Learning Linear Temporal Properties

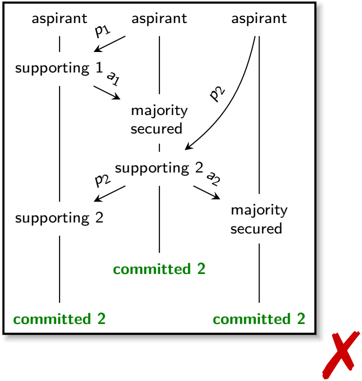
Daniel Neider

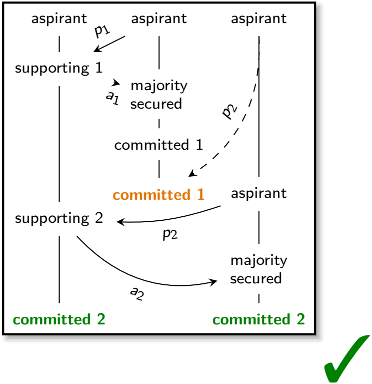
joint work with Ivan Gavran

MAX PLANCK INSTITUTE
FOR SOFTWARE SYSTEMS

Complexity, Algorithms, Automata and Logic Meet
Chennai Mathematical Institute, India
January 24, 2019

Apache Zookeeper Leader Election





Goal

Development of learning algorithms for formulas in Linear Temporal Logic (LTL) from data

Outline

1. Learning setup
2. A base algorithm for learning LTL formulas based on SAT solving
3. A performance improvement for the base algorithm based on learning decision trees
4. Future work

1. Learning Setup

Let $\mathcal{P}$ be a set of atomic propositions. Then,

- each $p \in \mathcal{P}$ is an LTL formula, and
- if φ and ψ are LTL formulas, so are $\neg\varphi$, $\varphi \lor \psi$, $\mathsf{X}\,\varphi$, $\varphi\,\mathsf{U}\,\psi$ (as well as $\varphi \land \psi$, $\varphi \rightarrow \psi$, $\mathsf{F}\,\varphi$, $\mathsf{G}\,\varphi$, and so on)

Let $\mathcal{P}$ be a set of atomic propositions. Then,

- each $p \in \mathcal{P}$ is an LTL formula, and
- if φ and ψ are LTL formulas, so are $\neg\varphi$, $\varphi \vee \psi$, $\mathsf{X}\,\varphi$, $\varphi \,\mathsf{U}\, \psi$ (as well as $\varphi \wedge \psi$, $\varphi \rightarrow \psi$, $\mathsf{F}\,\varphi$, $\mathsf{G}\,\varphi$, and so on)

We interpret LTL formulas over finite words $u \in (2^{\mathcal{P}})^*$

- For instance, $u = \{p\}\{q\}\emptyset\{p,q\}\emptyset\{p,q\}$

Let $\mathcal{P}$ be a set of atomic propositions. Then,

- each $p \in \mathcal{P}$ is an LTL formula, and
- if φ and ψ are LTL formulas, so are $\neg\varphi$, $\varphi \lor \psi$, $\mathsf{X}\,\varphi$, $\varphi\,\mathsf{U}\,\psi$ (as well as $\varphi \land \psi$, $\varphi \to \psi$, $\mathsf{F}\,\varphi$, $\mathsf{G}\,\varphi$, and so on)

We interpret LTL formulas over finite words $u \in (2^{\mathcal{P}})^*$

- For instance, $u = \{p\}\{q\}\emptyset\{p,q\}\emptyset\{p,q\}$

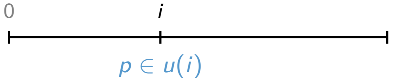
Example

Let $\mathcal{P} = \{p, q, r\}$

- $\neg p \lor q$
- $\mathsf{X}(p \land q)$
- $\neg p\,\mathsf{U}(q \land \mathsf{X}\,r)$

The semantics is defined by a satisfaction relation $\models$ as follows:

$u, i \models p \quad \Leftrightarrow \quad u$:



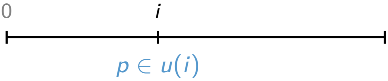
The semantics is defined by a satisfaction relation $\models$ as follows:

$$u, i \models p \quad \Leftrightarrow \quad u\colon \underset{0}{\vdash}\!\!-\!\!-\!\!-\!\!\underset{i}{+}\!\!-\!\!-\!\!-\!\!\dashv \qquad p \in u(i)$$

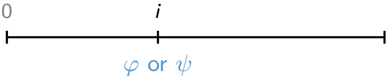
$$u, i \models \varphi \vee \psi \quad \Leftrightarrow \quad u\colon \underset{0}{\vdash}\!\!-\!\!-\!\!-\!\!\underset{i}{+}\!\!-\!\!-\!\!-\!\!\dashv \qquad \varphi \text{ or } \psi$$

The semantics is defined by a satisfaction relation $\models$ as follows:

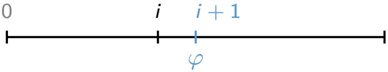
$u, i \models p \quad \Leftrightarrow \quad u$:



$u, i \models \varphi \lor \psi \quad \Leftrightarrow \quad u$:

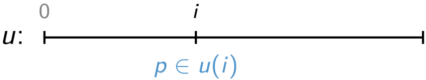


$u, i \models \mathrm{X}\, \varphi \quad \Leftrightarrow \quad u$:

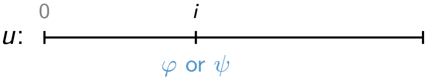


The semantics is defined by a satisfaction relation $\models$ as follows:

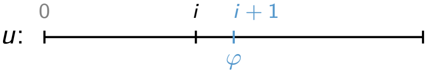
$u, i \models p \quad \Leftrightarrow$



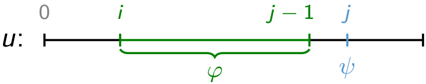
$u, i \models \varphi \lor \psi \quad \Leftrightarrow$



$u, i \models \mathrm{X}\, \varphi \quad \Leftrightarrow$

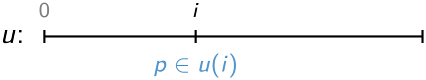


$u, i \models \varphi \,\mathrm{U}\, \psi \quad \Leftrightarrow$

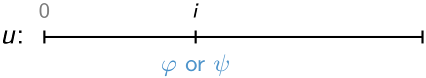


The semantics is defined by a satisfaction relation $\models$ as follows:

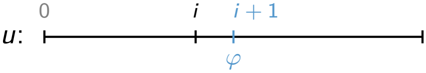
$u, i \models p \quad \Leftrightarrow$



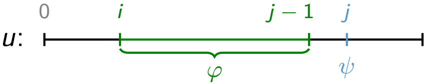
$u, i \models \varphi \vee \psi \quad \Leftrightarrow$



$u, i \models \mathsf{X}\, \varphi \quad \Leftrightarrow$



$u, i \models \varphi\, \mathsf{U}\, \psi \quad \Leftrightarrow$



Abbreviation: we write $u \models \psi$ for $u, 0 \models \psi$

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \land q)$

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \wedge q)$

$$u \quad = \quad \{p\} \quad \{q\} \quad \emptyset \quad \{p, q\} \quad \emptyset \quad \{p, q\}$$

Consider the LTL formula $\psi := \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	p						
2	q						
3	$p \wedge q$						
4	$\mathsf{X}(p \wedge q)$						

Consider the LTL formula $\psi := \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	p						
2	q						
3	$p \wedge q$						
4	$\mathsf{X}(p \wedge q)$						

Consider the LTL formula $\psi := \mathsf{X}(p \land q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	p						
2	q						
3	$p \land q$						
4	$\mathsf{X}(p \land q)$						

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1					
2	q						
3	$p \wedge q$						
4	$\mathsf{X}(p \wedge q)$						

Consider the LTL formula $\psi := \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	p	1					
2	q						
3	$p \wedge q$						
4	$\mathsf{X}(p \wedge q)$						

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \land q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1					
2	q						
3	$p \land q$						
4	$\mathsf{X}(p \land q)$						

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \land q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	p	1	0				
2	q						
3	$p \land q$						
4	$\mathsf{X}(p \land q)$						

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \land q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1	0	0	1	0	1
2	q						
3	$p \land q$						
4	$\mathsf{X}(p \land q)$						

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \land q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \land q$						
4	$\mathsf{X}(p \land q)$						

Consider the LTL formula $\psi := \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \wedge q$						
4	$\mathsf{X}(p \wedge q)$						

Consider the LTL formula $\psi := \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \wedge q$						
4	$\mathsf{X}(p \wedge q)$						

Consider the LTL formula $\psi := \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \wedge q$	0					
4	$\mathsf{X}(p \wedge q)$						

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \wedge q$	0	0	0	1	0	1
4	$\mathsf{X}(p \wedge q)$						

Consider the LTL formula $\psi := \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \wedge q$	0	0	0	1	0	1
4	$\mathsf{X}(p \wedge q)$						

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \land q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \land q$	0	0	0	1	0	1
4	$\mathsf{X}(p \land q)$						

Consider the LTL formula $\psi := \mathsf{X}(p \land q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \land q$	0	0	0	1	0	1
4	$\mathsf{X}(p \land q)$	0					

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \wedge q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \wedge q$	0	0	0	1	0	1
4	$\mathsf{X}(p \wedge q)$	0	0	1	0	1	0

Consider the LTL formula $\psi \coloneqq \mathsf{X}(p \land q)$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \land q$	0	0	0	1	0	1
4	$\mathsf{X}(p \land q)$	0	0	1	0	1	0

Sample

A *sample* is a pair $\mathcal{S} = (P, N)$ consisting of

- a finite set $P \subset (2^{\mathcal{P}})^*$ of *positive examples* (✓) and
- a finite set $N \subset (2^{\mathcal{P}})^*$ of *negative examples* (✗)

such that $P \cap N = \emptyset$

Sample

A *sample* is a pair $\mathcal{S} = (P, N)$ consisting of

- a finite set $P \subset (2^{\mathcal{P}})^*$ of *positive examples* (✓) and
- a finite set $N \subset (2^{\mathcal{P}})^*$ of *negative examples* (✗)

such that $P \cap N = \emptyset$

Size of an LTL Formula

The *size* of formula ψ is the number of its unique subformulas

Sample

A *sample* is a pair $\mathcal{S} = (P, N)$ consisting of

- a finite set $P \subset (2^{\mathcal{P}})^*$ of *positive examples* (✓) and
- a finite set $N \subset (2^{\mathcal{P}})^*$ of *negative examples* (✗)

such that $P \cap N = \emptyset$

Size of an LTL Formula

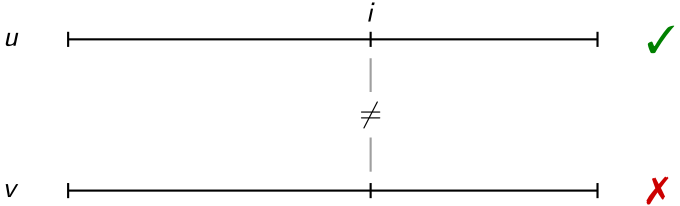
The *size* of formula ψ is the number of its unique subformulas

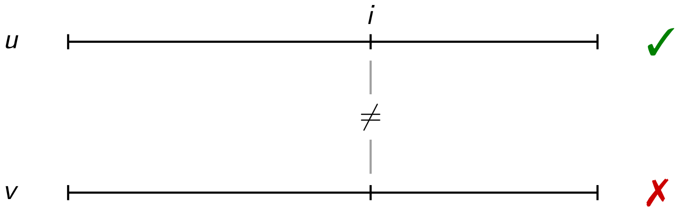
Learning of LTL Formulas

Given a sample $\mathcal{S}$, construct a minimal LTL formula ψ that is consistent with $\mathcal{S}$:

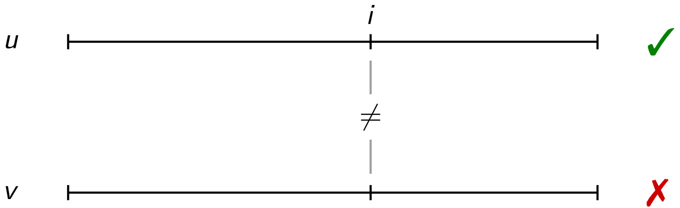
- $u \models \psi$ for each $u \in P$
- $u \not\models \psi$ for each $u \in N$

A Consistent Formula Always Exists

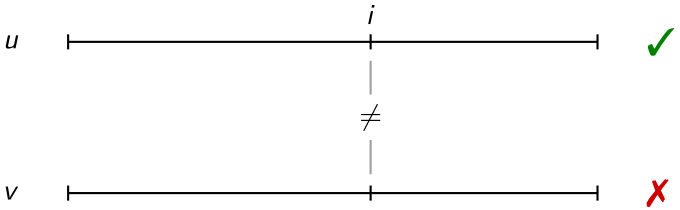




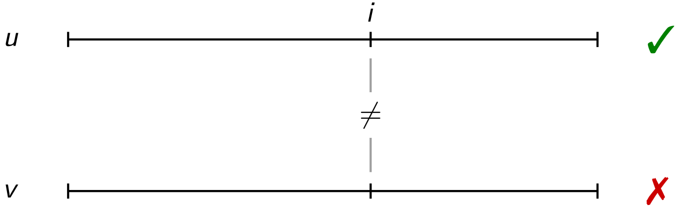
$$\left[\Big(\bigwedge_{p \in u(i)} p\Big) \wedge \Big(\bigwedge_{q \notin u(i)} \neg q\Big) \right]$$



$$\overbrace{\mathsf{X}\cdots\mathsf{X}}^{i\ \text{times}} \left[\Big(\bigwedge_{p\in u(i)} p \Big) \wedge \Big(\bigwedge_{q\notin u(i)} \neg q \Big) \right]$$



$$\bigvee_{u \in P} \bigwedge_{v \in N} \overbrace{X \cdots X}^{i \text{ times}} \left[\left(\bigwedge_{p \in u(i)} p \right) \wedge \left(\bigwedge_{q \notin u(i)} \neg q \right) \right]$$



$$\bigvee_{u\in P}\ \bigwedge_{v\in N}\ \overbrace{\mathsf{X}\cdots\mathsf{X}}^{i\ \text{times}} \left[\Big(\bigwedge_{p\in u(i)} p \Big) \wedge \Big(\bigwedge_{q\notin u(i)} \neg q \Big) \right]$$

But, this formula overfits the sample!

Given: a sample $\mathcal{S}$ and $n \in \mathbb{N}$

Question: does a consistent LTL formula of size n exist?

Given: a sample $\mathcal{S}$ and $n \in \mathbb{N}$

Question: does a consistent LTL formula of size n exist?

Membership in NP

- Nondeterministically guess an LTL formula of size n
- Verify that it is consistent with $\mathcal{S}$

Given: a sample $\mathcal{S}$ and $n \in \mathbb{N}$

Question: does a consistent LTL formula of size n exist?

Membership in NP

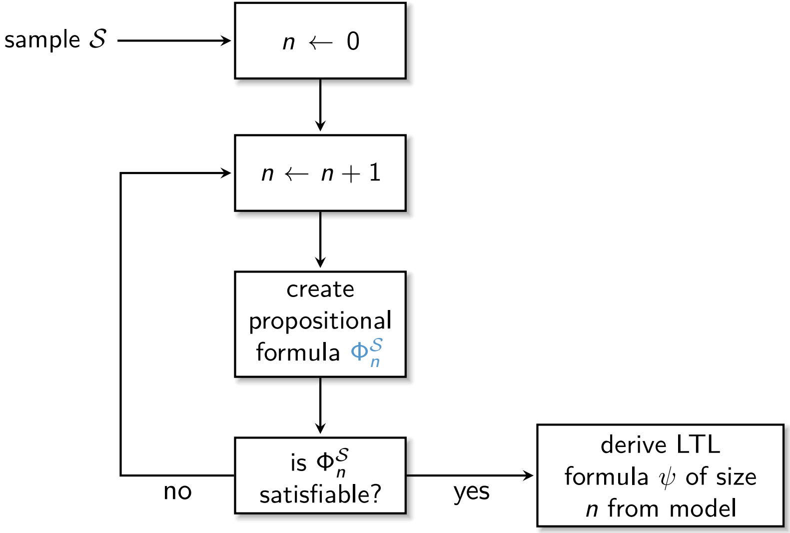
- Nondeterministically guess an LTL formula of size n
- Verify that it is consistent with $\mathcal{S}$

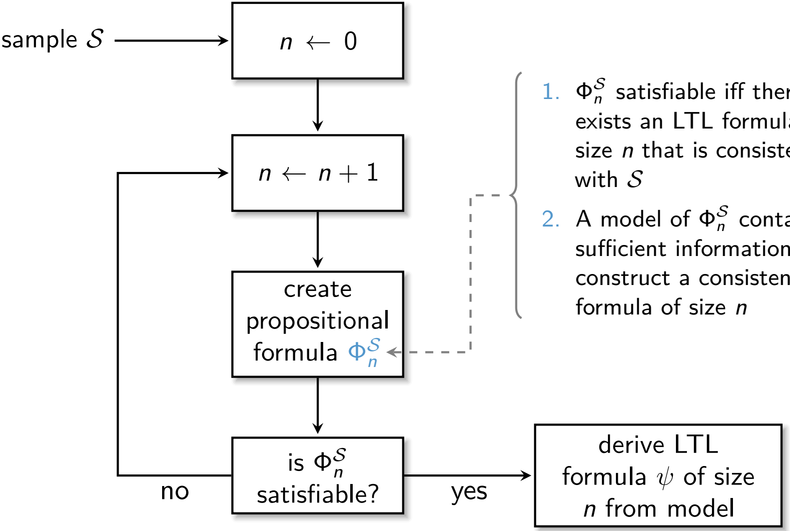
NP-Hardness?

We do not know whether the problem is NP-hard

- But we strongly suspect that it is
- What are samples requiring "large" consistent formulas?
  (proposal: $P = \{\{p\}^n\}$ and $N = \{\{p\}^{n+1}\}$)

2. A SAT-based Learning Algorithm


sample $\mathcal{S}$
$n \leftarrow 0$
$n \leftarrow n+1$
create propositional formula $\Phi_n^{\mathcal{S}}$
is $\Phi_n^{\mathcal{S}}$ satisfiable?
no
yes
derive LTL formula ψ of size n from model



1. $\Phi_n^{\mathcal{S}}$ satisfiable iff there exists an LTL formula of size n that is consistent with $\mathcal{S}$
2. A model of $\Phi_n^{\mathcal{S}}$ contains sufficient information to construct a consistent LTL formula of size n

$$\Phi_n^{\mathcal{S}} \coloneqq \boxed{\Phi_n^{structure}} \land \boxed{\Phi_n^{consistency}}$$

Encodes the structure of the LTL formula

$$\Phi_n^{\mathcal{S}} \coloneqq \boxed{\Phi_n^{structure}} \land \boxed{\Phi_n^{consistency}}$$

Ensures that the LTL formula is consistent with the sample

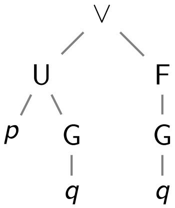
Encodes the structure of the LTL formula

$$\Phi_n^{\mathcal{S}} \coloneqq \boxed{\Phi_n^{structure}} \land \boxed{\Phi_n^{consistency}}$$

Ensures that the LTL formula is consistent with the sample

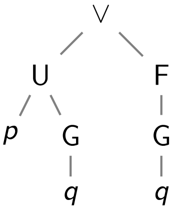
Consider the LTL formula $\varphi := (p \, \mathsf{U} (\mathsf{G}\, q)) \vee (\mathsf{F} (\mathsf{G}\, q))$

Consider the LTL formula $\varphi \coloneqq (p \, \mathsf{U} (\mathsf{G}\, q)) \vee (\mathsf{F} (\mathsf{G}\, q))$

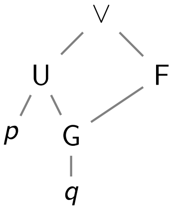


Syntax Tree

Consider the LTL formula $\varphi := (p \,\mathsf{U}(\mathsf{G}\, q)) \vee (\mathsf{F}(\mathsf{G}\, q))$

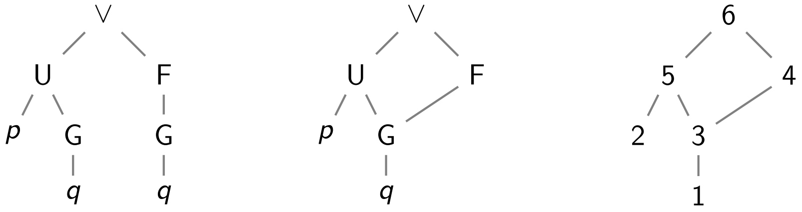


Syntax Tree



Syntax DAG

Consider the LTL formula $\varphi := (p\,\mathsf{U}(\mathsf{G}\,q)) \vee (\mathsf{F}(\mathsf{G}\,q))$

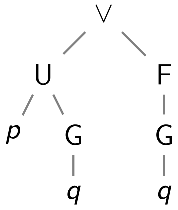


Syntax Tree

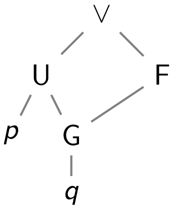
Syntax DAG

Identifiers

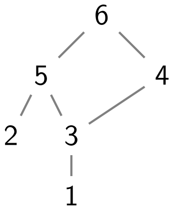
Consider the LTL formula $\varphi := (p\,\mathrm{U}(\mathrm{G}\,q)) \lor (\mathrm{F}(\mathrm{G}\,q))$



Syntax Tree



Syntax DAG



Identifiers

Variables

- $x_{i,\lambda}$ where $i \in \{1, \dots, n\}$ and $\lambda \in \Lambda$
- $l_{i,j}$ where $i \in \{2, \dots, n\}$ and $j \in \{1, \dots, i-1\}$
- $r_{i,j}$ where $i \in \{2, \dots, n\}$ and $j \in \{1, \dots, i-1\}$

Encoding the Structure of LTL Formulas

Variables

- $x_{i,\lambda}$ where $i \in \{1, \dots, n\}$ and $\lambda \in \Lambda$
- $l_{i,j}$ where $i \in \{2, \dots, n\}$ and $j \in \{1, \dots, i-1\}$
- $r_{i,j}$ where $i \in \{2, \dots, n\}$ and $j \in \{1, \dots, i-1\}$

Variables

- $x_{i,\lambda}$ where $i \in \{1, \ldots, n\}$ and $\lambda \in \Lambda$
- $l_{i,j}$ where $i \in \{2, \ldots, n\}$ and $j \in \{1, \ldots, i-1\}$
- $r_{i,j}$ where $i \in \{2, \ldots, n\}$ and $j \in \{1, \ldots, i-1\}$

Formula $\Phi_n^{structure}$

- For $i \in \{1, \ldots, n\}$, exactly one $x_{i,\lambda}$ can be set to `true`:

$$\left[\bigwedge_{1 \leq i \leq n} \bigvee_{\lambda \in \Lambda} x_{i,\lambda}\right] \wedge \left[\bigwedge_{1 \leq i \leq n} \bigwedge_{\lambda \neq \lambda' \in \Lambda} \neg x_{i,\lambda} \vee \neg x_{i,\lambda'}\right]$$

- For $i \in \{2, \ldots, n\}$, exactly one $l_{i,j}$ and exactly one $r_{i,j}$ can be set to `true`
- One of the variables $x_{1,p}$ for $p \in \mathcal{P}$ is set to `true`

Encodes the structure of the LTL formula

$$\Phi_n^{\mathcal{S}} := \boxed{\Phi_n^{structure}} \wedge \boxed{\Phi_n^{consistency}}$$

Ensures that the LTL formula is consistent with the sample

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	p	1	0	0	1	0	1
2	q	0	1	0	1	0	1
3	$p \wedge q$	0	0	0	1	0	1
4	$\mathsf{X}(p \wedge q)$	0	0	1	0	1	0

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	$x_{1,\lambda}$ / $\ell_{1,j}$ / $r_{1,j}$						
2	$x_{2,\lambda}$ / $\ell_{2,j}$ / $r_{2,j}$						
3	$x_{3,\lambda}$ / $\ell_{3,j}$ / $r_{3,j}$						
4	$x_{4,\lambda}$ / $\ell_{4,j}$ / $r_{4,j}$						

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p, q\}$	$\emptyset$	$\{p, q\}$
1	$x_{1,\lambda}$ / $\ell_{1,j}$ / $r_{1,j}$	$y^u_{1,0}$	$y^u_{1,1}$	$y^u_{1,2}$	$y^u_{1,3}$	$y^u_{1,4}$	$y^u_{1,5}$
2	$x_{2,\lambda}$ / $\ell_{2,j}$ / $r_{2,j}$	$y^u_{2,0}$	$y^u_{2,1}$	$y^u_{2,2}$	$y^u_{2,3}$	$y^u_{2,4}$	$y^u_{2,5}$
3	$x_{3,\lambda}$ / $\ell_{3,j}$ / $r_{3,j}$	$y^u_{3,0}$	$y^u_{3,1}$	$y^u_{3,2}$	$y^u_{3,3}$	$y^u_{3,4}$	$y^u_{3,5}$
4	$x_{4,\lambda}$ / $\ell_{4,j}$ / $r_{4,j}$	$y^u_{4,0}$	$y^u_{4,1}$	$y^u_{4,2}$	$y^u_{4,3}$	$y^u_{4,4}$	$y^u_{4,5}$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	$x_{1,\lambda}$ / $\ell_{1,j}$ / $r_{1,j}$	$y^u_{1,0}$	$y^u_{1,1}$	$y^u_{1,2}$	$y^u_{1,3}$	$y^u_{1,4}$	$y^u_{1,5}$
2	$x_{2,\lambda}$ / $\ell_{2,j}$ / $r_{2,j}$	$y^u_{2,0}$	$y^u_{2,1}$	$y^u_{2,2}$	$y^u_{2,3}$	$y^u_{2,4}$	$y^u_{2,5}$
3	$x_{3,\lambda}$ / $\ell_{3,j}$ / $r_{3,j}$	$y^u_{3,0}$	$y^u_{3,1}$	$y^u_{3,2}$	$y^u_{3,3}$	$y^u_{3,4}$	$y^u_{3,5}$
4	$x_{4,\lambda}$ / $\ell_{4,j}$ / $r_{4,j}$	$y^u_{4,0}$	$y^u_{4,1}$	$y^u_{4,2}$	$y^u_{4,3}$	$y^u_{4,4}$	$y^u_{4,5}$

$\wedge$-operator: $$\left[x_{3,\wedge} \wedge \ell_{3,1} \wedge r_{3,2}\right] \rightarrow \left[y^u_{3,1} \leftrightarrow (y^u_{1,1} \wedge y^u_{2,1})\right]$$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	$x_{1,\lambda}$ / $\ell_{1,j}$ / $r_{1,j}$	$y^u_{1,0}$	$y^u_{1,1}$	$y^u_{1,2}$	$y^u_{1,3}$	$y^u_{1,4}$	$y^u_{1,5}$
2	$x_{2,\lambda}$ / $\ell_{2,j}$ / $r_{2,j}$	$y^u_{2,0}$	$y^u_{2,1}$	$y^u_{2,2}$	$y^u_{2,3}$	$y^u_{2,4}$	$y^u_{2,5}$
3	$x_{3,\lambda}$ / $\ell_{3,j}$ / $r_{3,j}$	$y^u_{3,0}$	$y^u_{3,1}$	$y^u_{3,2}$	$y^u_{3,3}$	$y^u_{3,4}$	$y^u_{3,5}$
4	$x_{4,\lambda}$ / $\ell_{4,j}$ / $r_{4,j}$	$y^u_{4,0}$	$y^u_{4,1}$	$y^u_{4,2}$	$y^u_{4,3}$	$y^u_{4,4}$	$y^u_{4,5}$

X-operator: $\left[x_{3,\mathrm{X}} \land \ell_{3,2} \right] \rightarrow \left[y^u_{3,1} \leftrightarrow y^u_{2,2} \right]$

	u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	$x_{1,\lambda}$ / $\ell_{1,j}$ / $r_{1,j}$	$y^u_{1,0}$	$y^u_{1,1}$	$y^u_{1,2}$	$y^u_{1,3}$	$y^u_{1,4}$	$y^u_{1,5}$
2	$x_{2,\lambda}$ / $\ell_{2,j}$ / $r_{2,j}$	$y^u_{2,0}$	$y^u_{2,1}$	$y^u_{2,2}$	$y^u_{2,3}$	$y^u_{2,4}$	$y^u_{2,5}$
3	$x_{3,\lambda}$ / $\ell_{3,j}$ / $r_{3,j}$	$y^u_{3,0}$	$y^u_{3,1}$	$y^u_{3,2}$	$y^u_{3,3}$	$y^u_{3,4}$	$y^u_{3,5}$
4	$x_{4,\lambda}$ / $\ell_{4,j}$ / $r_{4,j}$	$y^u_{4,0}$	$y^u_{4,1}$	$y^u_{4,2}$	$y^u_{4,3}$	$y^u_{4,4}$	$y^u_{4,5}$

		u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	$x_{1,\lambda}$ / $\ell_{1,j}$ / $r_{1,j}$		$y^u_{1,0}$	$y^u_{1,1}$	$y^u_{1,2}$	$y^u_{1,3}$	$y^u_{1,4}$	$y^u_{1,5}$
2	$x_{2,\lambda}$ / $\ell_{2,j}$ / $r_{2,j}$		$y^u_{2,0}$	$y^u_{2,1}$	$y^u_{2,2}$	$y^u_{2,3}$	$y^u_{2,4}$	$y^u_{2,5}$
3	$x_{3,\lambda}$ / $\ell_{3,j}$ / $r_{3,j}$		$y^u_{3,0}$	$y^u_{3,1}$	$y^u_{3,2}$	$y^u_{3,3}$	$y^u_{3,4}$	$y^u_{3,5}$
4	$x_{4,\lambda}$ / $\ell_{4,j}$ / $r_{4,j}$		$y^u_{4,0}$	$y^u_{4,1}$	$y^u_{4,2}$	$y^u_{4,3}$	$y^u_{4,4}$	$y^u_{4,5}$

$$\Phi_n^{consistency} \quad := \quad \left[\bigwedge_{u \in P} \Phi_n^u \wedge y^u_{n,0} \right]$$

		u	$\{p\}$	$\{q\}$	$\emptyset$	$\{p,q\}$	$\emptyset$	$\{p,q\}$
1	$x_{1,\lambda}$ / $\ell_{1,j}$ / $r_{1,j}$		$y^u_{1,0}$	$y^u_{1,1}$	$y^u_{1,2}$	$y^u_{1,3}$	$y^u_{1,4}$	$y^u_{1,5}$
2	$x_{2,\lambda}$ / $\ell_{2,j}$ / $r_{2,j}$		$y^u_{2,0}$	$y^u_{2,1}$	$y^u_{2,2}$	$y^u_{2,3}$	$y^u_{2,4}$	$y^u_{2,5}$
3	$x_{3,\lambda}$ / $\ell_{3,j}$ / $r_{3,j}$		$y^u_{3,0}$	$y^u_{3,1}$	$y^u_{3,2}$	$y^u_{3,3}$	$y^u_{3,4}$	$y^u_{3,5}$
4	$x_{4,\lambda}$ / $\ell_{4,j}$ / $r_{4,j}$		$y^u_{4,0}$	$y^u_{4,1}$	$y^u_{4,2}$	$y^u_{4,3}$	$y^u_{4,4}$	$y^u_{4,5}$

$$\Phi_n^{consistency} := \left[\bigwedge_{u \in P} \Phi_n^u \wedge y^u_{n,0}\right] \wedge \left[\bigwedge_{u \in N} \Phi_n^u \wedge \neg y^u_{n,0}\right]$$

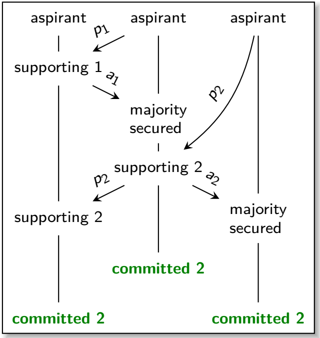
Remark

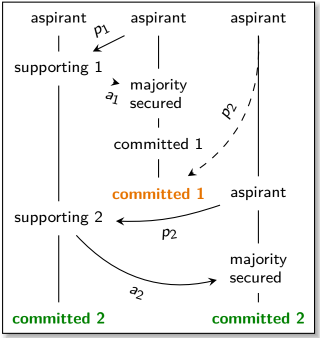
For a sample $\mathcal{S}$ and fixed n, the constructed propositional formula $\Phi_n^{\mathcal{S}}$ has the desired properties and ranges over $\mathcal{O}(n^2 + n \cdot |\mathcal{P}| + n \cdot |\mathcal{S}|)$ variables

Theorem

A minimal LTL formula consistent with a given sample $\mathcal{S}$ can be found by solving a sequence of SAT queries for $\Phi_n^{\mathcal{S}}$

Apache Zookeeper Leader Election





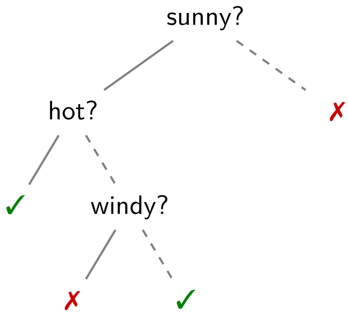
$$\psi \ := \ \neg rec(2,1) \,\mathsf{U}\, comm(1)$$

3. An Enhanced Algorithm Based on Decision Trees

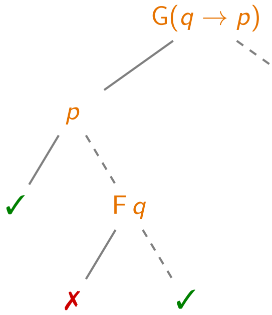
Sunny?	Hot?	Windy?	**Play golf or not**
false	true	true	✗
true	false	true	✗
false	false	false	✗
true	false	false	✓
true	true	true	✓

Decision Trees

Sunny?	Hot?	Windy?	**Play golf or not**
false	true	true	✗
true	false	true	✗
false	false	false	✗
true	false	false	✓
true	true	true	✓

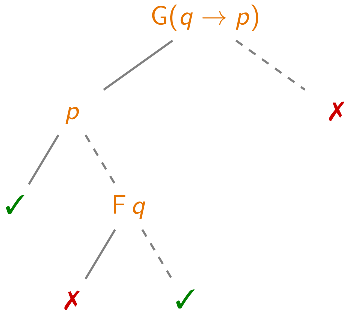


$\mathsf{G}(q \rightarrow p)$	p	$\mathsf{F}\,q$	**Classification**
false	true	true	✗
true	false	true	✗
false	false	false	✗
true	false	false	✓
true	true	true	✓

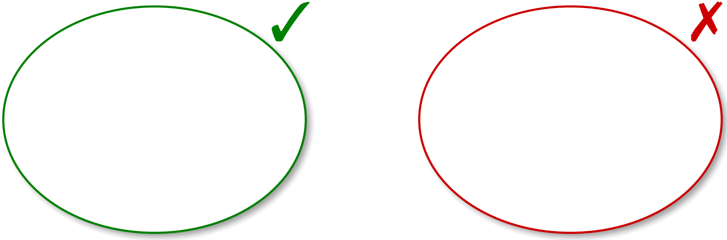


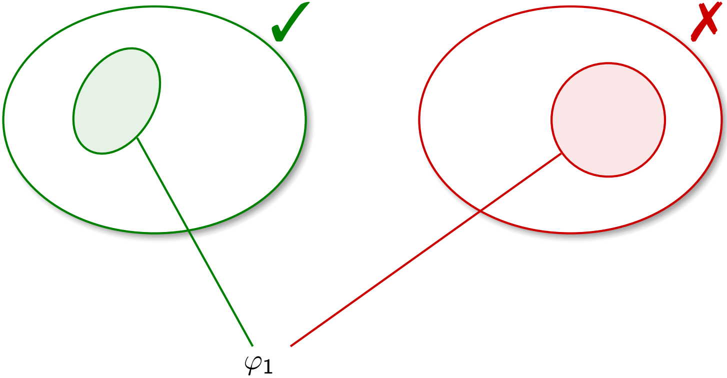
Decision Trees

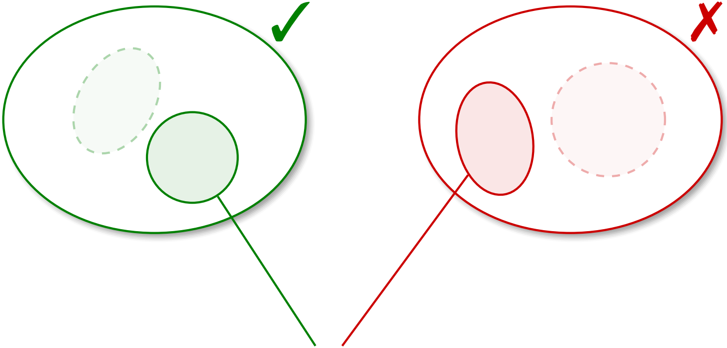
$\mathsf{G}(q \rightarrow p)$	p	$\mathsf{F}\,q$	**Classification**
false	true	true	✗
true	false	true	✗
false	false	false	✗
true	false	false	✓
true	true	true	✓



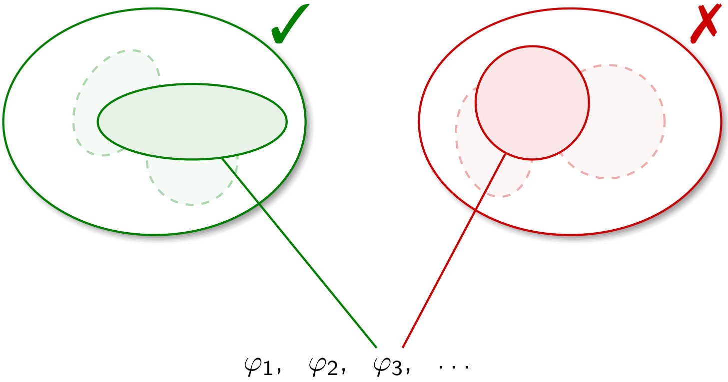
$$\big(\mathsf{G}(q \rightarrow p) \wedge p\big) \vee$$
$$\big(\mathsf{G}(q \rightarrow p) \wedge \neg p \wedge \neg\,\mathsf{F}\,q\big)$$




✓
✗
φ_1



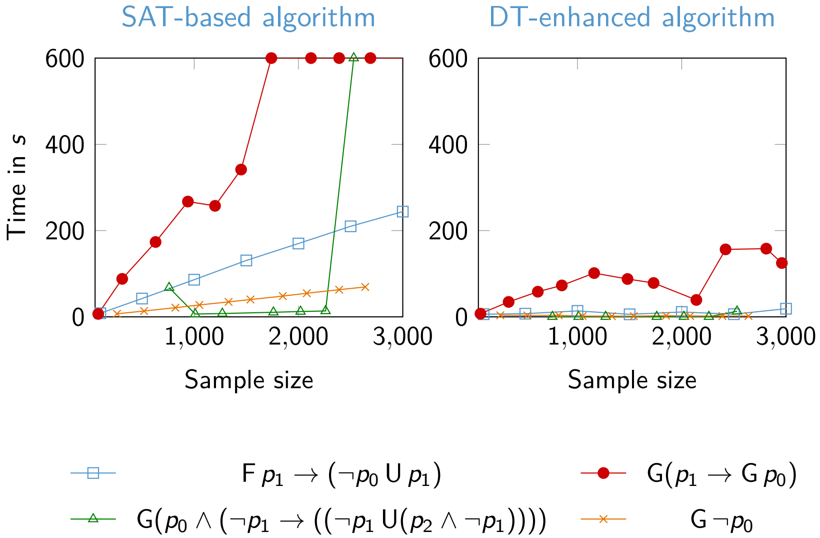
$\varphi_1, \quad \varphi_2$

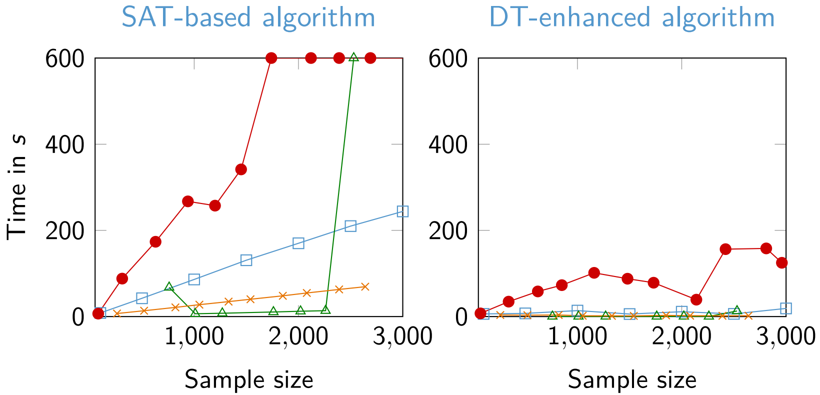


Strategy

Which sampling strategy to use?

Evaluation on Synthetic Benchmarks





Average increase in size: 1.41

4. Future Work

Try It!

You can download our tool from gitHub

`https://github.com/gergia/samples2LTL`

Ideas for Future Work

- Gain better understanding of the complexity
- Relax learning model, allow misclassification errors (MAX-SAT)
- Learning from positive data only
- Other declarative formalisms, such as (ω-)regular expressions, PSL, STL, CTL, UML Message Sequence Charts, etc.