

FIRST-ORDER AUTOMATA

Nicola Gigante

Free University of Bozen-Bolzano, Italy

iFM²

April 19, 2024

Udine, Italy

Infinite-state specification and verification

Model-checking and other verification techniques for **finite-state** systems have been extremely successful in the past decades.

However, many real-world scenarios require reasoning over **infinite-state** systems.

There are **many** infinite-state formalisms around.

- timed automata [AD94]
- hybrid automata [Alu+92]
- recursive state machines [Alu+05; BMP10]
- visibly pushdown languages [AM04]
- operator-precedence languages [Dro+17]
- FIFO machines [BZ83]
- counter machines [Woj99]
- Petri nets [Mur89; JK09]
- data-aware systems [DLV19; CGM13; Ghi+23]
- automata over infinite alphabets [Seg06; IX19]
- register automata [KF94]

And many **logics** to talk about them.

- precedence oriented temporal logic [CMP22]
- metric temporal logics [Koy90; LVR22; AH94]
- temporal logics with concrete domains [DQ21]
- temporal logics with arithmetics [Fel+23; Cim+20]
- separation logic [CYO01; BK18; Man20]
- constrained horn clauses [GB19]
- logics on data words [Boj+06; DLN07]
- signal temporal logic [MN04]
- first-order temporal logics [Kon+04]

And many **logics** to talk about them.

- precedence oriented temporal logic [CMP22]
- metric temporal logics [Koy90; LVR22; AH94]
- temporal logics with concrete domains [DQ21]
- temporal logics with arithmetics [Fel+23; Cim+20]
- separation logic [CYO01; BK18; Man20]
- constrained horn clauses [GB19]
- logics on data words [Boj+06; DLN07]
- signal temporal logic [MN04]
- **first-order temporal logics** [Kon+04]

First-order Temporal Logic

First-order temporal logics are expressive formalisms to specify properties of infinite-state systems:

- badly undecidable, however;
- used mostly in really tailored fragments or in **knowledge representation** applications where reasoning is not the main task;

Question

Can we make them **practical**, as a specification language for **verification** tasks, while maintaining a substantial fraction of its **expressive power**?

First-order automata

In the finite-state setting, **automata** have played the crucial role of the algorithmic powerhouse of verification and model checking.

Question

Can we define a similar tool for **first-order** temporal logic?

High-level overview of the talk:

- recap on propositional and first-order temporal logics
- past results: LTL modulo theories
- current directions: first-order automata
- conclusions

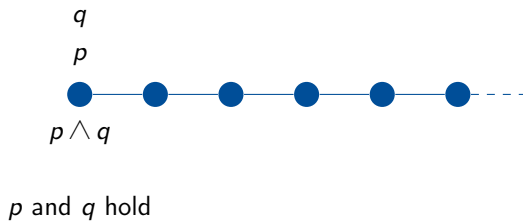
FIRST-ORDER TEMPORAL LOGICS

Let us recap the main concepts of Linear Temporal Logic:

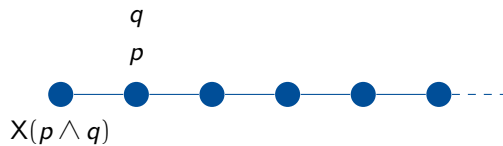
- modal logic interpreted over **discrete linear orders**
- traditionally, **infinite** linear orders
- recently, **finite** models gained attention [DV13]



Linear Temporal Logic

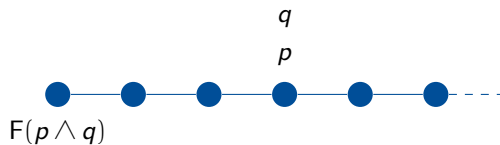


Linear Temporal Logic



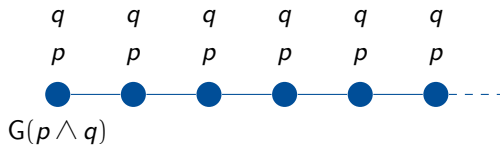
tomorrow p and q hold

Linear Temporal Logic



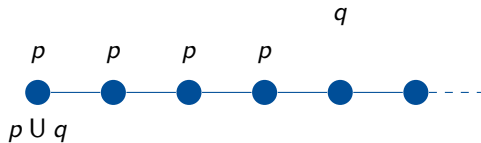
eventually p and q hold

Linear Temporal Logic



p and q hold **always**

Linear Temporal Logic



p holds **until** q holds

Linear Temporal Logic over Finite Traces



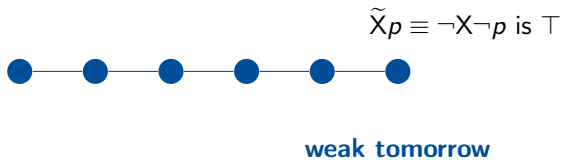
Linear Temporal Logic over Finite Traces



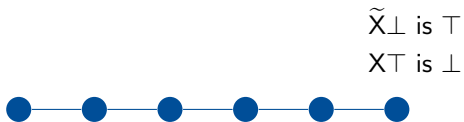
Linear Temporal Logic over Finite Traces



Linear Temporal Logic over Finite Traces



Linear Temporal Logic over Finite Traces



In FOLTL we mix classical first-order logic with LTL temporal operators.

$$\forall x \exists y G(p(x) \rightarrow F(q(x, y))) \quad G(\forall xy (p(x, y) \rightarrow \tilde{X}p(x, y + 1)))$$

$$\forall x X\neg p(x) \wedge X\exists x p(x)$$

FOLTL is interpreted over **sequences of first-order structures**.

- possibly over **multisorted** signatures;
- first-order features interpreted over the current structure;
- temporal operators to move from one structure to another;
- the **domain** of each sort is arbitrary but **constant** throughout the sequence.

FOLTL is famous for being **highly undecidable**, but, for a fragment F such that:

- the underlying FO fragment is decidable; and
- temporal formulas are **monodic**;
 - only one **free variable** inside temporal operators;
- constants in the signature are considered **rigid**;
- then FOLTL(F) is decidable [HWZ00]

Are there decidable fragments of FO?

- monadic fragment
- Bernays-Schönfinkel-Ramsey fragment ($\exists^*\forall^*/ =$)
- Ackermann fragment ($\exists^*\forall\exists^*$)
- Gödel-Kalmár-Schütte fragment ($\exists^*\forall\forall\exists^*$)
- Skolem fragment
- two-variable fragment
- guarded fragment
- separated fragment
- combinations of the above in multi-sorted signatures
- many description logics are decidable FO fragments in disguise

See [Voi19] for a detailed survey.

THE PAST:
LTL MODULO THEORIES

There have been many developments in the world of first-order theorem proving.

- many good solvers (SPASS, Vampire, E, *etc.*);
- not specific to decidable fragments, however.

Can we use these as a foundation of FOLTL reasoning? Yes.

- See *e.g.*, TSPASS, a solver for monodic first-order temporal logic

However, the **monodic** temporal fragment is limited for applications.

Imagine modeling an **update** query on an SQL database:

update R **set** $x = x - 1$ **where** $x > 0$

In FOLTL, it can become something like:

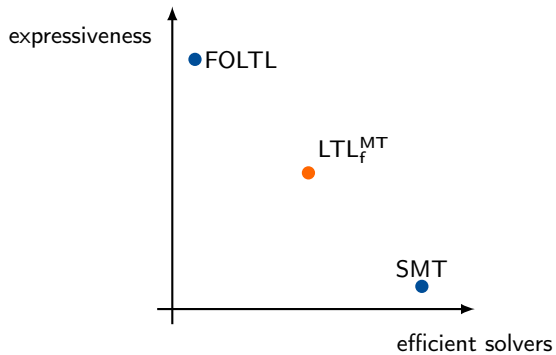
$$\forall k \forall x \left(\begin{array}{l} (R(k, x) \wedge x > 0 \rightarrow X(\neg R(k, x) \wedge R(k, x - 1))) \\ \wedge (R(k, x) \wedge x \leq 0 \rightarrow XR(k, x)) \end{array} \right)$$

We need to accept dealing with **semi-decidability**.

There is another approach at **first-order** reasoning:

- satisfiability modulo theories
- **bottom-up** approach:
 - let's add to SAT solvers bits of tractable first-order reasoning step by step
- SMT solvers are strong at quantifier-free reasoning over specific theories useful for program verification
 - linear integer arithmetics/real arithmetics
 - arrays
 - bitvectors
 - constrained Horn clauses
 - algebraic data types
- sometimes quantification works, as well

LTL_f^{MT} is a sweet spot in the specification of infinite-state properties. [IJCAI 22]



LTL_f^{MT} extends LTL_f by introducing **some** first-order elements.

- let Σ be a first-order signature and T a Σ -theory
- LTL_f^{MT} is interpreted over finite words of T -structures
- predicates and function symbols are **rigid**
- constants are **non-rigid**, *i.e.*, change over time
 - this is **not** covered by existing decidability results about FOLTL

LTL_f^{MT} extends LTL_f by introducing **some** first-order elements.

- terms can refer to the value of constants at the **next** or **previous** state.

$$t := c \mid x \mid f(t_1, \dots, t_n) \mid \bigcirc c \mid \ominus c \mid \triangleleft c \mid \triangleleft c$$

- LTL_f propositions are replaced by first-order sentences

$$\begin{aligned} \lambda &:= p(t_1, \dots, t_n) \mid \neg \lambda \mid \lambda \vee \lambda \mid \lambda \wedge \lambda \mid \exists x \lambda \mid \forall x \lambda \\ \phi &:= \lambda \mid \neg \phi \mid \phi \vee \phi \mid \phi \wedge \phi \mid X\phi \mid \tilde{X}\phi \mid Y\phi \mid Z\phi \\ &\quad \mid \phi U \phi \mid \phi R \phi \mid \phi S \phi \mid \phi T \phi \end{aligned}$$

LTL_f^{MT} extends LTL_f by introducing **some** first-order elements.

- terms can refer to the value of constants at the **next** or **previous** state.

$$t := c \mid x \mid f(t_1, \dots, t_n) \mid \bigcirc c \mid \ominus c \mid \triangleleft c \mid \triangleleft c$$

- LTL_f propositions are replaced by first-order sentences

$$\begin{aligned} \lambda &:= p(t_1, \dots, t_n) \mid \neg \lambda \mid \lambda \vee \lambda \mid \lambda \wedge \lambda \mid \exists x \lambda \mid \forall x \lambda \\ \phi &:= \lambda \mid \neg \phi \mid \phi \vee \phi \mid \phi \wedge \phi \mid X\phi \mid \tilde{X}\phi \mid Y\phi \mid Z\phi \\ &\quad \mid \phi U \phi \mid \phi R \phi \mid \phi S \phi \mid \phi T \phi \end{aligned}$$

LTL_f^{MT} extends LTL_f by introducing **some** first-order elements.

- terms can refer to the value of constants at the **next** or **previous** state.

$$t := c \mid x \mid f(t_1, \dots, t_n) \mid \bigcirc c \mid \ominus c \mid \triangleleft c \mid \triangleleft c$$

- LTL_f propositions are replaced by first-order sentences

$$\begin{aligned} \lambda &:= p(t_1, \dots, t_n) \mid \neg \lambda \mid \lambda \vee \lambda \mid \lambda \wedge \lambda \mid \exists x \lambda \mid \forall x \lambda \\ \phi &:= \lambda \mid \neg \phi \mid \phi \vee \phi \mid \phi \wedge \phi \mid X\phi \mid \tilde{X}\phi \mid Y\phi \mid Z\phi \\ &\quad \mid \phi U \phi \mid \phi R \phi \mid \phi S \phi \mid \phi T \phi \end{aligned}$$

LTL_f^{MT} extends LTL_f by introducing **some** first-order elements.

- terms can refer to the value of constants at the **next** or **previous** state.

$$t := c \mid x \mid f(t_1, \dots, t_n) \mid \bigcirc c \mid \ominus c \mid \triangleleft c \mid \triangleleft c$$

- LTL_f propositions are replaced by first-order sentences

$$\begin{aligned} \lambda &:= p(t_1, \dots, t_n) \mid \neg \lambda \mid \lambda \vee \lambda \mid \lambda \wedge \lambda \mid \exists x \lambda \mid \forall x \lambda \\ \phi &:= \lambda \mid \neg \phi \mid \phi \vee \phi \mid \phi \wedge \phi \mid X\phi \mid \tilde{X}\phi \mid Y\phi \mid Z\phi \\ &\quad \mid \phi U \phi \mid \phi R \phi \mid \phi S \phi \mid \phi T \phi \end{aligned}$$

$$G(a = 2b)$$

$$(a < b) \cup (b = 0)$$

$$G(a > 5) \wedge F(a = 0)$$

$$G(\exists x(a = 2x))$$

$$G(a = 2b) \quad (a < b) \cup (b = 0) \quad G(a > 5) \wedge F(a = 0)$$
$$G(\exists x(a = 2x))$$

$$G(a = 2b)$$

$$(a < b) \cup (b = 0)$$

$$G(a > 5) \wedge F(a = 0)$$

$$G(\exists x(a = 2x))$$

$$G(a = 2b)$$

$$(a < b) \cup (b = 0)$$

$$G(a > 5) \wedge F(a = 0)$$

$$G(\exists x(a = 2x))$$

$$G(a = 2b)$$

$$(a < b) \cup (b = 0)$$

$$G(a > 5) \wedge F(a = 0)$$

$$G(\exists x(a = 2x))$$

$$a = 0 \wedge ((\bigcirc a = a + 1) \cup a = 42)$$

$$b = 1 \wedge G(\sim b = b + 1 \wedge a = 2b)$$

$$G(p(a) \rightarrow \tilde{X}p(\triangleleft a + 1))$$

$$a = 0 \wedge ((\bigcirc a = a + 1) \cup a = 42)$$

$$b = 1 \wedge G(\sim b = b + 1 \wedge a = 2b)$$

$$G(p(a) \rightarrow \tilde{X}p(\triangleleft a + 1))$$

$$a = 0 \wedge ((\bigcirc a = a + 1) \cup a = 42)$$

$$b = 1 \wedge G(\sim b = b + 1 \wedge a = 2b)$$

$$G(p(a) \rightarrow \tilde{X}p(\triangleleft a + 1))$$

$$a = 0 \wedge ((\bigcirc a = a + 1) \cup a = 42)$$

$$b = 1 \wedge G(\sim b = b + 1 \wedge a = 2b)$$

$$G(p(a) \rightarrow \tilde{X}p(\triangleleft a + 1))$$

LTL_f^{MT} is **undecidable**, but:

- it is **semi-decidable**
- so we can get something useful, sometimes:
 - models of satisfiable formulas
 - counterexamples of specifications over bugged systems
- moreover, we have some interesting **decidable** fragments [ECAI 23]
- why **finite traces**?
 - on **infinite traces**, LTL^{MT} is not even semi-decidable.

Decidable fragments

(MC) Formulas over LRA where all **iteration conditions** are **monotonicity constraints**, *i.e.*, variable-to-variable or variable-to-constant comparisons, *e.g.*:

$$a < 0 \wedge b = 1 \wedge ((\bigcirc b > b \wedge \bigcirc a \leq a) \cup a = b)$$

(IPC) Formulas over LIA where all **iteration conditions** are **integer periodicity constraints**, *e.g.*:

$$(b \equiv_3 a) \cup (a > 42) \wedge F(a + b = c)$$

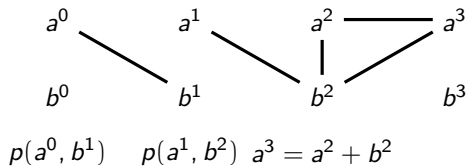
Iteration conditions:

- α in $\alpha \cup \beta$
- β in $\alpha \text{ R } \beta$

- (BL) **Bounded lookback** formulas, that generalize the above two by requiring that cross-state interaction is restricted to finitely many steps back. What does it mean?

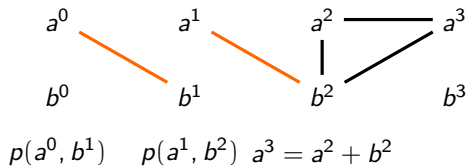
Bounded lookback formula

$$p(a, \bigcirc b) \cup (\bigcirc a = a + b)$$



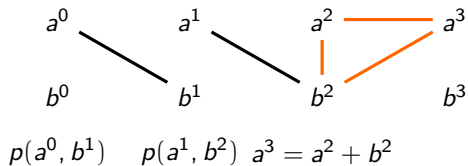
Bounded lookback formula

$$p(a, \bigcirc b) \cup (\bigcirc a = a + b)$$



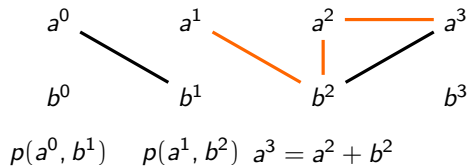
Bounded lookback formula

$$p(a, \bigcirc b) \cup (\bigcirc a = a + b)$$



Bounded lookback formula

$$p(a, \bigcirc b) \cup (\bigcirc a = a + b)$$



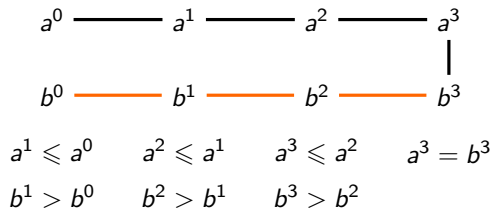
Non-BL formula

$$a < 0 \wedge b = 1 \wedge ((\bigcirc b > b \wedge \bigcirc a \leq a) \cup a = b)$$

$$\begin{array}{cccc} a^0 & \text{---} & a^1 & \text{---} & a^2 & \text{---} & a^3 \\ & & & & & & | \\ b^0 & \text{---} & b^1 & \text{---} & b^2 & \text{---} & b^3 \\ \\ a^1 \leq a^0 & & a^2 \leq a^1 & & a^3 \leq a^2 & & a^3 = b^3 \\ b^1 > b^0 & & b^2 > b^1 & & b^3 > b^2 & & \end{array}$$

Non-BL formula

$$a < 0 \wedge b = 1 \wedge ((\bigcirc b > b \wedge \bigcirc a \leq a) \cup a = b)$$



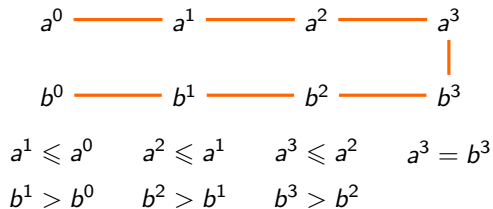
Non-BL formula

$$a < 0 \wedge b = 1 \wedge ((\bigcirc b > b \wedge \bigcirc a \leq a) \cup a = b)$$

$$\begin{array}{cccc} a^0 & \text{---} & a^1 & \text{---} & a^2 & \text{---} & a^3 \\ & & & & & & | \\ b^0 & \text{---} & b^1 & \text{---} & b^2 & \text{---} & b^3 \\ \\ a^1 \leq a^0 & & a^2 \leq a^1 & & a^3 \leq a^2 & & a^3 = b^3 \\ b^1 > b^0 & & b^2 > b^1 & & b^3 > b^2 & & \end{array}$$

Non-BL formula

$$a < 0 \wedge b = 1 \wedge ((\bigcirc b > b \wedge \bigcirc a \leq a) \cup a = b)$$



(NCS) Formulae without **cross-state comparisons**, *i.e.*, without any $\bigcirc c$, $\sim c$, *e.g.*:

$$(a > b \cup a + b = 2c) \wedge G(a + b > 0)$$

(FX) Formulas where the only temporal operators are F, X, and \tilde{X} , *e.g.*:

$$F(p(\bigcirc a) \wedge X(\neg p(a))) \wedge XF(r(a, b) \vee r(\bigcirc a, b))$$

The SAT encoding of Reynolds' tableau can be extended to an **SMT encoding** for LTL_f^{MT}.

- implemented in our BLACK solver¹ for arithmetic theories;
- given to **off-the-shelf** SMT solvers such as Z3 [MB08], cvc5 [Bar+22], *etc.*

¹<https://www.black-sat.org>

Which systems can we verify LTL_f^{MT} formulas on?

Knowledge-base driven Dynamic Systems (KDS):

- infinite-state transition systems

$$D = \langle K, I(X), T(X, X'), F(X) \rangle$$

- states are structures over the **first-order theory** K
- $I(X)$, $T(X, X')$, $F(X)$ are arbitrary **first-order** formulas over the theory K
 - **initial** states satisfying $I(X)$
 - **final** states satisfying $F(X)$
 - **transition relation** expressed by $T(X, X')$

Let D be a KDS and ϕ an LTL_f^{MT} formula:

- all the executions of a KDS D can be **represented** by an LTL_f^{MT} formula ψ_D
- **model-checking** of ϕ over D reduces to **satisfiability** of:

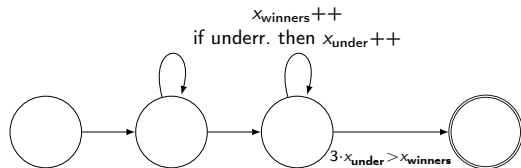
$$\gamma \equiv \psi_D \wedge \neg\phi$$

- if γ is **satisfiable**, the specification does not hold and the model is a counterexample
- if γ is **unsatisfiable**, the specification is valid over D
- this needs **non-rigid predicates**: only semi-decidable!

Some early experiments

Test setting:

- simulation of a company hiring process
- nondeterministic transitions:
 - dependent on arithmetic constraints
 - acting on unbounded relational data
- minimal length of the counterexamples dependent over scalable parameter N
- two modelings of the same system:
 - P_1 employs arithmetic constraints
 - P_2 avoids arithmetics, simulates constraints by other means
- two different properties for each variant



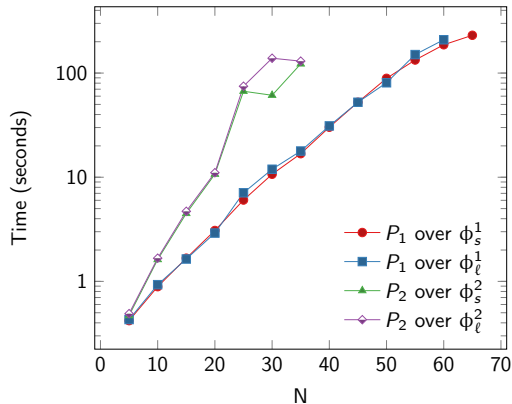
$$\phi_s^1 \equiv G(x_{state} = final \rightarrow 2x_{under} > x_{winners})$$

$$\phi_\ell^1 \equiv G \left(x_{state} = app \rightarrow F(x_{state} = final \wedge 2x_{under} > x_{winners}) \right)$$

Some early experiments

Results:

- 5 minutes timeout reached at $N = 70$
- **exponential** growth
 - but could be much worse, the problem is **undecidable!**
- liveness property not harder than the safety one
- system with **explicit arithmetics** faster to verify



THE PRESENT:
FIRST-ORDER AUTOMATA

LTL_f^{MT} is limited in many ways:

- many complex systems need **evolving predicates**;
- **nesting** of quantifiers and temporal operators is often essential
- we want to approach **full FOLTL**.

The need for automata

In the finite-state setting, automata are essential:

- operational counterpart of temporal logics
- basis of many algorithms and techniques

We want a class of automata for FOLTL.

What may an infinite-state first-order automaton look like?

- **states**: first-order structures over a state signature Γ ;
- **letters**: first-order structures over an alphabet signature Σ ;
- **initial states**: a class of Γ -structures;
- **final states**: a class of Γ -structures;
- **transition**: a relation between:
 - a Γ -structure (source state),
 - a Σ -structure (letter), and
 - another Γ -structure (dest. state).

But this is algorithmically untractable.

A **finite-state** automaton:

$$A = \langle \Sigma, Q, Q_0, \Delta, F \rangle$$

A **symbolic** finite-state automaton:

$$\mathcal{A} = \langle \Sigma, X, I(X), T(X, \Sigma, X'), F(X) \rangle$$

A **finite-state** automaton:

$$A = \langle \Sigma, Q, Q_0, \Delta, F \rangle$$

A **symbolic** finite-state automaton:

$$\mathcal{A} = \langle \Sigma, X, I(X), T(X, \Sigma, X'), F(X) \rangle$$

A **first-order** automaton:

$$A = \langle \Sigma, \Gamma, \phi_I, \phi_T, \phi_F \rangle$$

where:

- Σ is the **word** signature and Γ is the **state** signature;
- ϕ_I and ϕ_F are Γ -sentences;
- ϕ_T is a sentence over $\Gamma \cup \Sigma \cup \Gamma'$.

A **first-order** automaton:

$$A = \langle \Sigma, \Gamma, \phi_I, \phi_T, \phi_F \rangle$$

where:

- Σ is the **word** signature and Γ is the **state** signature;
- ϕ_I and ϕ_F are Γ -sentences;
- ϕ_T is a sentence over $\Gamma \cup \Sigma \cup \Gamma'$.

A word (sequence of Σ -structures) $\bar{\sigma} = \langle \sigma_0, \dots, \sigma_{n-1} \rangle$ is accepted iff there is a **trace** (sequence of Γ -structures) $\bar{\rho} = \langle \rho_0, \dots, \rho_n \rangle$ such that:

- $\rho_0 \models \phi_I$;
- $\rho_i \cup \sigma_i \cup \rho'_{i+1} \models \phi_T$;
- $\rho_n \models \phi_F$.

First-order T -regular languages

A set of first-order words whose structures come from a theory T is a **T -language**.

A T -language is **first-order T -regular** if there is a first-order automaton accepting it.

Theorem (Closure properties)

First-order T -regular languages are *closed* by:

- *union*;
- *intersection*;
- *concatenation*;
- *Kleene star*.

As an example we see how to prove concatenation.

We have two automata $A_1 = \langle \Sigma, \Gamma_1, \phi_0^1, \phi_T^1, \phi_F^1 \rangle$ and $A_2 = \langle \Sigma, \Gamma_2, \phi_0^2, \phi_T^2, \phi_F^2 \rangle$.

We want $A = \langle \Sigma, \Gamma, \phi_0, \phi_T, \phi_F \rangle$ such that $\mathcal{L}(A) = \mathcal{L}(A_1) \cdot \mathcal{L}(A_2)$.

- suppose for a moment that ϕ_T can be an **existential second-order** sentence;
- then, what about:

$$\begin{aligned} \phi_T \equiv & (p \wedge \phi_T^1 \wedge p') \vee (\neg p \wedge \phi_T^2 \wedge \neg p') \vee \\ & \left(p \wedge \neg p' \wedge \phi_F^1 \wedge \exists \Gamma'' (\phi_0^2[\Gamma/\Gamma''] \wedge \phi_T^2[\Gamma/\Gamma'']) \right) \end{aligned}$$

with $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \{p\}$.

What about these existential second-order quantifiers?

$$\begin{aligned} \phi_T \equiv & (p \wedge \phi_T^1 \wedge p') \vee (\neg p \wedge \phi_T^2 \wedge \neg p') \vee \\ & \left(p \wedge \neg p' \wedge \phi_F^1 \wedge \exists \Gamma'' (\phi_0^2[\Gamma/\Gamma''] \wedge \phi_T^2[\Gamma/\Gamma'']) \right) \end{aligned}$$

What about these existential second-order quantifiers?

$$\phi_T \equiv \exists \Gamma'' \left[\begin{array}{l} (p \wedge \phi_T^1 \wedge p') \vee (\neg p \wedge \phi_T^2 \wedge \neg p') \vee \\ (p \wedge \neg p' \wedge \phi_F^1 \wedge (\phi_0^2[\Gamma/\Gamma''] \wedge \phi_T^2[\Gamma/\Gamma''])) \end{array} \right]$$

What about these existential second-order quantifiers?

$$\phi_T \equiv \exists \Gamma'' \left[\begin{array}{l} (p \wedge \phi_T^1 \wedge p') \vee (\neg p \wedge \phi_T^2 \wedge \neg p') \vee \\ (p \wedge \neg p' \wedge \phi_F^1 \wedge (\phi_0^2[\Gamma/\Gamma''] \wedge \phi_T^2[\Gamma/\Gamma''])) \end{array} \right]$$

Now the predicates in Γ'' can be added to Γ itself:

- the semantics of the automata will require the existence of their interpretation;
- we get back a first-order automaton on an extended state signature.

What about these existential second-order quantifiers?

$$\phi_T \equiv \left[\begin{array}{l} (p \wedge \phi_T^1 \wedge p') \vee (\neg p \wedge \phi_T^2 \wedge \neg p') \vee \\ \left(p \wedge \neg p' \wedge \phi_F^1 \wedge (\phi_0^2[\Gamma/\Gamma''] \wedge \phi_T^2[\Gamma/\Gamma'']) \right) \end{array} \right]$$

Now the predicates in Γ'' can be added to Γ itself:

- the semantics of the automata will require the existence of their interpretation;
- **we get back a first-order automaton on an extended state signature.**

Complementation is missing from the picture:

- easy for **deterministic** automata;
- so **determinization** is the key;
- how to symbolically represent the subset construction of a first-order automaton?
 - we are working on it...

A major feature of automata is that they capture temporal logic.

Can we capture FOLTL with first-order automata?

We need some ingredients.

Definition (Stepped normal form)

Given an FOLTL formula ϕ , the *step normal form* of ϕ , denoted $\text{snf}(\phi)$, is the formula defined recursively as follows:

- $\text{snf}(p(t_1, \dots, t_n)) = p(t_1, \dots, t_n)$ and $\text{snf}(t_1 = t_2) = (t_1 = t_2)$;
- $\text{snf}(Qx\phi) = Qx\text{snf}(\phi)$, where $Q \in \{\forall, \exists\}$ and x is a first-order variable;
- $\text{snf}(\neg\phi) = \neg\text{snf}(\phi)$;
- $\text{snf}(\phi_1 \circ \phi_2) = \text{snf}(\phi_1) \circ \text{snf}(\phi_2)$, where $\circ \in \{\wedge, \vee\}$;
- $\text{snf}(\circ\phi) = \circ\phi$ where $\circ \in \{X, Y, \tilde{X}, Z\}$;
- $\text{snf}(\phi_1 U \phi_2) = \text{snf}(\phi_2) \vee (\text{snf}(\phi_1) \wedge X(\phi_1 U \phi_2))$;
- $\text{snf}(\phi_1 R \phi_2) = \text{snf}(\phi_2) \wedge (\text{snf}(\phi_1) \vee \tilde{X}(\phi_1 R \phi_2))$.

Definition (Closure)

The *closure* of a FOLTL sentence ϕ is the set $C(\phi)$ defined as follows:

- $X\phi \in C(\phi)$;
- $\psi \in C(\phi)$ for any subformula ψ of ϕ (including itself);
- for any $\phi_1 \cup \phi_2 \in C(\phi)$ we have $X(\phi_1 \cup \phi_2) \in C(\phi)$;
- for any $\phi_1 \text{ R } \phi_2 \in C(\phi)$ we have $\tilde{X}(\phi_1 \text{ R } \phi_2) \in C(\phi)$.

We also define the following subsets of the closure:

$$XS = \{xs_\psi \mid X\psi \in C(\phi)\}$$

$$\tilde{X}S = \{ws_\psi \mid \tilde{X}\psi \in C(\phi)\}$$

where xs_ψ and ws_ψ are predicates of the arity n corresponding to the number of free first-order variables in ψ .

Let ϕ be a FOLTL sentence. The automaton of ϕ is $\mathcal{A}(\phi) = \langle \Sigma, \Gamma, \phi_0, \phi_T, \phi_F \rangle$ where:

- the state signature is $\Gamma = XS \cup \tilde{X}S$;
- the initial and final conditions are:

$$\phi_0 \equiv xS_\phi$$

$$\phi_F \equiv \bigwedge_{ws_\psi \in \tilde{X}S} \forall \bar{x}. ws_\psi(\bar{x}) \wedge \bigwedge_{xs_\psi \in XS} \forall \bar{x}. \neg xs_\psi(\bar{x})$$

- the transition relation is:

$$\phi_T \equiv \bigwedge_{s_\psi \in XS \cup \tilde{X}S} \forall \bar{x}. [s_\psi(\bar{x}) \leftrightarrow \text{snf}'_S(\psi(\bar{x}))]$$

Checking (non-)emptiness

Of course checking emptiness of first-order automata is **undecidable**.

However, we can state a semi-algorithm for non-emptiness with some assumptions.

Checking (non-)emptiness

For $k > 0$, let:

$$\llbracket A \rrbracket_k^F \equiv \phi_0^0 \wedge \bigwedge_{i=0}^{k-1} \phi_T^k \wedge \phi_F^k$$

```
1: procedure NonEmpty(A)
2:    $k \leftarrow 0$ 
3:   while true do
4:     if  $\llbracket A \rrbracket_k^F$  is satisfiable then
5:       return not empty
6:     end if
7:      $k \leftarrow k + 1$ 
8:   end while
9: end procedure
```

Recall the result from the literature we cited before.

For a fragment F such that:

- the underlying FO fragment is decidable; and
- temporal formulas are **monodic**;
 - only one **free variable** inside temporal operators;
- constants in the signature are considered **rigid**;
- then FOLTL(F) is decidable [HWZ00]

Can we recover this result from first-order automata? **Yes.**

- if we follow the encoding we notice that **monodic** FOLTL sentences translate into automata with **monadic** state signature Γ ;
- monadic predicates are easier to deal with (e.g., monadic FO is decidable);
- we can translate a monadic Γ into a Γ containing only **propositions**;
- emptiness for this kind of automata is decidable (they are almost finite-state).

CONCLUSIONS

Our satisfiability checking tool BLACK is currently being refactored to support full FOLTL through first-order automata.

- experimental results soon, maybe another iFM² talk?

Can we go beyond the simple unraveling semi-algorithm shown before?

For better **practical** results we need more **theory**.

Interesting paths we are exploring:

- fixpoint computation based on **second-order quantifier elimination** [GSS08];
- encoding into **constrained Horn clauses** [GB19];

What we have not seen today

We can add **past operators** to FOLTL.

- if we start from **pure-past** FOLTL sentences, we directly obtain **deterministic** first-order automata;
- this happens with pure-past LTL as well and is crucial for **reactive synthesis**;
- 2EXPTIME for LTL, EXPTIME for pure-past LTL, because of this fact;
- can this help with **reactive synthesis** for FOLTL specifications?

We want first-order temporal logics to become practical tools for infinite-state specification and verification.

- long way ahead;
- little steps already done;
- first-order automata as a reasoning and algorithmic tool;
- many theoretical and algorithmic developments missing yet.

We want first-order temporal logics to become practical tools for infinite-state specification and verification.

- long way ahead;
- little steps already done;
- first-order automata as a reasoning and algorithmic tool;
- many theoretical and algorithmic developments missing yet.

Thank you!
Questions?

REFERENCES

- [AD94] Rajeev Alur and David L. Dill. 'A Theory of Timed Automata'. In: *Theor. Comput. Sci.* 126.2 (1994), pp. 183–235. doi: 10.1016/0304-3975(94)90010-8.
- [AH94] Rajeev Alur and Thomas A. Henzinger. 'A Really Temporal Logic'. In: *Journal of the ACM* 41.1 (1994), pp. 181–204. doi: 10.1145/174644.174651.
- [Alu+05] Rajeev Alur, Michael Benedikt, Kousha Etessami, Patrice Godefroid, Thomas W. Reps and Mihalis Yannakakis. 'Analysis of recursive state machines'. In: *ACM Trans. Program. Lang. Syst.* 27.4 (2005), pp. 786–818. doi: 10.1145/1075382.1075387.
- [Alu+92] Rajeev Alur, Costas Courcoubetis, Thomas A. Henzinger and Pei-Hsin Ho. 'Hybrid Automata: An Algorithmic Approach to the Specification and Verification of Hybrid Systems'. In: *Hybrid Systems*. Ed. by Robert L. Grossman, Anil Nerode, Anders P. Ravn and Hans Rischel. Vol. 736. Lecture Notes in Computer Science. Springer, 1992, pp. 209–229. doi: 10.1007/3-540-57318-6_30.

- [AM04] Rajeev Alur and P. Madhusudan. 'Visibly pushdown languages'. In: *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*. Ed. by László Babai. ACM, 2004, pp. 202–211. doi: 10.1145/1007352.1007390.
- [Bar+22] Haniel Barbosa, Clark W. Barrett, Martin Brain, Gereon Kremer, Hanna Lachnitt, Makai Mann, Abdalrhman Mohamed, Mudathir Mohamed, Aina Niemetz, Andres Nötzli, Alex Ozdemir, Mathias Preiner, Andrew Reynolds, Ying Sheng, Cesare Tinelli and Yoni Zohar. 'cvc5: A Versatile and Industrial-Strength SMT Solver'. In: *Proceedings of the 28th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Ed. by Dana Fisman and Grigore Rosu. Vol. 13243. Lecture Notes in Computer Science. Springer, 2022, pp. 415–442. doi: 10.1007/978-3-030-99524-9_24.

- [BK18] James Brotherston and Max I. Kanovich. ‘On the Complexity of Pointer Arithmetic in Separation Logic’. In: *Proceedings of the 16th Asian Symposium on Programming Languages and Systems*. Ed. by Sukyoung Ryu. Vol. 11275. Lecture Notes in Computer Science. Springer, 2018, pp. 329–349. doi: 10.1007/978-3-030-02768-1_18.
- [BMP10] Massimo Benerecetti, Stefano Minopoli and Adriano Peron. ‘Analysis of Timed Recursive State Machines’. In: *Proceedings of the 17th International Symposium on Temporal Representation and Reasoning*. Ed. by Nicolas Markey and Jef Wijsen. IEEE Computer Society, 2010, pp. 61–68. doi: 10.1109/TIME.2010.10.
- [Boj+06] Mikolaj Bojanczyk, Anca Muscholl, Thomas Schwentick, Luc Segoufin and Claire David. ‘Two-Variable Logic on Words with Data’. In: *21th IEEE Symposium on Logic in Computer Science (LICS 2006), 12-15 August 2006, Seattle, WA, USA, Proceedings*. IEEE Computer Society, 2006, pp. 7–16. doi: 10.1109/LICS.2006.51.
- [BZ83] Daniel Brand and Pitro Zafiropulo. ‘On Communicating Finite-State Machines’. In: *J. ACM* 30.2 (1983), pp. 323–342. doi: 10.1145/322374.322380.

- [CGM13] Diego Calvanese, Giuseppe De Giacomo and Marco Montali. 'Foundations of data-aware process analysis: a database theory perspective'. In: *Proceedings of the 32nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*. Ed. by Richard Hull and Wenfei Fan. ACM, 2013, pp. 1–12. doi: 10.1145/2463664.2467796.
- [Cim+20] Alessandro Cimatti, Alberto Griggio, Enrico Magnago, Marco Roveri and Stefano Tonetta. 'SMT-based satisfiability of first-order LTL with event freezing functions and metric operators'. In: *Inf. Comput.* 272 (2020), p. 104502. doi: 10.1016/j.ic.2019.104502. url: <https://doi.org/10.1016/j.ic.2019.104502>.
- [CMP22] Michele Chiari, Dino Mandrioli and Matteo Pradella. 'A First-Order Complete Temporal Logic for Structured Context-Free Languages'. In: *Log. Methods Comput. Sci.* 18.3 (2022). doi: 10.46298/lmcs-18(3:11)2022.

- [CYO01] Cristiano Calcagno, Hongseok Yang and Peter W. O’Hearn. ‘Computability and Complexity Results for a Spatial Assertion Language for Data Structures’. In: *Proceedings of the 21st Conference on Foundations of Software Technology and Theoretical Computer Science*. Ed. by Ramesh Hariharan, Madhavan Mukund and V. Vinay. Vol. 2245. Lecture Notes in Computer Science. Springer, 2001, pp. 108–119. doi: 10.1007/3-540-45294-X_10.
- [DLN07] Stéphane Demri, Ranko Lazic and David Nowak. ‘On the freeze quantifier in Constraint LTL: Decidability and complexity’. In: *Inf. Comput.* 205.1 (2007), pp. 2–24. doi: 10.1016/j.ic.2006.08.003.
- [DLV19] Alin Deutsch, Yuliang Li and Victor Vianu. ‘Verification of Hierarchical Artifact Systems’. In: *ACM Trans. Database Syst.* 44.3 (2019), 12:1–12:68. doi: 10.1145/3321487.
- [DQ21] Stéphane Demri and Karin Quaas. ‘Concrete domains in logics: a survey’. In: *ACM SIGLOG News* 8.3 (2021), pp. 6–29. doi: 10.1145/3477986.3477988. url: <https://doi.org/10.1145/3477986.3477988>.

- [Dro+17] Manfred Droste, Stefan Dück, Dino Mandrioli and Matteo Pradella. ‘Weighted Operator Precedence Languages’. In: *Proceedings of the 42nd International Symposium on Mathematical Foundations of Computer Science*. Ed. by Kim G. Larsen, Hans L. Bodlaender and Jean-François Raskin. Vol. 83. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017, 31:1–31:15. doi: 10.4230/LIPIcs.MFCS.2017.31.
- [DV13] Giuseppe De Giacomo and Moshe Y. Vardi. ‘Linear Temporal Logic and Linear Dynamic Logic on Finite Traces’. In: *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*. Ed. by Francesca Rossi. IJCAI/AAAI, 2013, pp. 854–860.
- [ECAI 23] Luca Geatti, Alessandro Gianola, Nicola Gigante and Sarah Winkler. ‘Decidable Fragments of LTLf Modulo Theories’. In: *Proceedings of the 26th European Conference on Artificial Intelligence*. 2023.

References

- [Fel+23] Paolo Felli, Marco Montali, Fabio Patrizi and Sarah Winkler. ‘Monitoring Arithmetic Temporal Properties on Finite Traces’. In: *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023*. Ed. by Brian Williams, Yiling Chen and Jennifer Neville. AAAI Press, 2023, pp. 6346–6354. doi: 10.1609/aaai.v37i5.25781.
- [GB19] Arie Gurfinkel and Nikolaj S. Bjørner. ‘The Science, Art, and Magic of Constrained Horn Clauses’. In: *Proceedings of the 21st International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*. IEEE, 2019, pp. 6–10. doi: 10.1109/SYNASC49474.2019.00010.
- [Ghi+23] Silvio Ghilardi, Alessandro Gianola, Marco Montali and Andrey Rivkin. ‘Safety Verification and Universal Invariants for Relational Action Bases’. In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*. ijcai.org, 2023, pp. 3248–3257. doi: 10.24963/ijcai.2023/362.
- [GSS08] Dov M. Gabbay, Renate A. Schmidt and Andrzej Szalas. *Second-Order Quantifier Elimination - Foundations, Computational Aspects and Applications*. Vol. 12. Studies in logic : Mathematical logic and foundations. College Publications, 2008. isbn: 978-1-904987-56-7.

- [HWZ00] Ian M. Hodkinson, Frank Wolter and Michael Zakharyashev. ‘Decidable fragment of first-order temporal logics’. In: *Ann. Pure Appl. Log.* 106.1-3 (2000), pp. 85–134. doi: 10.1016/S0168-0072(00)00018-X.
- [IJCAI 22] Luca Geatti, Alessandro Gianola and Nicola Gigante. ‘Linear Temporal Logic Modulo Theories over Finite Traces’. In: *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*. 2022, pp. 2641–2647. doi: 10.24963/ijcai.2022/366.
- [IX19] Radu Iosif and Xiao Xu. ‘Alternating Automata Modulo First Order Theories’. In: *Computer Aided Verification - 31st International Conference, CAV 2019, New York City, NY, USA, July 15-18, 2019, Proceedings, Part II*. Ed. by Isil Dillig and Serdar Tasiran. Vol. 11562. Lecture Notes in Computer Science. Springer, 2019, pp. 43–63. doi: 10.1007/978-3-030-25543-5_3.
- [JK09] Kurt Jensen and Lars Michael Kristensen. *Coloured Petri Nets - Modelling and Validation of Concurrent Systems*. Springer, 2009. doi: 10.1007/b95112.

- [KF94] Michael Kaminski and Nissim Francez. 'Finite-Memory Automata'. In: *Theor. Comput. Sci.* 134.2 (1994), pp. 329–363. doi: 10.1016/0304-3975(94)90242-9.
- [Kon+04] Roman Kontchakov, Carsten Lutz, Frank Wolter and Michael Zakharyashev. 'Temporalising Tableaux'. In: *Stud Logica* 76.1 (2004), pp. 91–134. doi: 10.1023/B:STUD.0000027468.28935.6d.
- [Koy90] Ron Koymans. 'Specifying Real-Time Properties with Metric Temporal Logic'. In: *Real-Time Systems* 2.4 (1990), pp. 255–299. doi: 10.1007/BF01995674.
- [LVR22] Jianwen Li, Moshe Y. Vardi and Kristin Y. Rozier. 'Satisfiability checking for Mission-time LTL (MLTL)'. In: *Inf. Comput.* 289.Part (2022), p. 104923. doi: 10.1016/j.ic.2022.104923.
- [Man20] Alessio Mansutti. 'Reasoning with separation logics: complexity, expressive power, proof systems'. PhD thesis. University of Paris-Saclay, France, 2020. url: <https://tel.archives-ouvertes.fr/tel-03094373>.

- [MB08] Leonardo Mendonça de Moura and Nikolaj S. Bjørner. 'Z3: An Efficient SMT Solver'. In: *Proceedings of the 14th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Ed. by C. R. Ramakrishnan and Jakob Rehof. Vol. 4963. Lecture Notes in Computer Science. Springer, 2008, pp. 337–340. doi: 10.1007/978-3-540-78800-3_24. url: https://doi.org/10.1007/978-3-540-78800-3%5C_24.
- [MN04] Oded Maler and Dejan Nickovic. 'Monitoring Temporal Properties of Continuous Signals'. In: *Proceedings of Formal FORMATS 2004 and FTRTFT 2004*. Ed. by Yassine Lakhnech and Sergio Yovine. Vol. 3253. Lecture Notes in Computer Science. Springer, 2004, pp. 152–166. doi: 10.1007/978-3-540-30206-3_12.
- [Mur89] Tadao Murata. 'Petri nets: Properties, analysis and applications'. In: *Proc. IEEE* 77.4 (1989), pp. 541–580. doi: 10.1109/5.24143.
- [Seg06] Luc Segoufin. 'Automata and Logics for Words and Trees over an Infinite Alphabet'. In: *Proceedings of the 20th International Workshop on Computer Science Logic*. Ed. by Zoltán Ésik. Vol. 4207. Lecture Notes in Computer Science. Springer, 2006, pp. 41–57. doi: 10.1007/11874683_3.

- [Voi19] Marco Voigt. 'Decidable fragments of first-order logic and of first-order linear arithmetic with uninterpreted predicates'. PhD thesis. Saarland University, Saarbrücken, Germany, 2019. url: <https://publikationen.sulb.uni-saarland.de/handle/20.500.11880/27767>.
- [Woj99] Arkadiusz Wojna. 'Counter Machines'. In: *Inf. Process. Lett.* 71.5-6 (1999), pp. 193–197. doi: 10.1016/S0020-0190(99)00116-7.