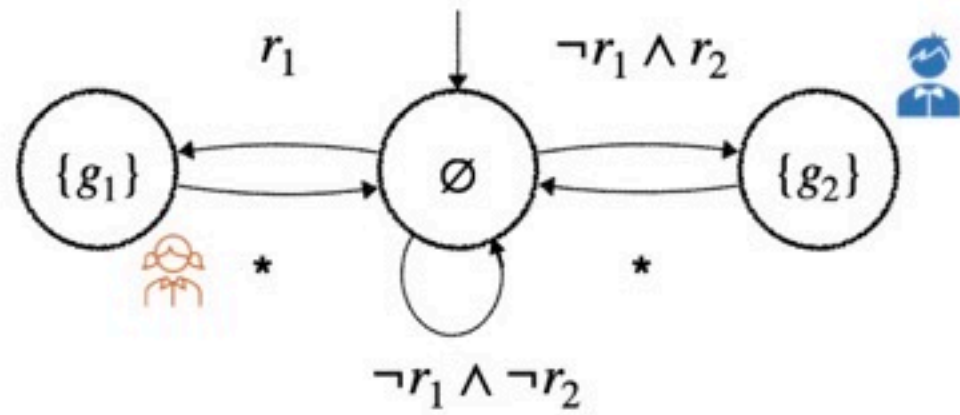
Reactive systems

Inputs trigger outputs

Infinite computations

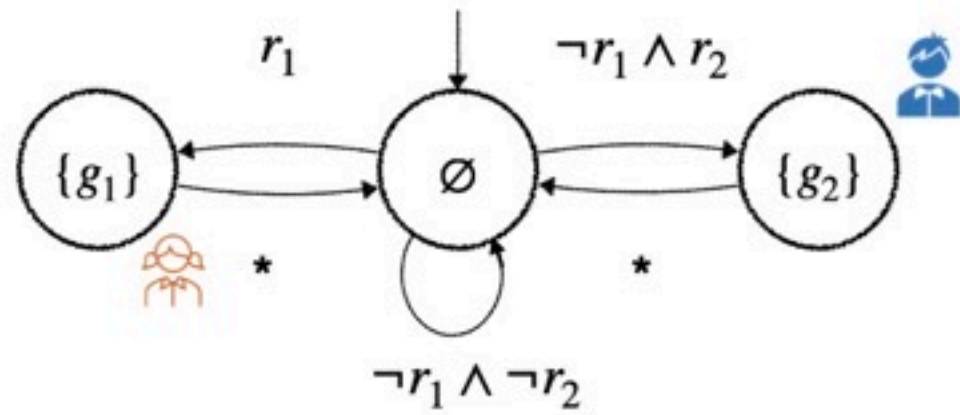


Setting

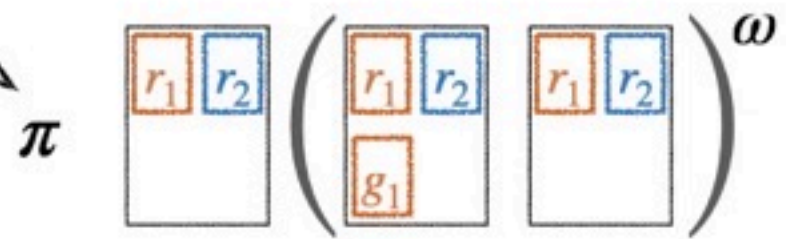


 “eventually g_2 ”

Setting

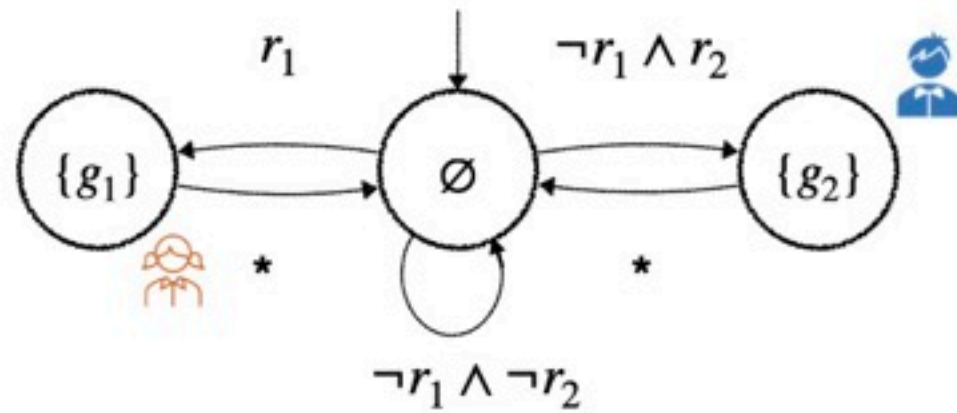


X counterexample

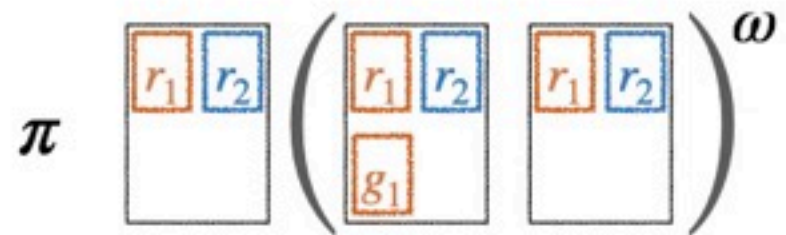


“eventually g_2 ”

Setting



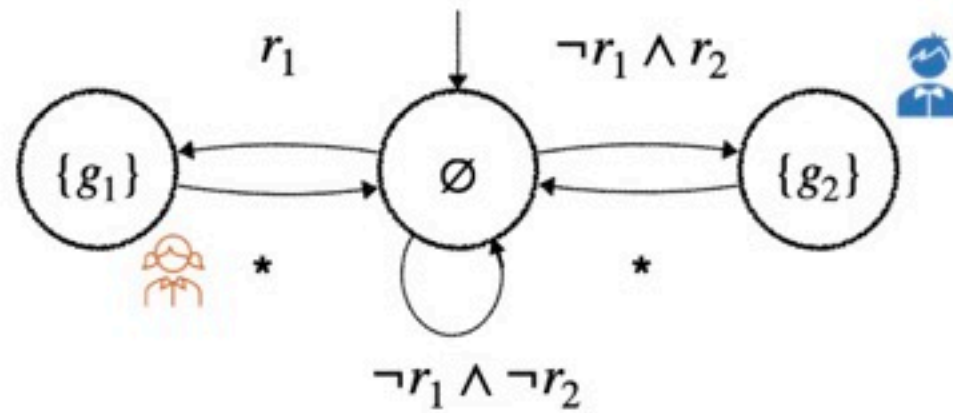
Is “always r_1 ”
the cause for “always not g_2 ” on π ?



 “eventually g_2 ”

Setting

Causes over input sequences
Analyse the system dynamics

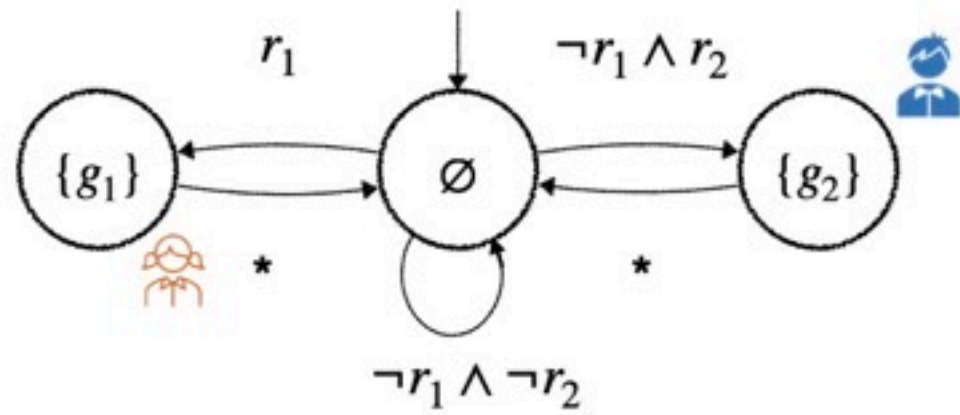


Is “always r_1 ”
the cause for “always not g_2 ” on π ?

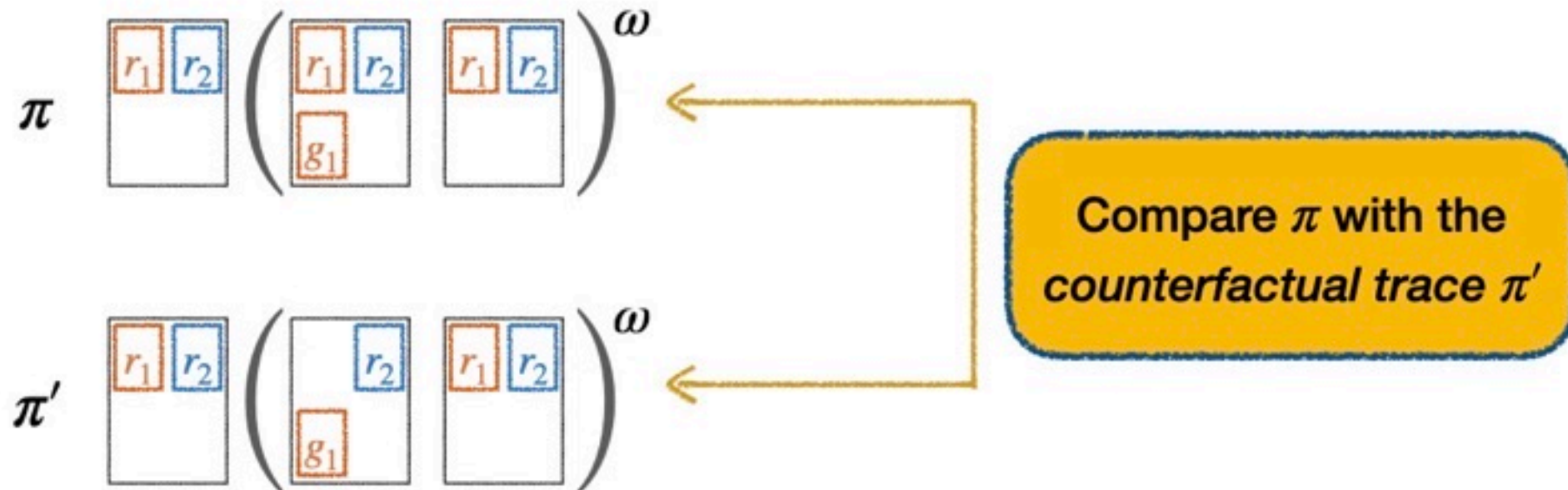


 “eventually g_2 ”

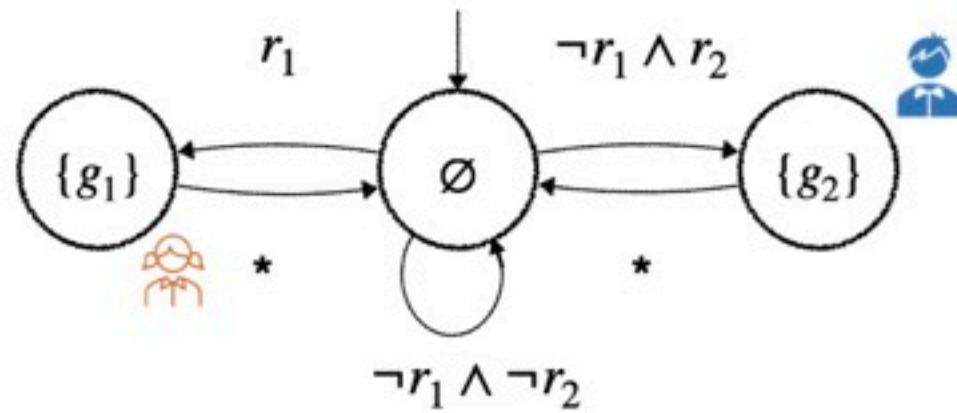
Setting



Is “always r_1 ”
the cause for “always not g_2 ” on π ?

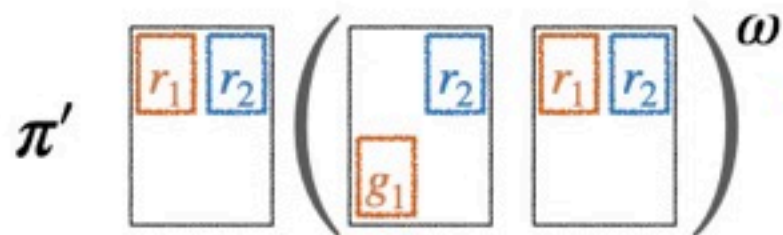
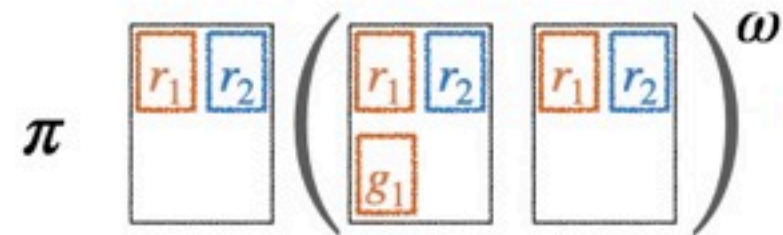


Setting



Is “always r_1 ”
the cause for “always not g_2 ” on π ?

Hyperproperties



Compare π with the
counterfactual trace π'

Overview

Halpern and Pearl's actual causality

Hyperproperties and how they allow us to talk
about causality

Overview

Halpern and Pearl's actual causality

Hyperproperties and how they allow us to talk about causality

Actual causality for reactive systems

Causes as sets of events

Causes as temporal properties

Algorithmic challenge:
How to compute causes?

Semantic challenge:
How to define causes?

Overview

Halpern and Pearl's actual causality

Hyperproperties and how they allow us to talk about causality



Actual causality for reactive systems

Causes as sets of events

Causes as temporal properties

Algorithmic challenge:
How to compute causes?

Semantic challenge:
How to define causes?

Actual Causality

Causal analysis is confusing!

Joseph Y. Halpern, Judea Pearl:

Causes and Explanations: A Structural-Model Approach: Part 1: Causes. UAI 2001: 194-202

Causes and Explanations: A Structural-Model Approach. Part I: Causes

Author(s): Joseph Y. Halpern and Judea Pearl

Source: *The British Journal for the Philosophy of Science*, Dec., 2005, Vol. 56, No. 4 (Dec., 2005), pp. 843-887

Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

A Modification of the Halpern-Pearl Definition of Causality

Joseph Y. Halpern*
Cornell University

Actual Causality



Actual Causality

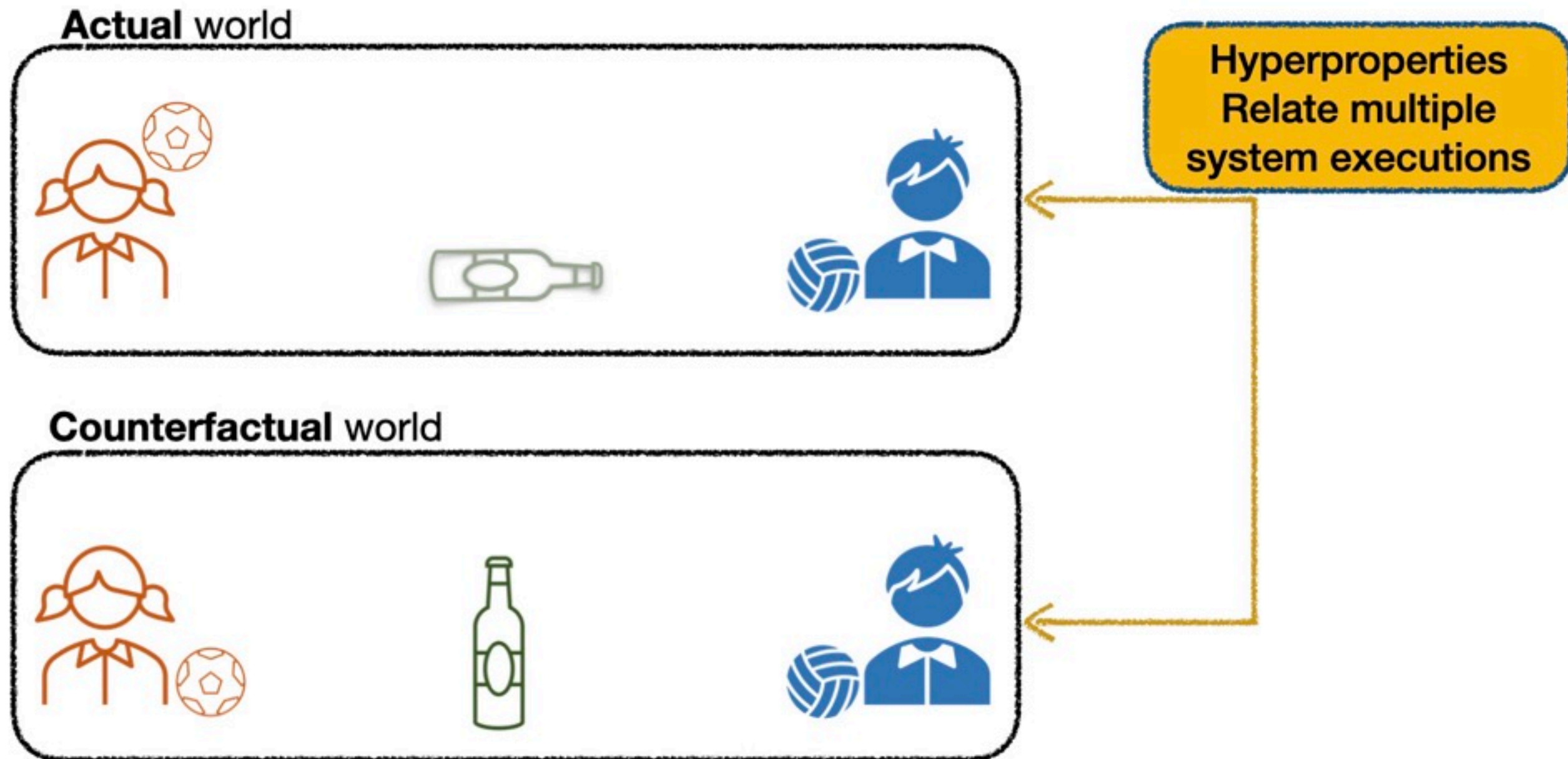
Actual world



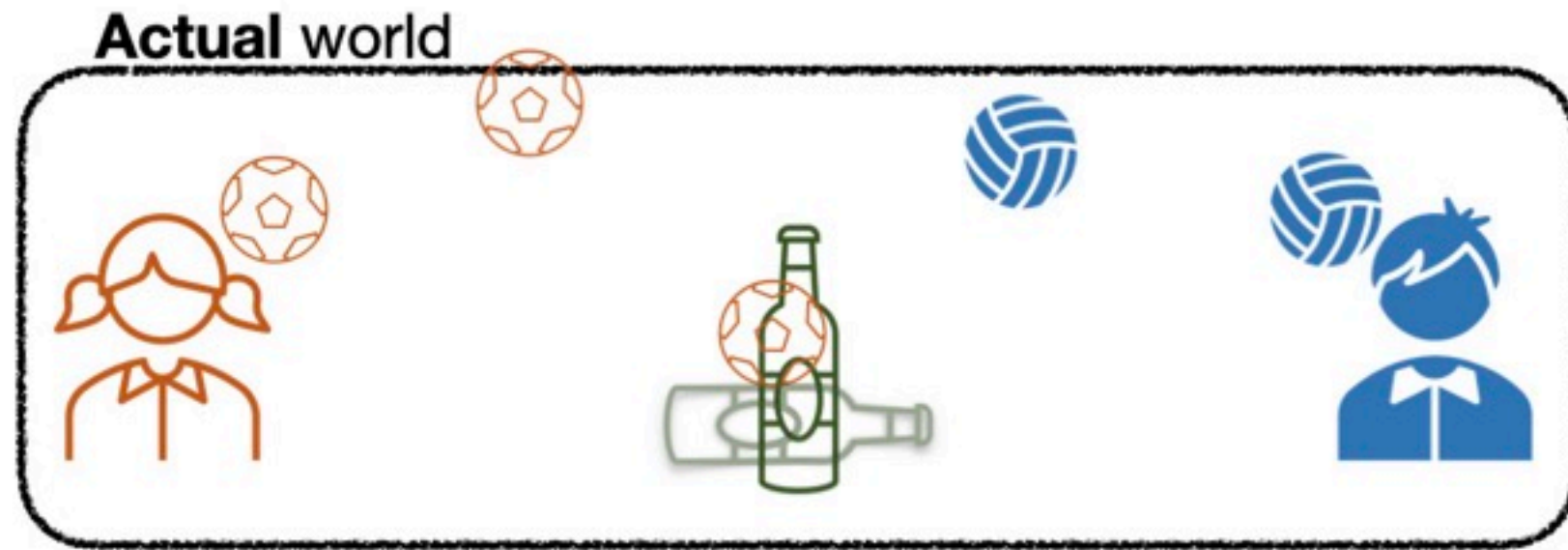
Counterfactual world



Actual Causality



Actual Causality



Actual Causality

Actual world



Counterfactual world

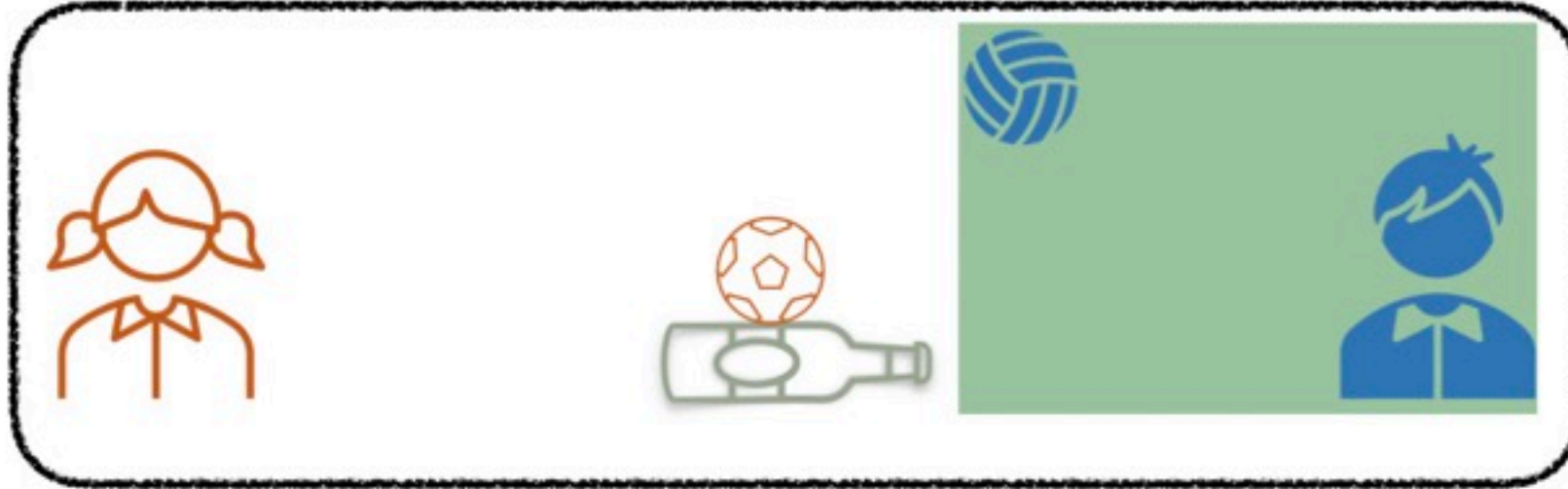


Billy's Ball breaks the bottle!

Actual Causality

Contingencies

Actual world



preemption of causes

Counterfactual world + contingency



Actual Causality



AC1: the cause appears in the actual world



AC2: for every counterfactual world there exists a contingency where **effect** does not hold



AC3: this is a minimal cause

Overview

Halpern and Pearl's actual causality

Hyperproperties and how they allow us to talk about causality



Actual causality for reactive systems

Causes as sets of events

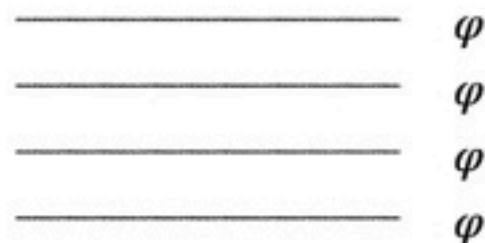
Causes as temporal properties

Algorithmic challenge:
How to compute causes?

Semantic challenge:
How to define causes?

Hyperproperties

- Extend trace properties (e.g., in LTL) to system properties



Linear Temporal Logic – LTL

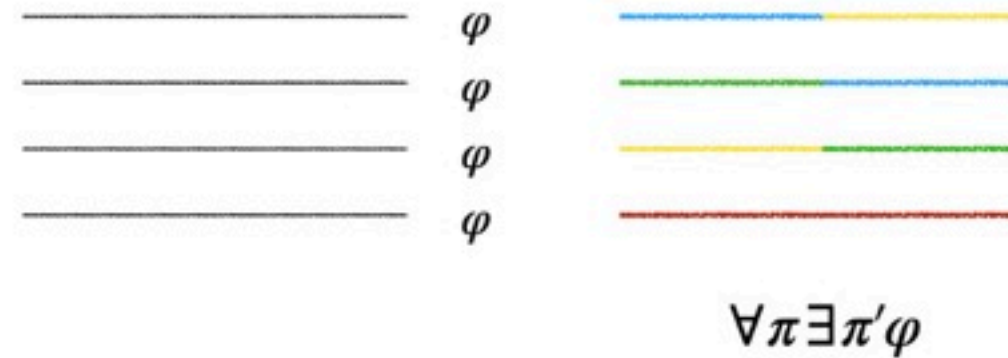
$\square p$ – p holds at every timepoint

$\diamond p$ – p eventually holds

$\bigcirc p$ – p holds at the next timepoint

Hyperproperties

- Extend trace properties (e.g., in LTL) to system properties
- Reason about sets of traces



Linear Temporal Logic – LTL

- $\square p$ – p holds at every timepoint
- $\diamond p$ – p eventually holds
- $\bigcirc p$ – p holds at the next timepoint

Hyperproperties

- Extend trace properties (e.g., in LTL) to system properties
- Reason about sets of traces

HyperLTL – extending LTL with trace quantification

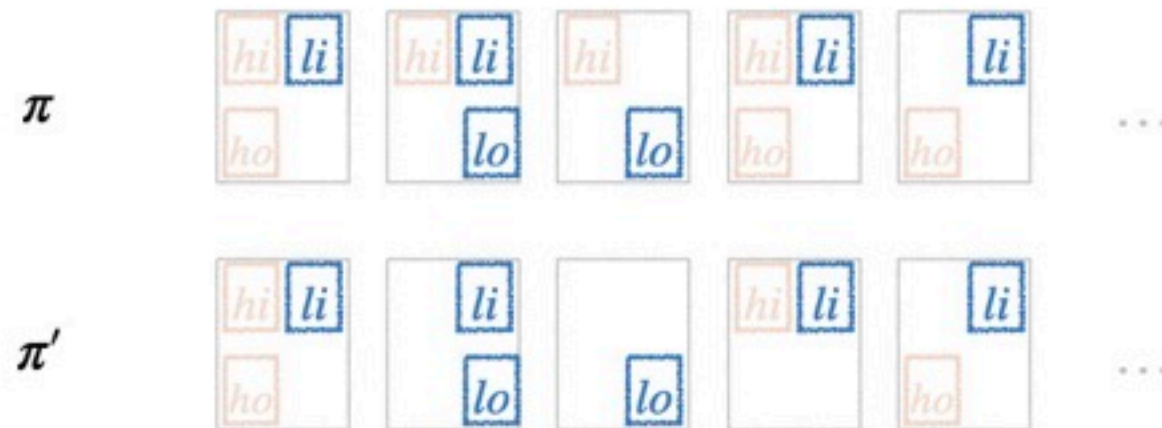
- Observational determinism: $\forall \pi \forall \pi' \square (li_{\pi} \leftrightarrow li_{\pi'}) \rightarrow \square (lo_{\pi} \leftrightarrow lo_{\pi'})$

Hyperproperties

- Extend trace properties (e.g., in LTL) to system properties
- Reason about sets of traces

HyperLTL — extending LTL with trace quantification

- Observational determinism: $\forall \pi \forall \pi' \square (li_{\pi} \leftrightarrow li_{\pi'}) \rightarrow \square (lo_{\pi} \leftrightarrow lo_{\pi'})$



Hyperproperties

- Extend trace properties (e.g., in LTL) to system properties
- Reason about sets of traces



Information-flow properties
[Observational
determinism]



Robustness



Symmetry



Causality

Hyperproperties

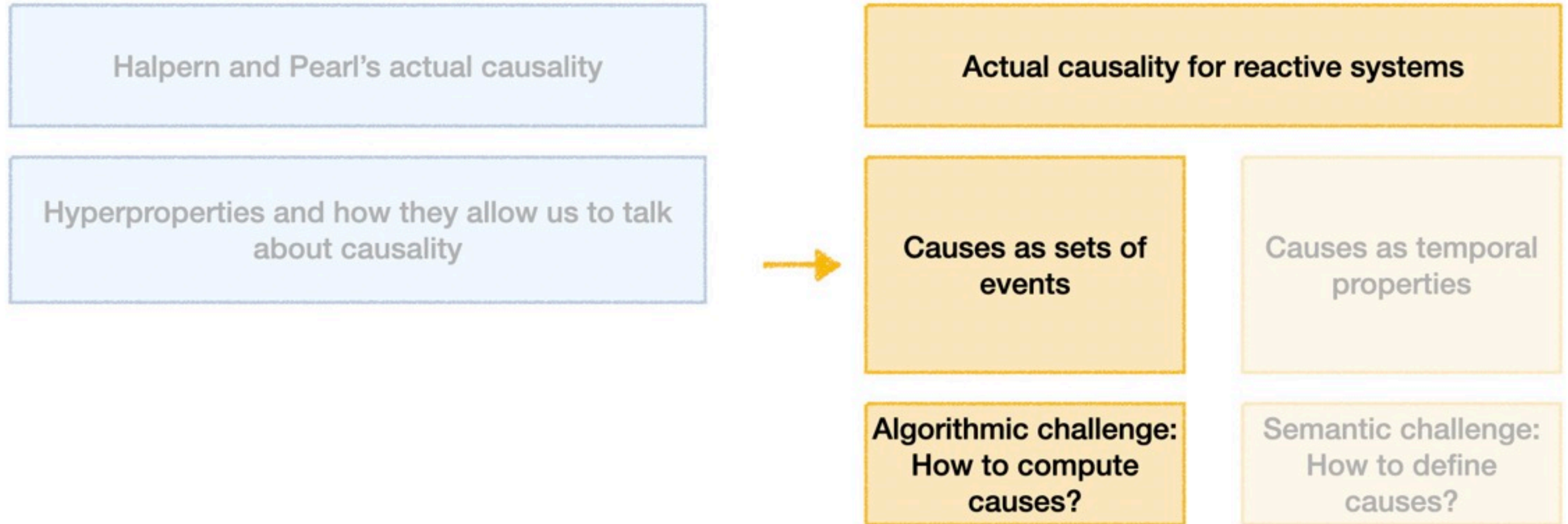
1. Use Hyperproperties to reason about causality



2. Explain Hyperproperty violations

```
Outputfiles
Generated Alger  Generated Dot  Counter Example
1 in_0@0=0
2 in_1@0=1
3 I:F(And(Or(Neg(light0))(Neg(light1)))(Or(light0)(light1)))...284@0=1
4 I:remember_state@0=0
5 l light_0@0=0
6 state[0]_0@0=0
7 state[1]_0@0=0
8 l light_1@0=0
9 state[0]_1@0=0
10 state[1]_1@0=0
11 sink@0=0
12 init@0=0
13 entered_lasso@0=0
14 L:F(And(Or(Neg(light0))(Neg(light1)))(Or(light0)(light1)))...284@0=0
15 l0_copy@0=0
16 l1_copy@0=0
17 l2_copy@0=0
18 l3_copy@0=0
19 l4_copy@0=0
20 l5_copy@0=0
21 l6_copy@0=0
22 l7_copy@0=0
23 L_MH:F(And(Or(Neg(light0))(Neg(light1)))(Or(light0)(light1)))...284@0=0
24 in_0@1=0
25 in_1@1=0
26 I:F(And(Or(Neg(light0))(Neg(light1)))(Or(light0)(light1)))...284@1=0
27 I:remember_state@1=0
28 l light_0@1=1
29 state[0]_0@1=0
30 state[1]_0@1=1
31 l light_1@1=0
32 state[0]_1@1=1
33 state[1]_1@1=0
34 sink@1=0
35 init@1=1
36 entered_lasso@1=0
37 L:F(And(Or(Neg(light0))(Neg(light1)))(Or(light0)(light1)))...284@1=1
```

Overview



Overview

For trace properties:

Explaining Counterexamples Using Causality.

Beer, Ben-David, Chockler, Orni, and Trefler. (CAV 2009).

For Hyperproperties:

Explaining Hyperproperty Violations

Coenen, Dachsel, Finkbeiner, Frenkel, Hahn, Horak, Metzger and Siber. (CAV 2022)



Actual causality for reactive systems

Causes as sets of events

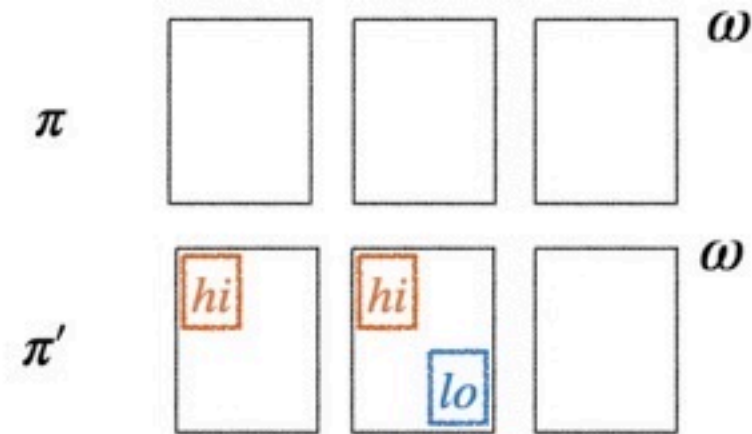
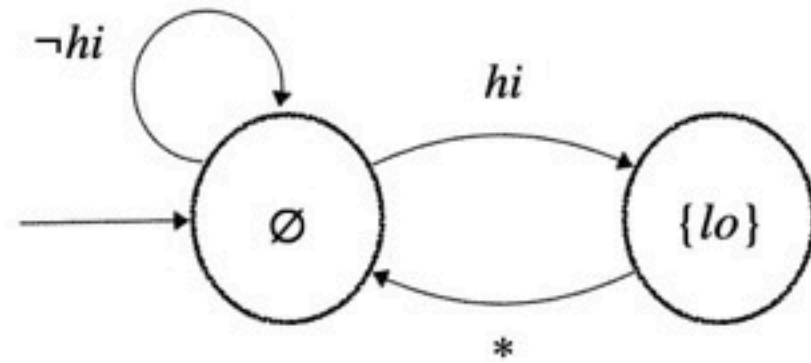
Causes as temporal properties

**Algorithmic challenge:
How to compute causes?**

Semantic challenge:
How to define causes?

Explaining Hyperproperty Violations

$$\forall \pi \forall \pi' \square (li_{\pi} \leftrightarrow li_{\pi'}) \rightarrow \square (lo_{\pi} \leftrightarrow lo_{\pi'})$$



WHY?

Actual Causality



SAT: the cause appears in the actual world



CF: for every counterfactual world there exists a contingency where **effect** does not hold



MIN: this is a minimal cause

Actual Causality for Hyperproperties



SAT



CF : for every counterfactual world
there exists a contingency where **effect**
does not hold



MIN

Actual Causality for Hyperproperties

\forall^* prefix:

$$\psi = \forall \pi_1 \forall \pi_2 \exists \pi'_1 \exists \pi'_2 . \varphi$$

$$\neg \psi = \exists \pi_1 \exists \pi_2 \forall \pi'_1 \forall \pi'_2 . \neg \varphi$$



- Effect: a violation of a **Hyperproperty** ψ

Actual Causality for Hyperproperties

\forall^* prefix:

$$\psi = \forall \pi_1 \forall \pi_2 \exists \pi'_1 \exists \pi'_2 . \varphi$$

$$\neg \psi = \exists \pi_1 \exists \pi_2 \forall \pi'_1 \forall \pi'_2 . \neg \varphi$$



- Effect: a violation of a **Hyperproperty** ψ



- Actual World: a set Γ of counterexample traces

Lasso-shaped:
A finite representation
of infinite traces

Actual Causality for Hyperproperties

\forall^* prefix:

$$\psi = \forall \pi_1 \forall \pi_2 \exists \pi'_1 \exists \pi'_2 . \varphi$$

$$\neg \psi = \exists \pi_1 \exists \pi_2 \forall \pi'_1 \forall \pi'_2 . \neg \varphi$$



- Effect: a violation of a **Hyperproperty** ψ



- Actual World: a set Γ of counterexample traces

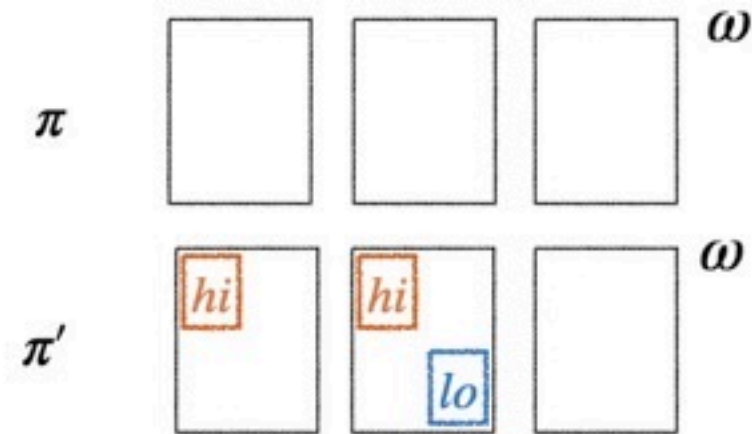
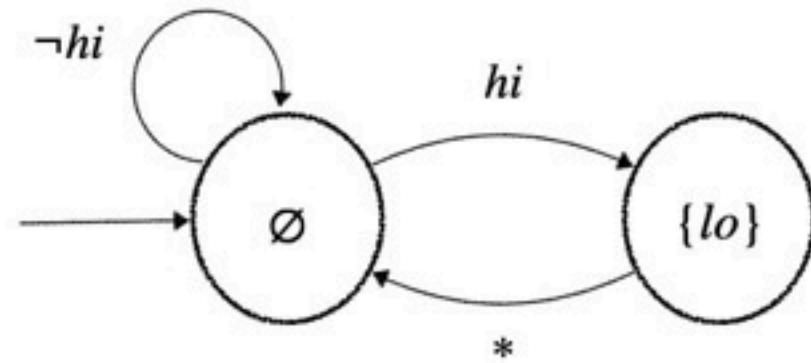


- Cause: set of events on the set of traces

Lasso-shaped:
A finite representation
of infinite traces

Explaining Hyperproperty Violations

$$\forall \pi \forall \pi' \square (li_{\pi} \leftrightarrow li_{\pi'}) \rightarrow \square (lo_{\pi} \leftrightarrow lo_{\pi'})$$

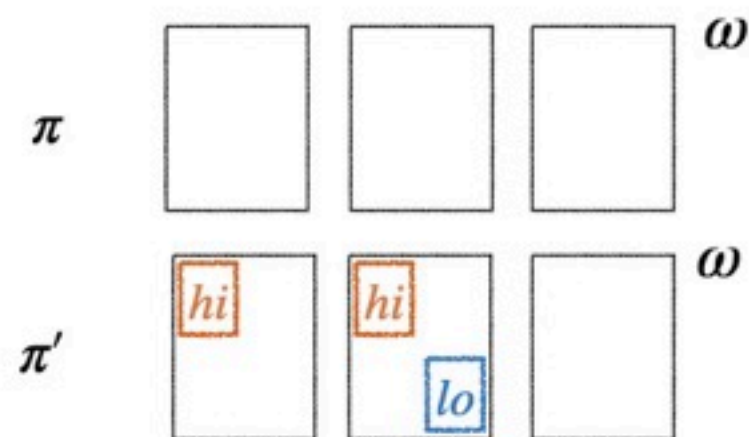
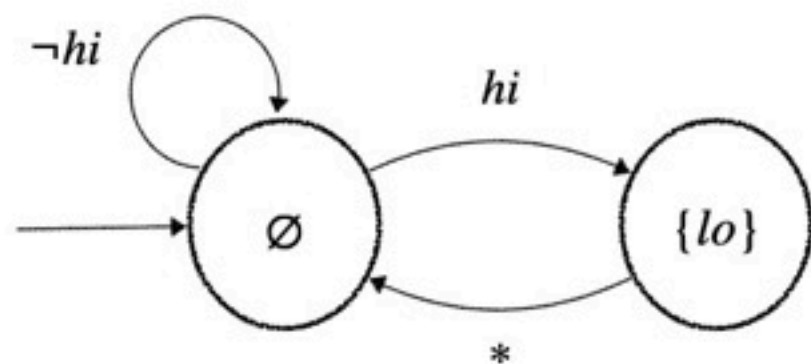


WHY?

Explaining Hyperproperty Violations

CF: \forall counterfactual \exists contingency s.t. φ holds

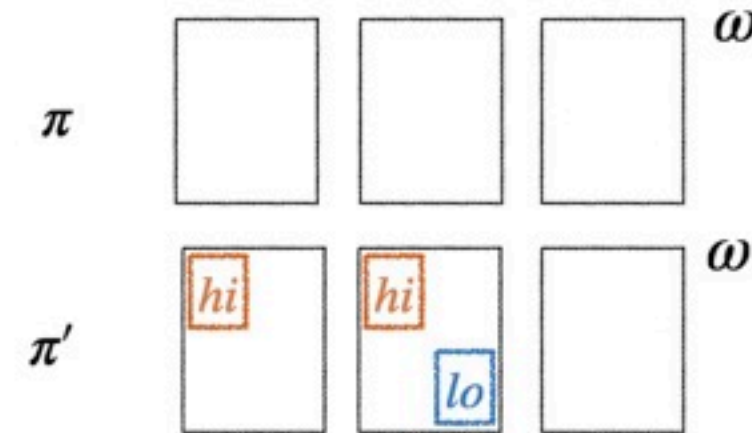
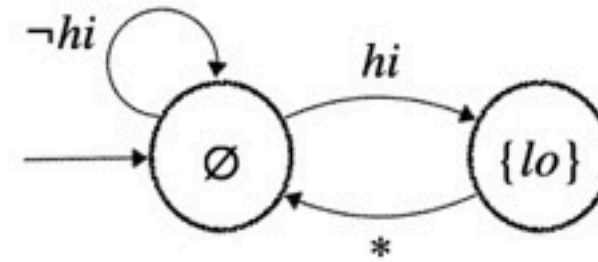
$$\forall \pi \forall \pi' \boxed{\square (li_{\pi} \leftrightarrow li_{\pi'}) \rightarrow \square (lo_{\pi} \leftrightarrow lo_{\pi'})}^{\varphi}$$



Explaining Hyperproperty Violations

CF: \forall counterfactual \exists contingency s.t. φ holds

$$\forall \pi \forall \pi' \square (li_{\pi} \leftrightarrow li_{\pi'}) \rightarrow \square (lo_{\pi} \leftrightarrow lo_{\pi'})$$



$$C = \{ \langle hi, 0, \pi' \rangle \}$$

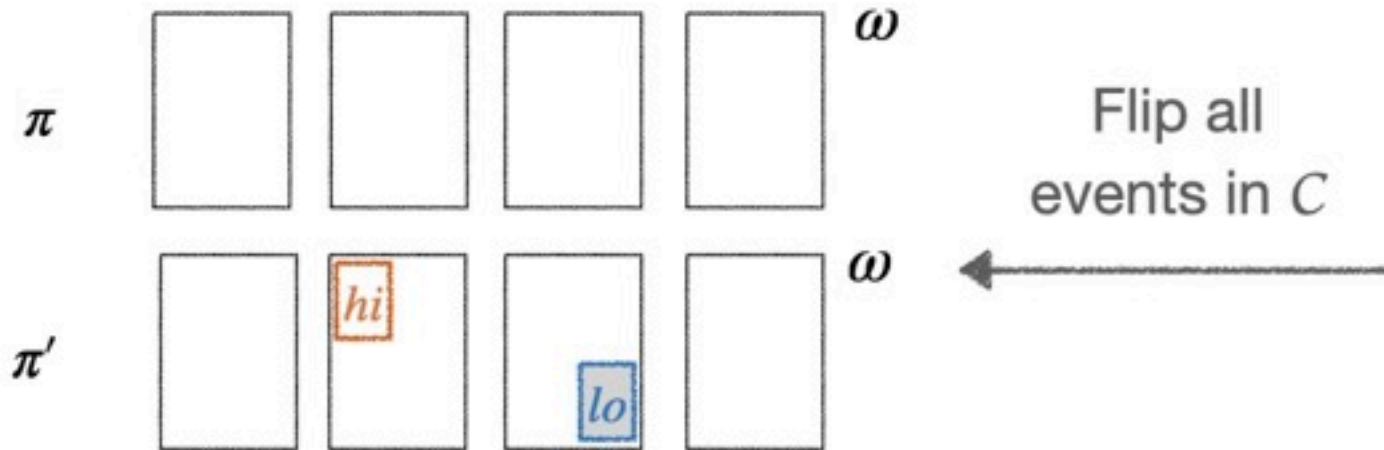
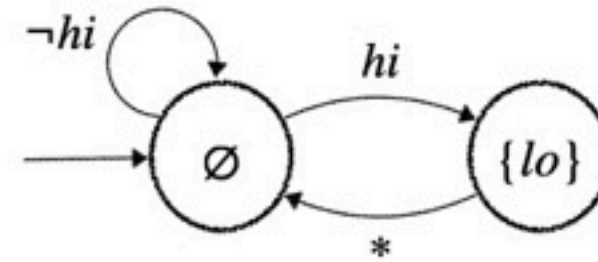
$$\Gamma = (\pi, \pi')$$

Cause - set of events

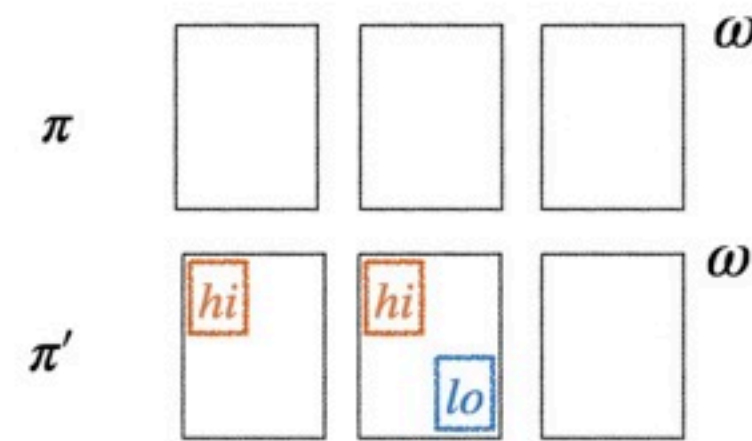
Explaining Hyperproperty Violations

CF: \forall counterfactual \exists contingency s.t. φ holds

$$\forall \pi \forall \pi' \square (li_{\pi} \leftrightarrow li_{\pi'}) \rightarrow \square (lo_{\pi} \leftrightarrow lo_{\pi'})$$



$intervene(\Gamma, C, \emptyset)$



$$C = \{ \langle hi, 0, \pi' \rangle \}$$

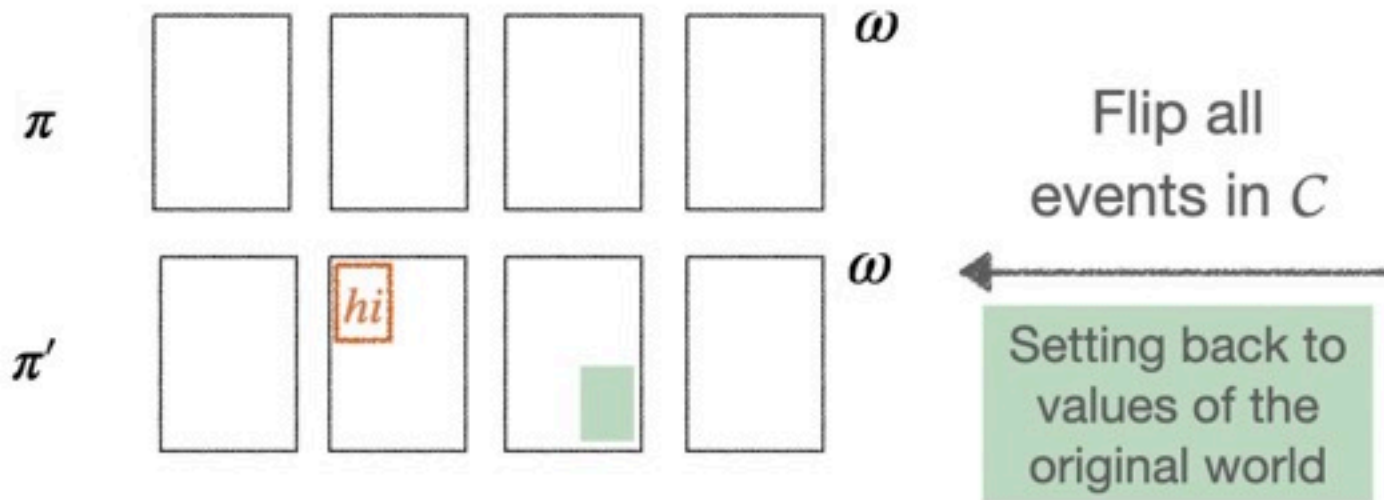
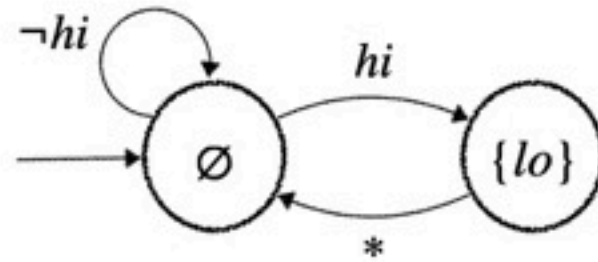
$$\Gamma = (\pi, \pi')$$

Cause - set of events

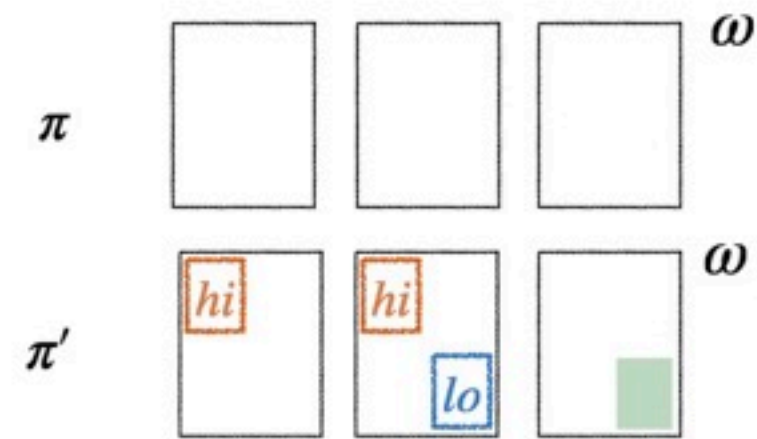
Explaining Hyperproperty Violations

CF: \forall counterfactual \exists contingency s.t. φ holds

$$\forall \pi \forall \pi' \boxed{\square (li_{\pi} \leftrightarrow li_{\pi'}) \rightarrow \square (lo_{\pi} \leftrightarrow lo_{\pi'})} \checkmark$$



$$intervene(\Gamma, C, \{\langle lo, 2, \pi' \rangle\})$$

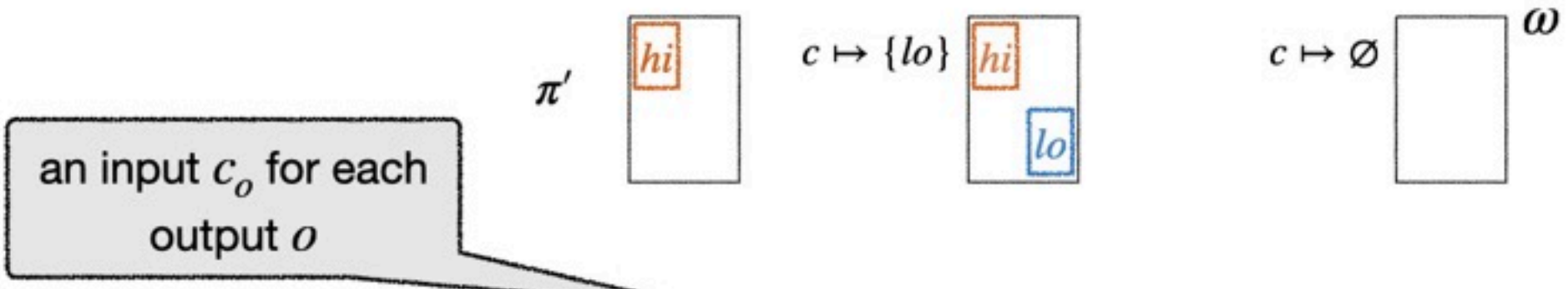
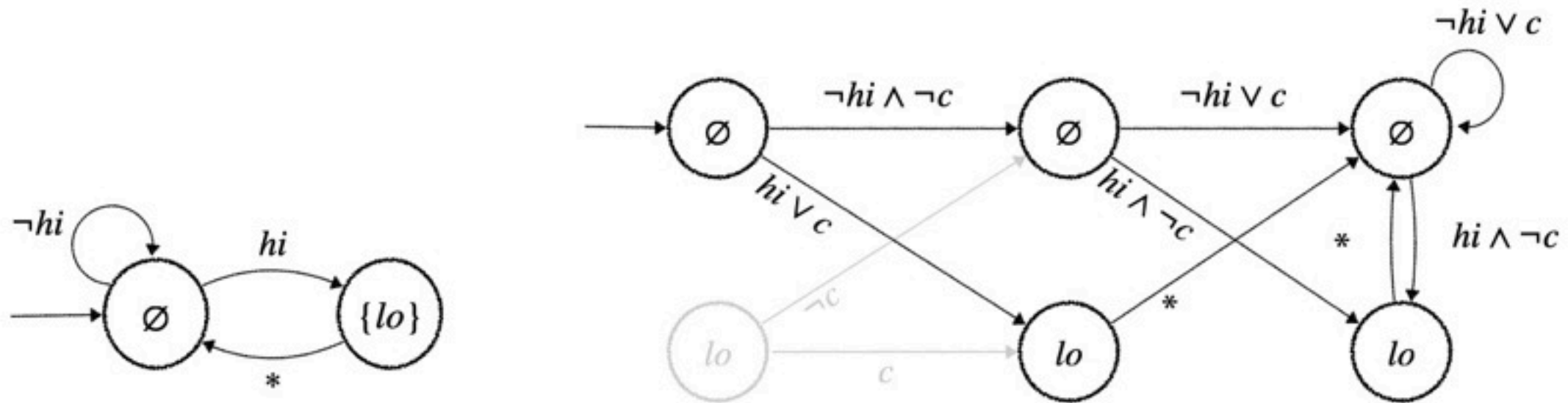


$$C = \{\langle hi, 0, \pi' \rangle\}$$

$$\Gamma = (\pi, \pi')$$

Cause - set of events

Computing Contingencies



Counterfactual automaton: additional inputs $[c]$ to set a contingency

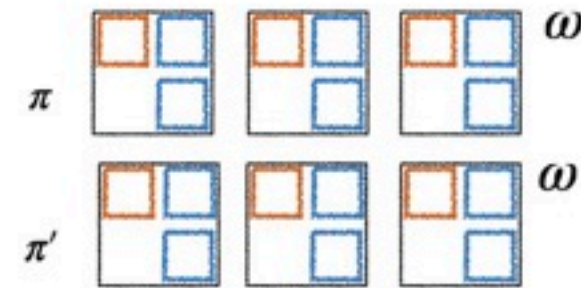
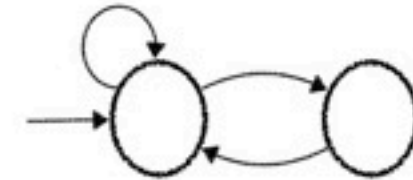
Actual Causality for Hyperproperties

Find C such that

SAT: $\Gamma \models C$

CF: \forall counterfactual \exists contingency s.t. φ holds

MIN: no subset of C satisfies SAT & CF



$$\forall \pi \forall \pi' \square (li_{\pi} \leftrightarrow li'_{\pi}) \rightarrow \square (lo_{\pi} \leftrightarrow lo'_{\pi})$$

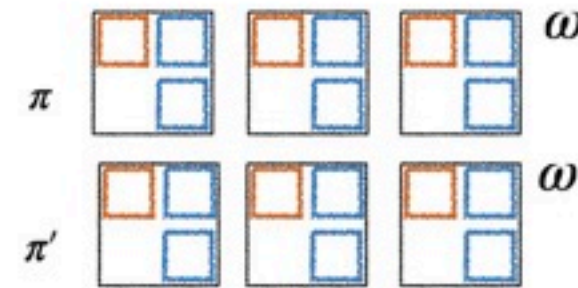
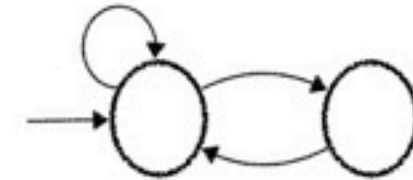
Actual Causality for Hyperproperties

Find C such that

SAT: $\Gamma \models C$

CF: \forall counterfactual \exists contingency s.t. φ holds

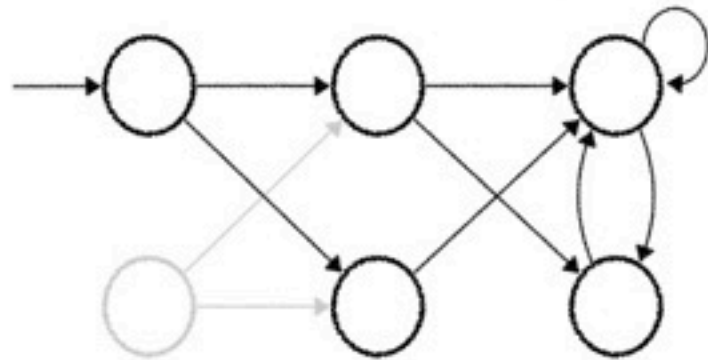
MIN: no subset of C satisfies SAT & CF



$$\forall \pi \forall \pi' \square (li_{\pi} \leftrightarrow li'_{\pi}) \rightarrow \square (lo_{\pi} \leftrightarrow lo'_{\pi})$$

Finding a cause as a hyperproperty

$$\exists \pi_1 \exists \pi_2 . \forall \pi'_1 \forall \pi'_2 . \psi_{cause}$$



Events on π_1, π_2 correspond to the cause
 π'_1, π'_2 represent other possible (not minimal)
 causes

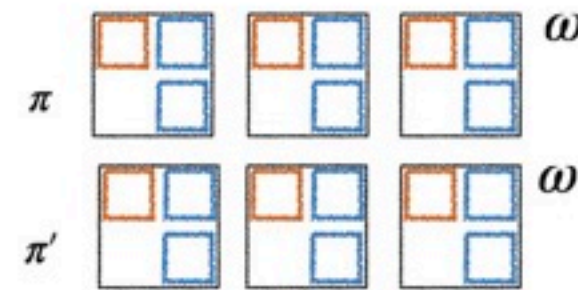
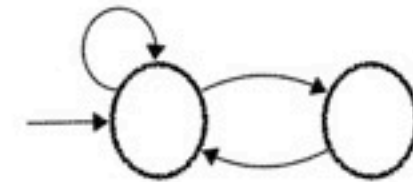
Actual Causality for Hyperproperties

Find C such that

SAT: $\Gamma \models C$

CF: \forall counterfactual \exists contingency s.t. φ holds

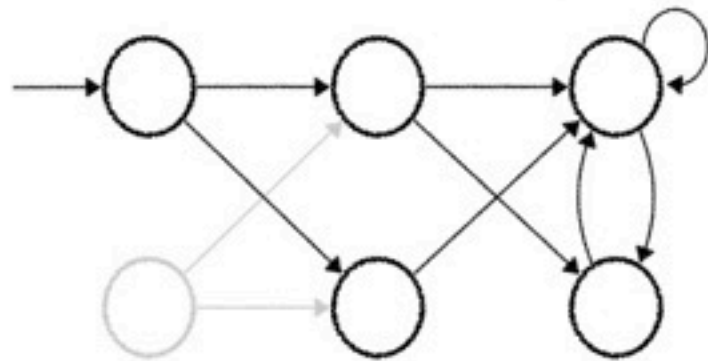
MIN: no subset of C satisfies SAT & CF



$$\forall \pi \forall \pi' \square (li_{\pi} \leftrightarrow li'_{\pi}) \rightarrow \square (lo_{\pi} \leftrightarrow lo'_{\pi})$$

Finding a cause as a hyperproperty

$$\exists \pi_1 \exists \pi_2 . \forall \pi'_1 \forall \pi'_2 . \psi_{cause}$$



HyperLTL
model
checking

Events on π_1, π_2 correspond to the cause
 π'_1, π'_2 represent other possible (not minimal)
causes

Experiments



Instance	$ \Gamma $	$ \varphi $	$\#(C)$	time(ms)
Running example (paper)	10	9	2	55
Security in & out	35	19	8	798
Drone example 1	24	19	5	367
Drone example 2	18	36	3	256
Asymmetric arbiter '19	28	35	10	490
Asymmetric arbiter	72	35	24	1480

```

OutputFiles
Generated Alger  Generated Doc  Counter Example
1 in_000-0
2 in_000-1
3 L-F(And(Or(Neg(Light0))(Neg(Light1)))(Or(Light0)(Light1)))...28498-1
4 remember_state00-0
5 light_000-0
6 state[R]_000-0
7 state[L]_000-0
8 light_100-0
9 state[R]_100-0
10 state[L]_100-0
11 state00-0
12 init00-0
13 entered_1ass00-0
14 L-F(And(Or(Neg(Light0))(Neg(Light1)))(Or(Light0)(Light1)))...28498-0
15 copy00-0
16 copy00-0
17 copy00-0
18 copy00-0
19 copy00-0
20 copy00-0
21 copy00-0
22 copy00-0
23 L_M(F(And(Or(Neg(Light0))(Neg(Light1)))(Or(Light0)(Light1)))...28498-0
24 in_001-0
25 in_001-0
26 L-F(And(Or(Neg(Light0))(Neg(Light1)))(Or(Light0)(Light1)))...28498-0
27 remember_state00-0
28 light_001-0
29 state[R]_001-0
30 state[L]_001-0
31 light_101-0
32 state[R]_101-0
33 state[L]_101-0
34 state01-0
35 init01-0
36 entered_1ass01-0
37 L-F(And(Or(Neg(Light0))(Neg(Light1)))(Or(Light0)(Light1)))...28498-1
    
```



$\{ \langle hi, 0, \pi' \rangle \}$
 $\{ \langle \neg hi, 0, \pi \rangle \}$

Overview

Halpern and Pearl's actual causality

Hyperproperties and how they allow us to talk about causality

Counterfactual automaton &
Reduction to HyperLTL model-checking

Actual causality for reactive systems

Causes as sets of events

Causes as temporal properties

**Algorithmic challenge:
How to compute causes?**

Semantic challenge:
How to define causes?

Overview

Halpern and Pearl's actual causality

Hyperproperties and how they allow us to talk about causality

Actual causality for reactive systems

Causes as sets of events

Causes as temporal properties ←

Algorithmic challenge:
How to compute causes?

**Semantic challenge:
How to define causes?**

Overview

Restricted temporal properties:

Causality Checking for Complex System Models. Leitner-Fischer, Leue. (VMCAI 2013)

General LTL formulas:

- *Temporal Causality in Reactive Systems.* Coenen, Finkbeiner, Frenkel, Hahn, Metzger and Siber. (ATVA 2022)
- *Checking and Sketching Causes on Temporal Sequences.* Beutner, Finkbeiner, Frenkel, Siber. (ATVA 2023)
- *Synthesis of Temporal Causality.* Finkbeiner, Frenkel, Metzger, Siber. (CAV 2024)

Actual causality for reactive systems

Causes as sets of events

Causes as temporal properties

Algorithmic challenge:
How to compute causes?

Semantic challenge:
How to define causes?



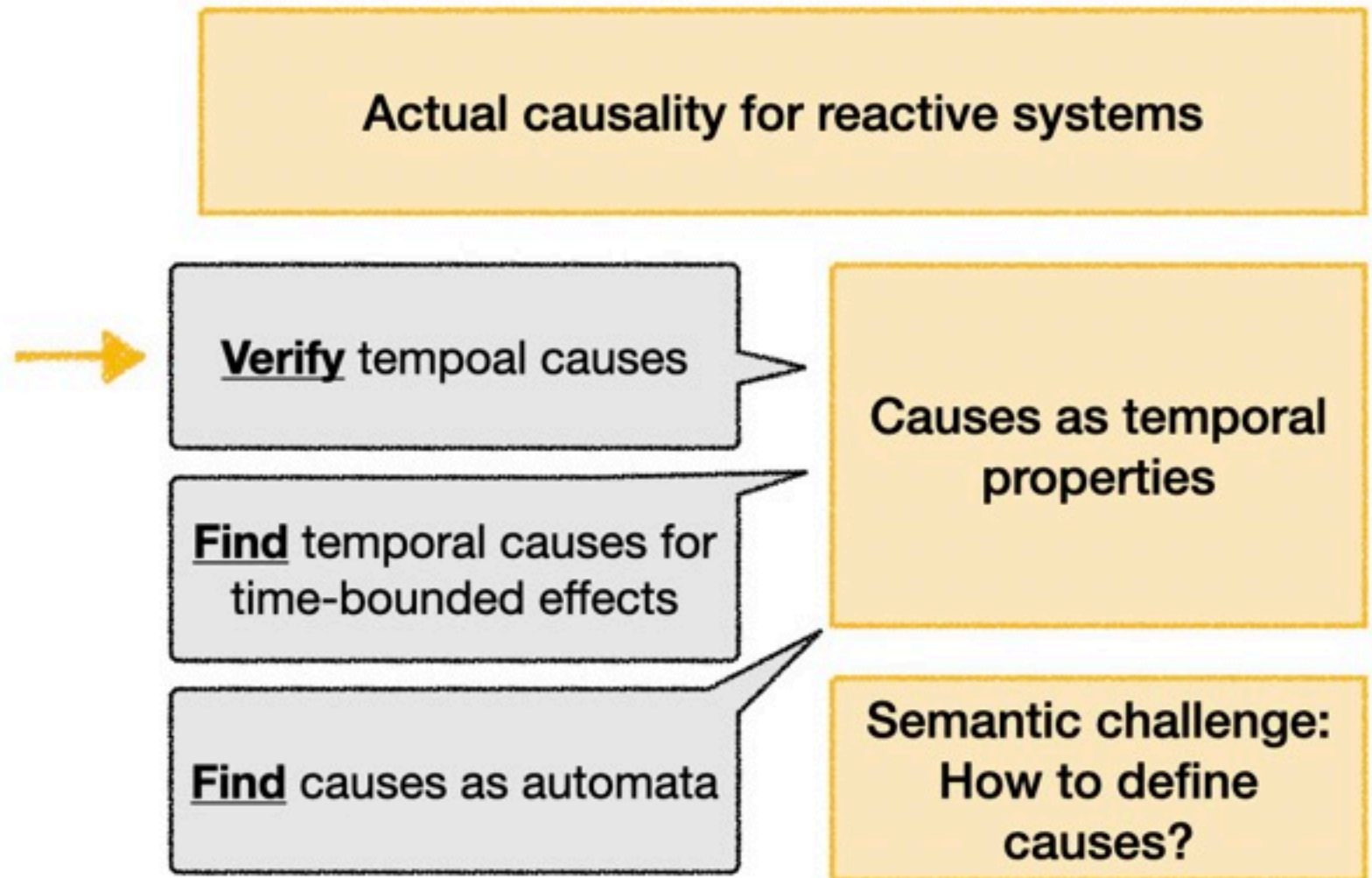
Overview

Restricted temporal properties:




Causality Checking for Complex System Models. Leitner-Fischer, Leue. (VMCAI 2013)

General LTL formulas:

- *Temporal Causality in Reactive Systems.* Coenen, Finkbeiner, Frenkel, Hahn, Metzger and Siber. (ATVA 2022)
- *Checking and Sketching Causes on Temporal Sequences.* Beutner, Finkbeiner, Frenkel, Siber. (ATVA 2023)
- *Synthesis of Temporal Causality.* Finkbeiner, Frenkel, Metzger, Siber. (CAV 2024)



Causes as Trace Properties

-  • Effect: a violation of an ω -regular property ψ
-  • Actual World: a counterexample trace
-  • Cause: an ω -regular property

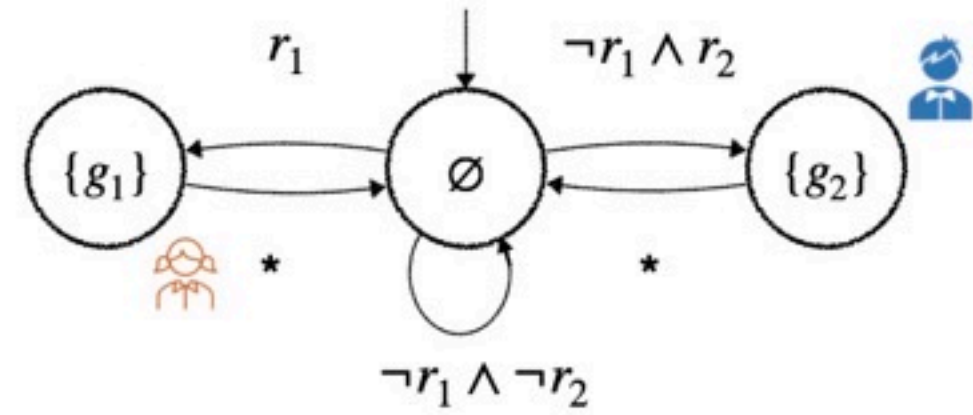
Lasso-shaped

Quantified Propositional Temporal Logic – QPTL

LTL + quantification over propositions

$\exists q . q \wedge \square (q \leftrightarrow \bigcirc \neg q) \wedge \square (q \rightarrow a)$ – “ a holds at every odd position”

Causes as Trace Properties

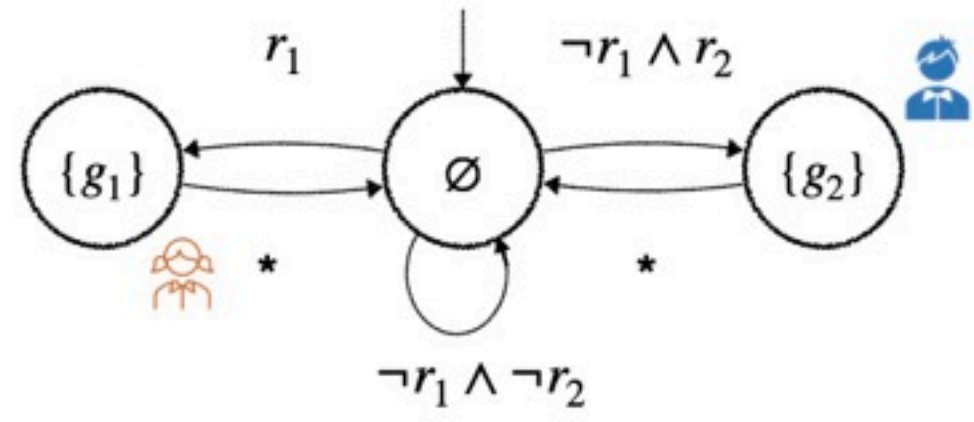


$$\pi \left[\begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \hline \hline \end{array} \left(\begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline g_1 & \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \\ \hline \end{array} \right)^\omega \not\models \diamond g_2$$

Causes as Trace Properties

CF: \forall counterfactual

Closest input sequences
s.t. C doesn't hold



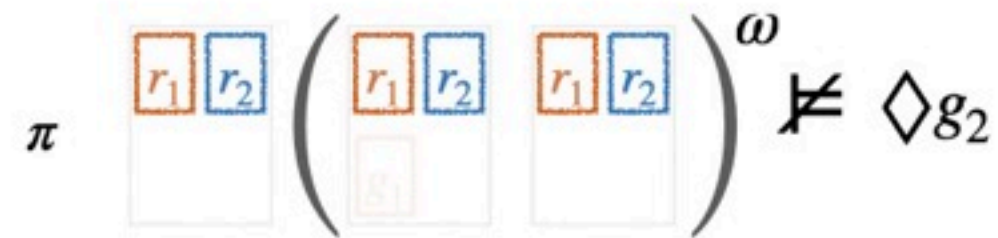
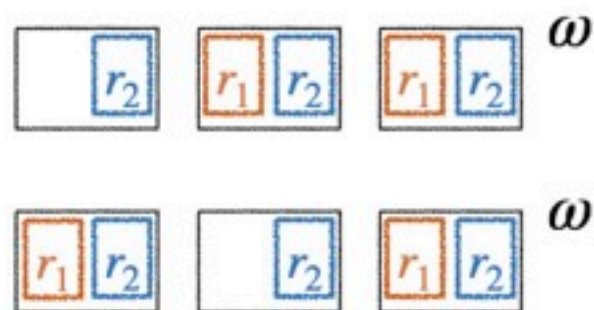
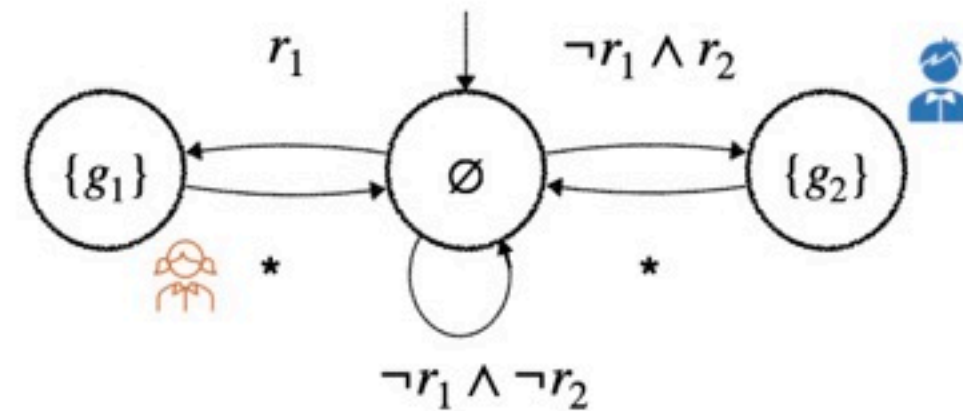
$$\pi \left(\begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \left(\begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \right)^\omega \not\models \diamond g_2$$

Causes as Trace Properties

CF: \forall counterfactual

$$C = r_1 \wedge \bigcirc r_1$$

$$\neg C = \neg r_1 \vee \bigcirc \neg r_1$$



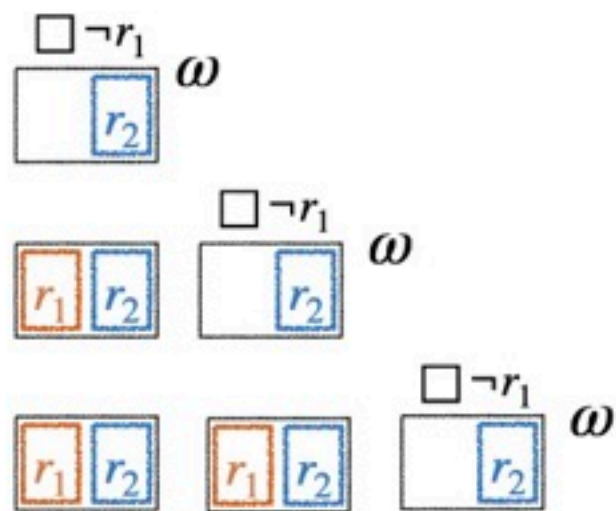
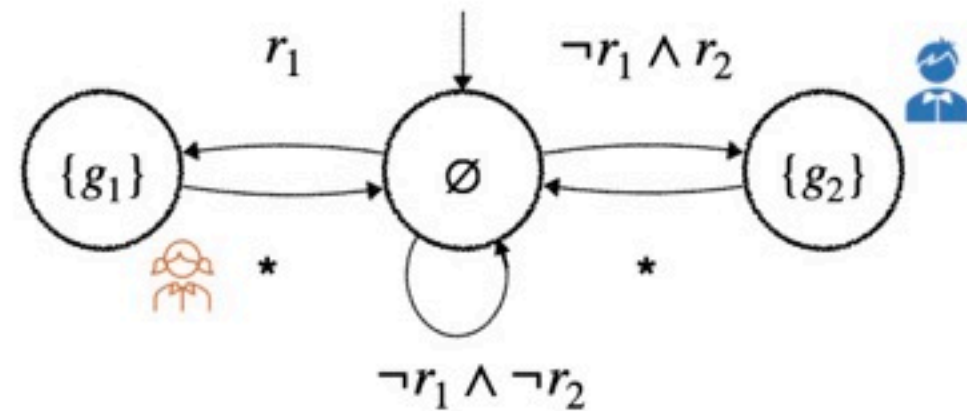
Closest input sequences
s.t. C doesn't hold

Causes as Trace Properties

CF: \forall counterfactual

$$C = \Box \Diamond r_1$$

$$\neg C = \Diamond \Box \neg r_1$$



$$\pi \left(\begin{matrix} r_1 & r_2 \\ \hline & \end{matrix} \left(\begin{matrix} r_1 & r_2 \\ \hline g_1 & \end{matrix} \quad \begin{matrix} r_1 & r_2 \\ \hline & \end{matrix} \right)^\omega \right) \not\models \Diamond g_2$$

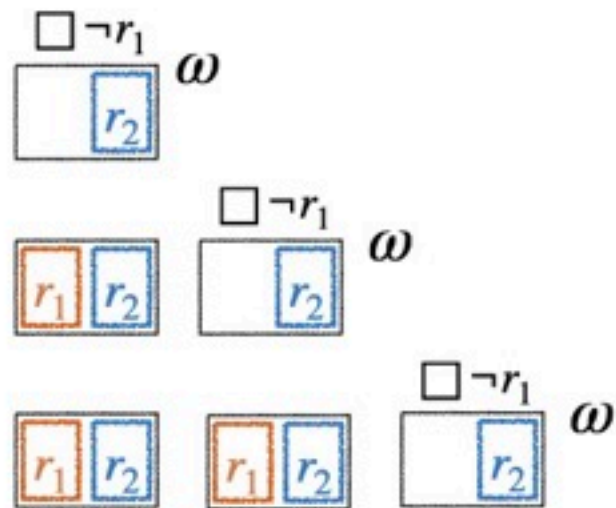
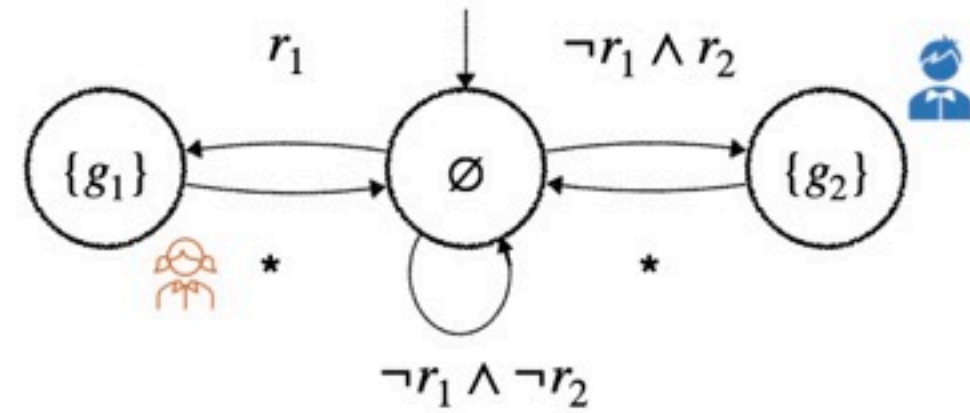
Closest input sequences
s.t. C doesn't hold

Causes as Trace Properties

CF: \forall counterfactual

$$C = \Box \Diamond r_1$$

$$\neg C = \Diamond \Box \neg r_1$$



$$\pi \left(\begin{matrix} [r_1] [r_2] \\ [] \end{matrix} \left(\begin{matrix} [r_1] [r_2] \\ [g_1] \end{matrix} \begin{matrix} [r_1] [r_2] \\ [] \end{matrix} \right) \right)^\omega \not\equiv \Diamond g_2$$

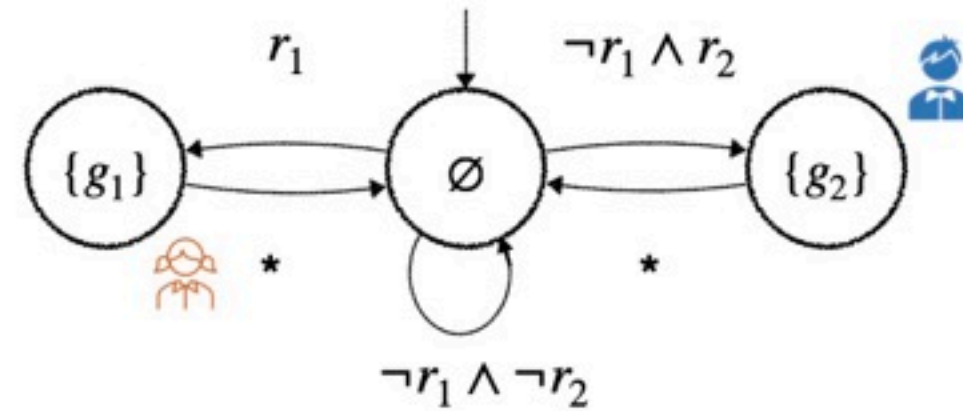
Closest input sequences
s.t. C doesn't hold

Compare traces that have the same **rejection structure**

Causes as Trace Properties

CF: \forall counterfactual

Closest input sequences
s.t. C doesn't hold



HyperQPTL formula

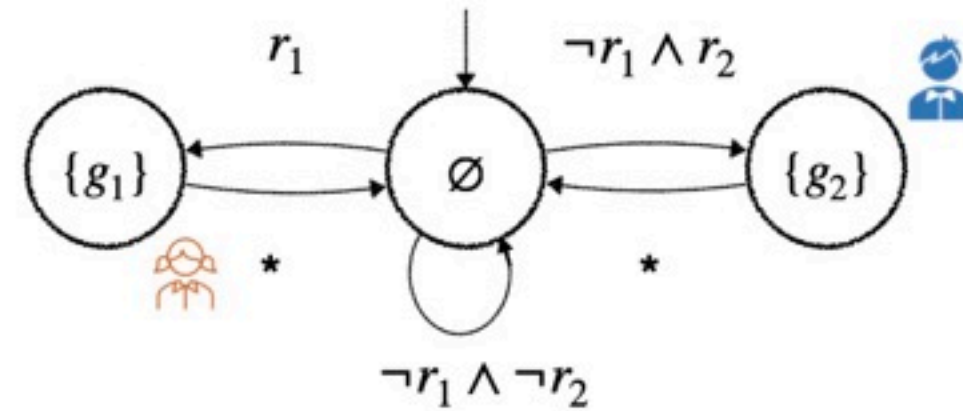
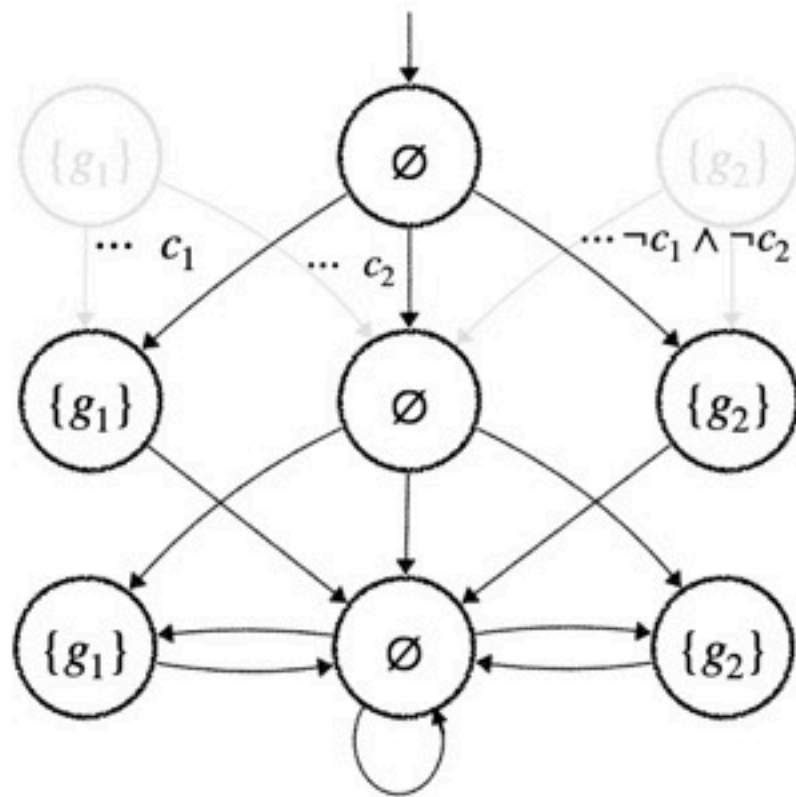
$\Psi_{struct}(\pi_1, \pi_2)$: π_1, π_2 satisfy all sub-formulas of C at the same positions

$$\pi \begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \left(\begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \right)^\omega \not\models \diamond g_2$$

Compare traces that have the same **rejection structure**

Causes as Trace Properties

CF: \forall counterfactual \exists contingency s.t. $\diamond g_2$ holds



$$\pi \left(\begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \left(\begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline g_1 \\ \hline \end{array} \begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \right)^\omega \neq \diamond g_2$$

Counterfactual automaton additional inputs $[c_1, c_2]$ set a contingency

Causes as Trace Properties

Given a candidate
cause C , verify:

SAT: $\pi \models C$

CF: \forall counterfactual \exists contingency s.t. φ holds

MIN: There is no C' such that $C' \rightarrow C$ and $C' \models \text{SAT} \ \& \ \text{CF}$

$\Rightarrow C$ is a cause of
 φ on π

Causes as Trace Properties

Given a candidate cause C , verify:

SAT: $\pi \models C$

CF: \forall counterfactual \exists contingency s.t. φ holds

MIN: There is no C' such that $C' \rightarrow C$ and $C' \models \text{SAT} \ \& \ \text{CF}$

$\Rightarrow C$ is a cause of φ on π

SAT

Verify C on π

CF

Counterfactuals: traces with the same rejection structure
Contingencies: using the counterfactual automaton

MIN

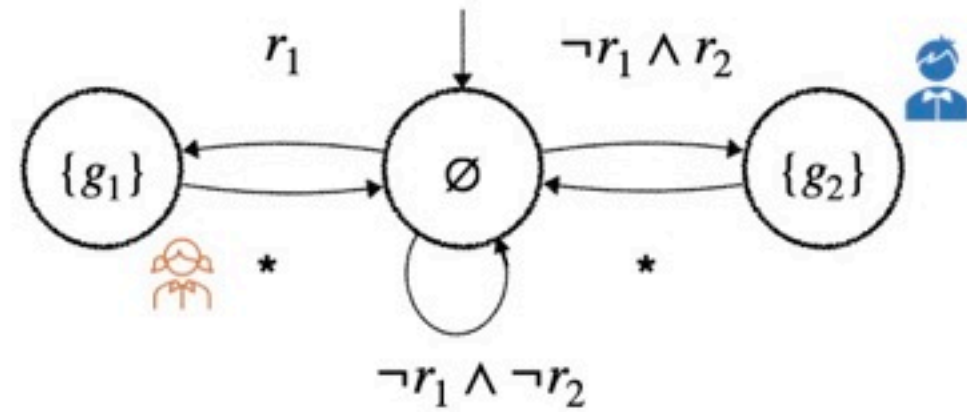
No lasso-shaped trace can be removed from C

Decidable via
HyperQPTL model-
checking!



Example: Unfair Arbiter

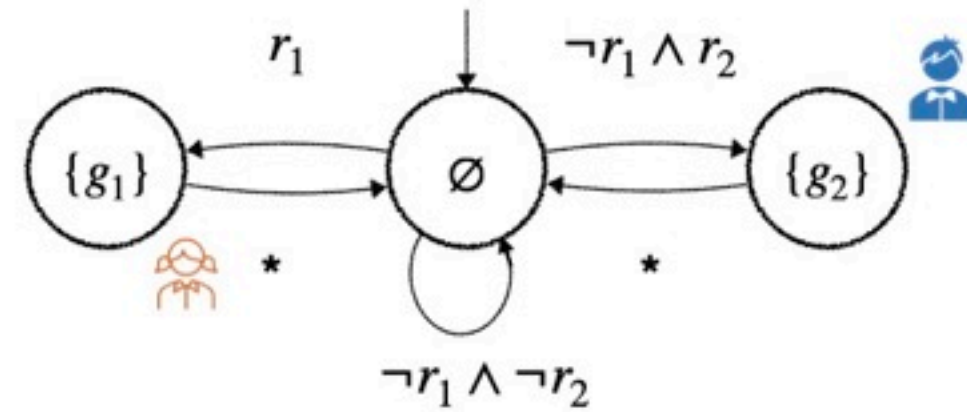
X Is $\Box r_1$ the cause for $\Box \neg g_2$?



$$\pi \begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \left(\begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline g_1 & \end{array} \begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \right)^\omega \not\models \Diamond g_2$$

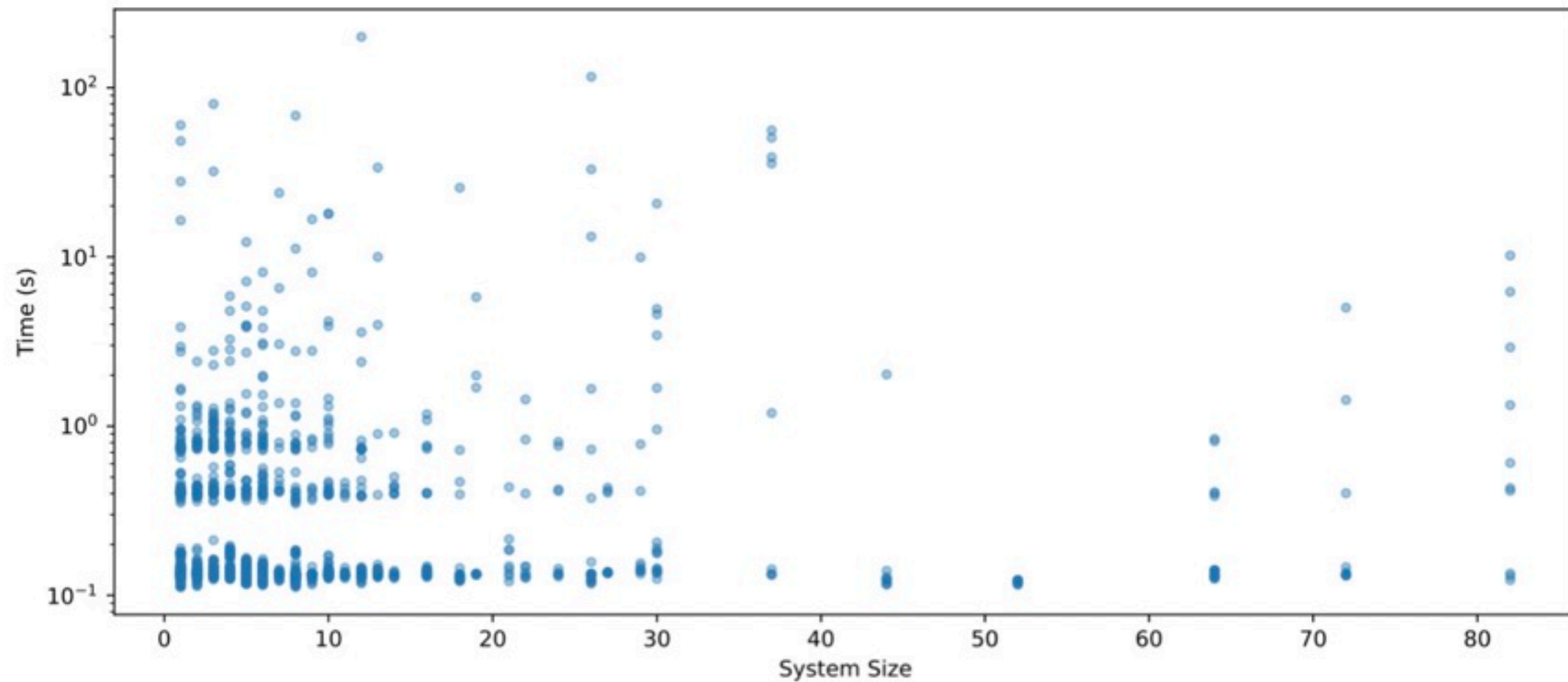
Example: Unfair Arbiter

- ✓ r_1 holds at every odd position
is a cause for $\Box \neg g_2$ on π



$$\pi \quad \begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \left(\begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline g_1 & \end{array} \quad \begin{array}{|c|c|} \hline r_1 & r_2 \\ \hline \end{array} \right)^\omega \not\models \Diamond g_2$$

CATS: a Tool for Checking and Sketching Causes on Temporal Sequences



Randomly generated cause, effect, and trace

Choose with/out contingencies

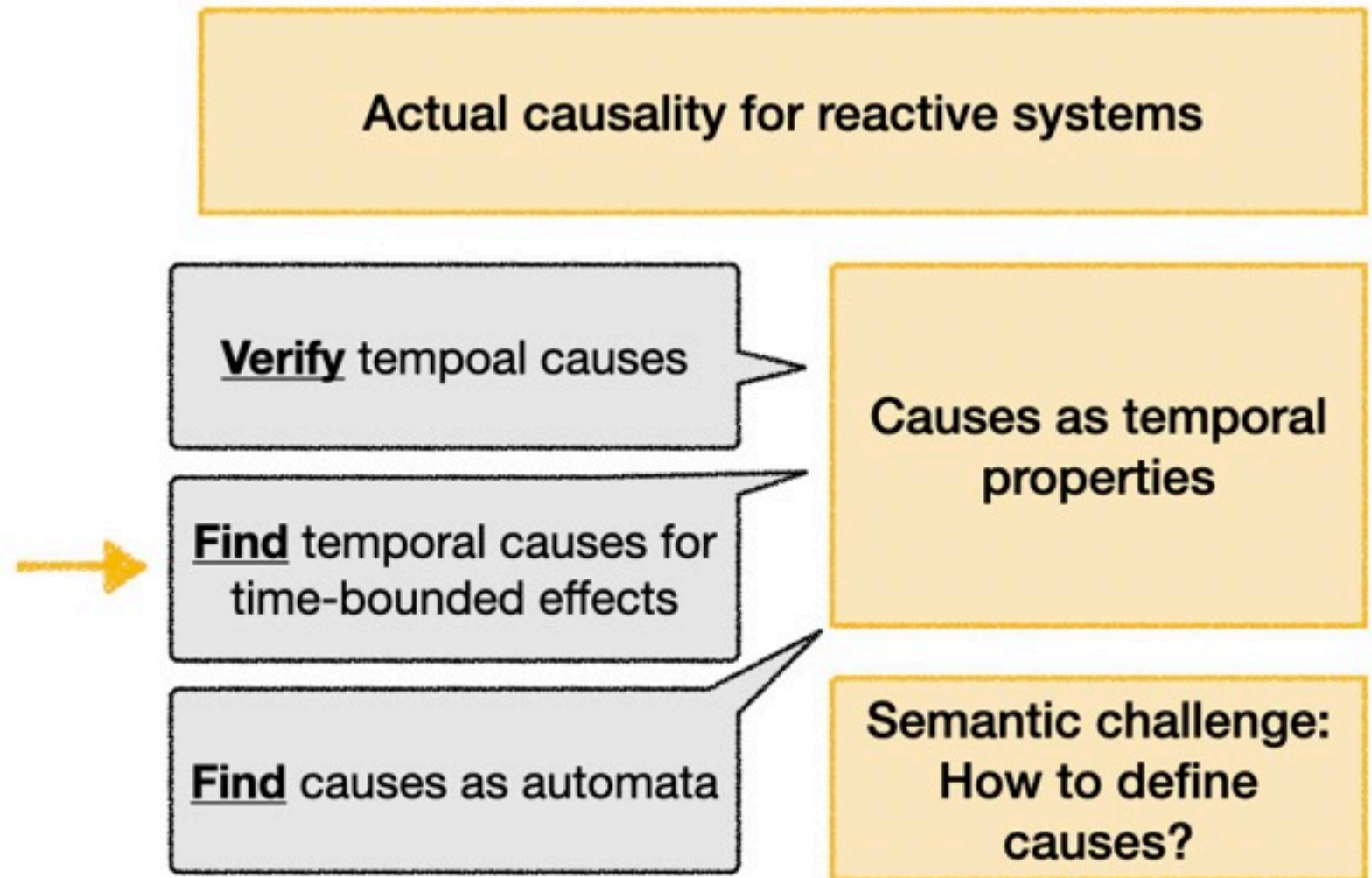
Overview

Restricted temporal properties:

Causality Checking for Complex System Models. Leitner-Fischer, Leue. (VMCAI 2013)

General LTL formulas:

- *Temporal Causality in Reactive Systems.* Coenen, Finkbeiner, Frenkel, Hahn, Metzger and Siber. (ATVA 2022)
- *Checking and Sketching Causes on Temporal Sequences.* Beutner, Finkbeiner, Frenkel, Siber. (ATVA 2023)
- *Synthesis of Temporal Causality.* Finkbeiner, Frenkel, Metzger, Siber. (CAV 2024)



CATS: a Tool for Checking and Sketching Causes on Temporal Sequences

Effect: ○○○ $p \wedge q$

CATS: a Tool for Checking and Sketching Causes on Temporal Sequences

Effect: $\bigcirc\bigcirc\bigcirc p \wedge q$

Cause can only "talk"
about first 3 time-steps

CATS: a Tool for Checking and Sketching Causes on Temporal Sequences

Effect: ○○○ $p \wedge q$

Enumerate all possible causes
and **find** the cause

Cause can only “talk”
about first 3 time-steps

Overview

Restricted temporal properties:

Causality Checking for Complex System Models. Leitner-Fischer, Leue. (VMCAI 2013)

General LTL formulas:

- *Temporal Causality in Reactive Systems.* Coenen, Finkbeiner, Frenkel, Hahn, Metzger and Siber. (ATVA 2022)
- *Checking and Sketching Causes on Temporal Sequences.* Beutner, Finkbeiner, Frenkel, Siber. (ATVA 2023)
- *Synthesis of Temporal Causality.* Finkbeiner, Frenkel, Metzger, Siber. (CAV 2024)



Actual causality for reactive systems

Verify temporal causes

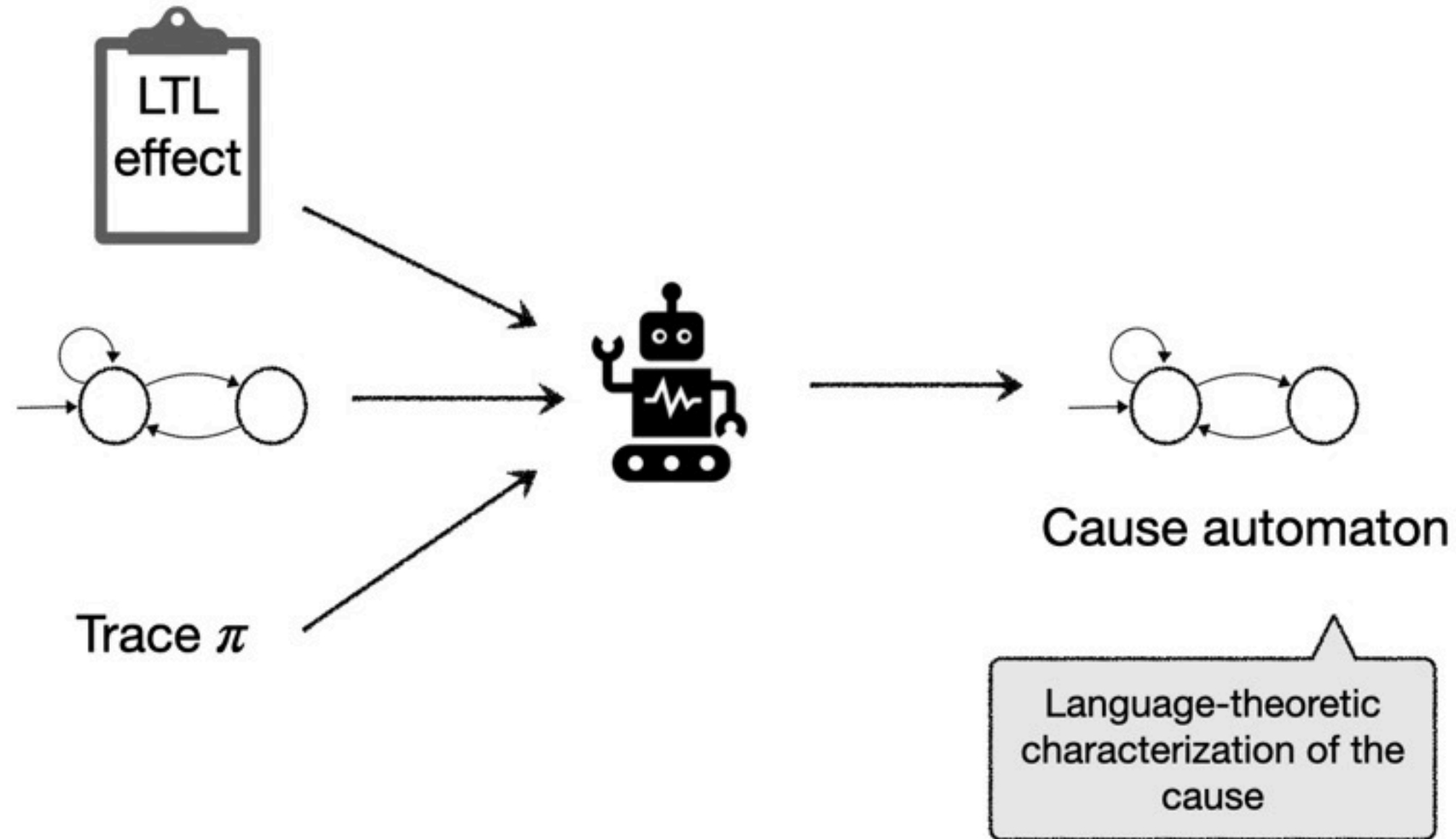
Find temporal causes for time-bounded effects

Find causes as automata

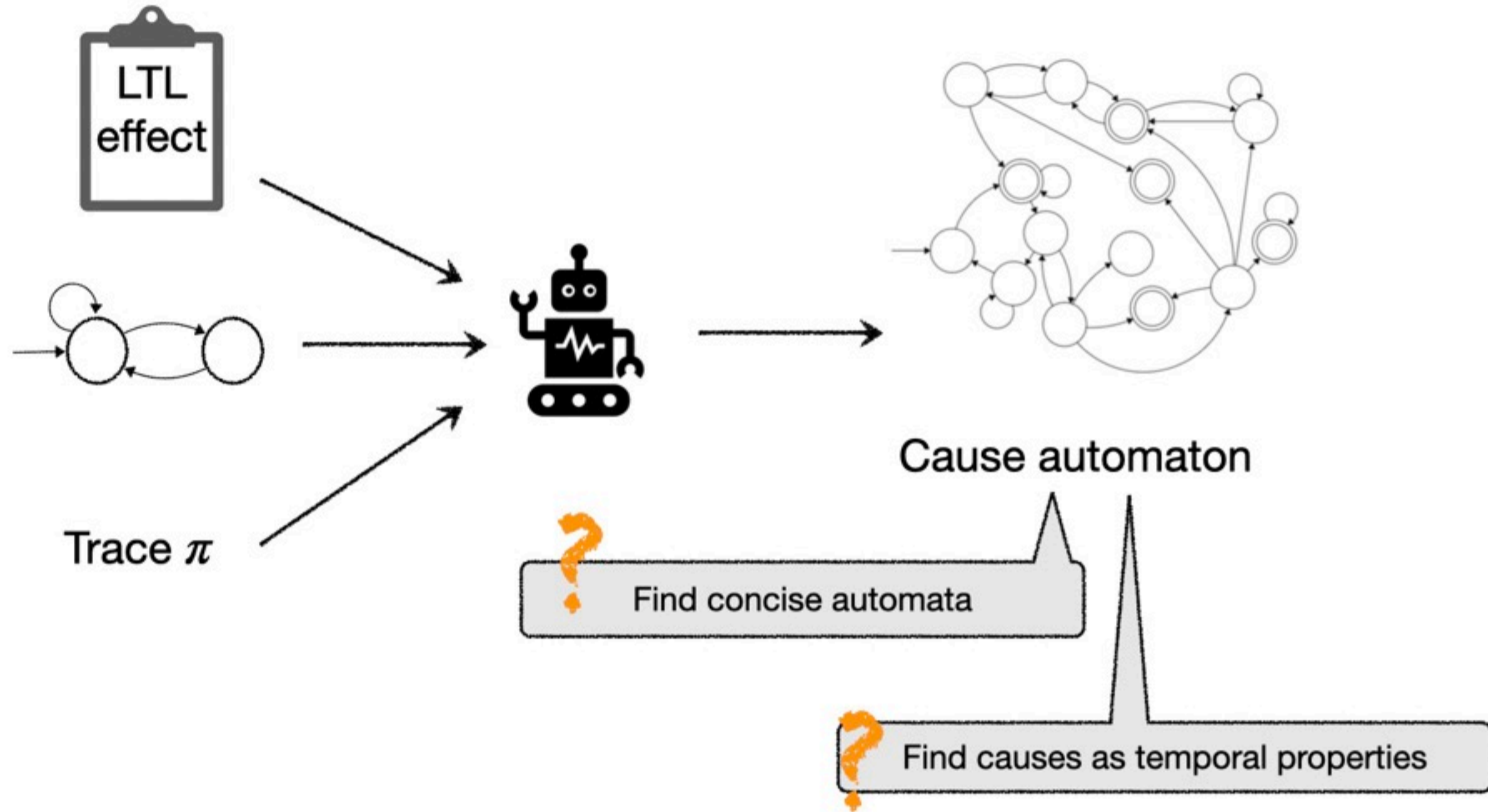
Causes as temporal properties

Semantic challenge:
How to define causes?

Find Causes Using Automata



Find Causes Using Automata



Summary

Actual causality for reactive systems

Formalism for expressing HP causality for reactive systems

Causes as sets of events

Causes as temporal properties

Algorithm for finding all causes for Hyperproperties

Verify temporal causes

Find temporal causes for time-bounded effects

Find causes as automata

Open Questions?

Find concise causes as automata

Find temporal causes for temporal effects

Define temporal causality for hyperproperties

Verify & find temporal causes for hyperproperties

Applications: explainability & repair