Deciding WS1S Using an Automata-based Approach

Tomáš Fiedor^{1,2} Lukáš Holík²

¹Red Hat, Czech Republic

Ondřej Lengál² Tomáš Vojnar²

²Brno University of Technology, Czech Republic

Vienna UT, 2015

Weak monadic second-order logic of one successor:

- second-order ⇒ quantification over relations;
- monadic \Rightarrow the relations are unary i.e. sets;
- weak \Rightarrow the sets are finite;
- of one successor \Rightarrow reasoning about linear structures.

Weak monadic second-order logic of one successor:

- second-order ⇒ quantification over relations;
- monadic \Rightarrow the relations are unary i.e. sets;
- weak ⇒ the sets are finite;
- of one successor \Rightarrow reasoning about linear structures.

Extensions of WS1S:

- WSkS with k successors;
- SkS allows quantification over infinite sets;
- M2L(str) allows quantification over infinite (but bounded) sets.

Weak monadic second-order logic of one successor:

- second-order ⇒ quantification over relations;
- monadic \Rightarrow the relations are unary i.e. sets;
- weak ⇒ the sets are finite;
- of one successor \Rightarrow reasoning about linear structures.

Extensions of WS1S:

- WSkS with k successors;
- SkS allows quantification over infinite sets;
- M2L(str) allows quantification over infinite (but bounded) sets.
- Corresponds to finite automata [Büchi'60].

Weak monadic second-order logic of one successor:

- second-order ⇒ quantification over relations;
- monadic ⇒ the relations are unary i.e. sets;
- weak ⇒ the sets are finite;
- of one successor \Rightarrow reasoning about linear structures.

Extensions of WS1S:

- WSkS with k successors;
- SkS allows quantification over infinite sets;
- M2L(str) allows quantification over infinite (but bounded) sets.
- Corresponds to finite automata [Büchi'60].

Decidable, but NONELEMENTARY:

- tower of exponentials of height given by quantifier alternations.
 - Deciding WS1S via DFAs: determinization, complementation, ...

The MONA Tool

■ MONA – an automata-based WS1S/WSkS decision procedure:

- semi-symbolic DFAs/DTAs: MTBDDs used to encode transitions,
- efficient on many formulae obtained in various applications.
- Used in tools for checking complex shape invariants:
 - Pointer Assertion Logic Engine (PALE),
 - STRucture ANd Data (STRAND).
- Various other applications:
 - other kinds of program and protocol verification, parsing, synthesis, linguistics, multimedia, ...

The MONA Tool

■ MONA – an automata-based WS1S/WSkS decision procedure:

- semi-symbolic DFAs/DTAs: MTBDDs used to encode transitions,
- efficient on many formulae obtained in various applications.
- Used in tools for checking complex shape invariants:
 - Pointer Assertion Logic Engine (PALE),
 - STRucture ANd Data (STRAND).
- Various other applications:
 - other kinds of program and protocol verification, parsing, synthesis, linguistics, multimedia, ...
- However, sometimes the complexity strikes back:
 - unavoidable in general,
 - one can try to push the usability border further:
 - using the recent advancements in non-deterministic automata.

Syntax and Semantics of WS1S

Minimal syntax:

- Let X, Y, ... be 2nd-order variables.
- Terms: $\psi ::= X \subseteq Y \mid \operatorname{Sing}(X) \mid X = \{0\} \mid X = \sigma(Y)$
- Formulae: $\varphi ::= \psi \mid \varphi \land \varphi \mid \varphi \lor \varphi \mid \neg \varphi \mid \exists X. \varphi$

Syntax and Semantics of WS1S

Minimal syntax:

- Let X, Y, ... be 2nd-order variables.
- Terms: $\psi ::= X \subseteq Y \mid \text{Sing}(X) \mid X = \{0\} \mid X = \sigma(Y)$
- Formulae: $\varphi ::= \psi \mid \varphi \land \varphi \mid \varphi \lor \varphi \mid \neg \varphi \mid \exists X.\varphi$

A note on semantics:

- Variables interpreted as finite subsets of ℕ.
- Singleton Sing(X): $X = \{n\}$ for some $n \in \mathbb{N}$.
- Successor $X = \sigma(Y)$: $Y = \{n\}$ and $X = \{n+1\}$ for some $n \in \mathbb{N}$.

$$\varphi \Rightarrow \psi \quad \stackrel{\textit{def}}{\Leftrightarrow} \quad \neg \varphi \lor \psi$$

$$\begin{array}{ll} \varphi \Rightarrow \psi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \varphi \lor \psi \\ \forall \mathbf{X}. \varphi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \exists \mathbf{X}. \neg \varphi \end{array}$$

$$\begin{array}{ll} \varphi \Rightarrow \psi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \varphi \lor \psi \\ \forall X.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \exists X.\neg \varphi \\ \exists x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \exists X.\text{Sing}(X) \land \varphi \end{array}$$

$$\begin{array}{lll} \varphi \Rightarrow \psi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \varphi \lor \psi \\ \forall X.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \exists X.\neg \varphi \\ \exists x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \exists X.\text{Sing}(X) \land \varphi \\ \forall x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \forall X.\text{Sing}(X) \Rightarrow \varphi \end{array}$$

$$\begin{split} \varphi \Rightarrow \psi & \stackrel{\text{def}}{\Leftrightarrow} \quad \neg \varphi \lor \psi \\ \forall X.\varphi & \stackrel{\text{def}}{\Leftrightarrow} \quad \neg \exists X.\neg \varphi \\ \exists x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} \quad \exists X.\text{Sing}(X) \land \varphi \\ \forall x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} \quad \forall X.\text{Sing}(X) \Rightarrow \varphi \\ X = Y & \stackrel{\text{def}}{\Leftrightarrow} \quad X \subseteq Y \land Y \subseteq X \end{split}$$

$$\begin{array}{cccc} \varphi \Rightarrow \psi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \varphi \lor \psi \\ \forall X.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \exists X.\neg \varphi \\ \exists x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \exists X.\text{Sing}(X) \land \varphi \\ \forall x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \forall X.\text{Sing}(X) \Rightarrow \varphi \\ X = Y & \stackrel{\text{def}}{\Leftrightarrow} & X \subseteq Y \land Y \subseteq X \\ X = \emptyset & \stackrel{\text{def}}{\Leftrightarrow} & \forall Z.X \subseteq Z \end{array}$$

$$\begin{array}{cccc} \varphi \Rightarrow \psi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \varphi \lor \psi \\ \forall X.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \exists X.\neg \varphi \\ \exists x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \exists X.\text{Sing}(X) \land \varphi \\ \forall x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \forall X.\text{Sing}(X) \Rightarrow \varphi \\ X = Y & \stackrel{\text{def}}{\Leftrightarrow} & X \subseteq Y \land Y \subseteq X \\ X = \emptyset & \stackrel{\text{def}}{\Leftrightarrow} & \forall Z.X \subseteq Z \\ x \in Y & \stackrel{\text{def}}{\Leftrightarrow} & \text{Sing}(X) \land X \subseteq Y \end{array}$$

$$\begin{array}{lll} \varphi \Rightarrow \psi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \varphi \lor \psi \\ \forall X.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \neg \exists X.\neg \varphi \\ \exists x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \exists X.\text{Sing}(X) \land \varphi \\ \forall x.\varphi & \stackrel{\text{def}}{\Leftrightarrow} & \forall X.\text{Sing}(X) \Rightarrow \varphi \\ X = Y & \stackrel{\text{def}}{\Leftrightarrow} & X \subseteq Y \land Y \subseteq X \\ X = \emptyset & \stackrel{\text{def}}{\Leftrightarrow} & \forall Z.X \subseteq Z \\ x \in Y & \stackrel{\text{def}}{\Leftrightarrow} & \text{Sing}(X) \land X \subseteq Y \\ x \leq y & \stackrel{\text{def}}{\Leftrightarrow} & \forall X. (y \in X \land (\forall z.\sigma(z) \in X \Rightarrow z \in X)) \Rightarrow x \in X \end{array}$$

Models of WS1S formulae:

• maps of the variables to finite subsets of \mathbb{N} .

Models of WS1S formulae:

• maps of the variables to finite subsets of N.

Such sets can be encoded as binary strings:

		Index:	012345	0123456		01234567	
•	$\{1, 4, 5\} \rightarrow$	Membership:	x√xx√√ ,	X√XX√√X	or	X √ XX √ √ XX	
		Encoding:	010011	0100110		01001100	

Models of WS1S formulae:

• maps of the variables to finite subsets of N.

Such sets can be encoded as binary strings:

- Note that any number of zeros can always be added to the right!

Models of WS1S formulae:

• maps of the variables to finite subsets of N.

Such sets can be encoded as binary strings:

- Note that any number of zeros can always be added to the right!

For each variable, we have one track in the alphabet:

• e.g., $\begin{bmatrix} 0\\0 \end{bmatrix}$ is a symbol.

Models of WS1S formulae:

• maps of the variables to finite subsets of N.

Such sets can be encoded as binary strings:

- Note that any number of zeros can always be added to the right!

For each variable, we have one track in the alphabet:

- e.g., $\begin{vmatrix} 0 \\ 0 \end{vmatrix}$ is a symbol.
- A natural encoding for sets of transitions over such symbols:
 - multi-terminal BDDs.

Models of WS1S formulae:

• maps of the variables to finite subsets of N.

Such sets can be encoded as binary strings:

- $\bullet \ \{1,4,5\} \rightarrow \begin{array}{cccc} \text{Index:} & 012345 & 0123456 & 01234567 \\ \text{Membership:} & \textbf{x} \swarrow \textbf{x} \checkmark \checkmark \checkmark & \textbf{or} & \textbf{x} \And \textbf{x} \checkmark \checkmark \textbf{x} \checkmark \checkmark \textbf{x} \\ \text{Encoding:} & 010011 & 0100110 & 01001100 \end{array}$
- Note that any number of zeros can always be added to the right!

For each variable, we have one track in the alphabet:

- e.g., $\begin{bmatrix} 0\\0 \end{bmatrix}$ is a symbol.
- A natural encoding for sets of transitions over such symbols:
 - multi-terminal BDDs.

Example:
$$\{X_1 \mapsto \emptyset, X_2 \mapsto \{4, 2\}\} \rightsquigarrow X_1: \begin{bmatrix} 0 \\ X_2: \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
.

 $X \subseteq Y$

(X is a subset of Y)



$$\blacksquare \begin{array}{ccc} X \mapsto \{ 2, 4 \} \\ Y \mapsto \{1,2,3,4\} \models X \subseteq Y \quad \rightsquigarrow \quad \begin{array}{c} X \colon \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix}^* \subseteq L(\mathcal{A}_{X \subseteq Y}) \end{array}$$

T. Fiedor, L. Holík, O. Lengál, T. Vojnar

WS1S through NFA

Vienna UT, 2015 7 / 25



T. Fiedor, L. Holík, O. Lengál, T. Vojnar

WS1S through NFA







Logical connectives mapped to automata operations.

• More complex formulae handled by composition of automata.

Logical connectives mapped to automata operations.

- More complex formulae handled by composition of automata.
- Example:

$$\neg (X \subseteq Y) \land \exists Z.(\operatorname{Sing}(Z) \lor \exists W.W = \sigma(Z))$$
$$| \qquad | \qquad | \\ \mathcal{A}_3 \qquad \mathcal{A}_2 \qquad \mathcal{A}_1$$

Logical connectives mapped to automata operations.

• More complex formulae handled by composition of automata.

Logical connectives mapped to automata operations.

• More complex formulae handled by composition of automata.

Logical connectives mapped to automata operations.

• More complex formulae handled by composition of automata.



Logical connectives mapped to automata operations.

• More complex formulae handled by composition of automata.



Logical connectives mapped to automata operations.

• More complex formulae handled by composition of automata.



Handling Quantification

- Quantification handled by projection of a certain track.
 - Introduces non-determinism to automata.
 - Alternations require negation, hence determinization ~> high cost.

Handling Quantification

- Quantification handled by projection of a certain track.
 - Introduces non-determinism to automata.
 - Alternations require negation, hence determinization ~> high cost.
- An additional issue with projection of variables in an automaton:
 - after removing some tracks, some models need not be accepted:
 - some zero suffixes need not be present.
 - One needs to adjust the final states.


Handling Quantification

- Quantification handled by projection of a certain track.
 - Introduces non-determinism to automata.
 - Alternations require negation, hence determinization ~> high cost.
- An additional issue with projection of variables in an automaton:
 - after removing some tracks, some models need not be accepted:
 - some zero suffixes need not be present.
 - One needs to adjust the final states.



Handling Quantification

- Quantification handled by projection of a certain track.
 - Introduces non-determinism to automata.
 - Alternations require negation, hence determinization ~> high cost.
- An additional issue with projection of variables in an automaton:
 - after removing some tracks, some models need not be accepted:
 - some zero suffixes need not be present.
 - One needs to adjust the final states.



Handling Quantification

- Quantification handled by projection of a certain track.
 - Introduces non-determinism to automata.
 - Alternations require negation, hence determinization ~> high cost.
- An additional issue with projection of variables in an automaton:
 - after removing some tracks, some models need not be accepted:
 - some zero suffixes need not be present.
 - One needs to adjust the final states.



- We consider only formulae in Prenex Normal Form (∃PNF).
- We focus on dealing with the prefix of alternating quantifiers.

- We consider only formulae in Prenex Normal Form (∃PNF).
- We focus on dealing with the prefix of alternating quantifiers.
- For a formula φ with *m* quantifier alternations over sets of variables X_i ⊆ X,



- We consider only formulae in Prenex Normal Form (∃PNF).
- We focus on dealing with the prefix of alternating quantifiers.
- For a formula φ with *m* quantifier alternations over sets of variables X_i ⊆ X,



- \rightarrow construct a hierarchical family of automata defined as follows:
 - A_{\varphi_0}: a composition of atomic automata described before,

•
$$\mathcal{A}_{\varphi_m} = (\underbrace{2^{2}}_{m}, \Delta_m, I_m, F_m)$$
: described next.

2Qn

- We consider only formulae in Prenex Normal Form (∃PNF).
- We focus on dealing with the prefix of alternating quantifiers.
- For a formula φ with *m* quantifier alternations over sets of variables X_i ⊆ X,



- \rightarrow construct a hierarchical family of automata defined as follows:
 - \mathcal{A}_{φ_0} : a composition of atomic automata described before,

•
$$\mathcal{A}_{\varphi_m} = (\underline{2^2}, \Delta_m, I_m, F_m)$$
: described next.

Intuition: on-the-fly projection and subset construction for all *m* levels (instead of doing it one-by-one), with antichain pruning.

T. Fiedor, L. Holík, O. Lengál, T. Vojnar

WS1S through NFA

Correspondence between formulae and automata:

• A formula φ is valid iff $L(\mathcal{A}_{\varphi}) = \Sigma^*$.

Correspondence between formulae and automata:

- A formula φ is valid iff $L(\mathcal{A}_{\varphi}) = \Sigma^*$.
- A formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) \neq \emptyset$.

Correspondence between formulae and automata:

- A formula φ is valid iff $L(\mathcal{A}_{\varphi}) = \Sigma^*$.
- A formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) \neq \emptyset$.
- A formula φ is unsatisfiable iff $L(\mathcal{A}_{\varphi}) = \emptyset$.

Correspondence between formulae and automata:

- A formula φ is valid iff $L(\mathcal{A}_{\varphi}) = \Sigma^*$.
- A formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) \neq \emptyset$.
- A formula φ is unsatisfiable iff $L(\mathcal{A}_{\varphi}) = \emptyset$.

A key observation for ground formulae:

The symbols degenerate to the empty one: $\Sigma = \{[]\}$.

Correspondence between formulae and automata:

- A formula φ is valid iff $L(\mathcal{A}_{\varphi}) = \Sigma^*$.
- A formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) \neq \emptyset$.
- A formula φ is unsatisfiable iff $L(\mathcal{A}_{\varphi}) = \emptyset$.

A key observation for ground formulae:

- The symbols degenerate to the empty one: $\Sigma = \{[]\}$.
- A ground formula φ is valid iff it is satisfiable iff $L(\mathcal{A}_{\varphi}) = []^*$.

Correspondence between formulae and automata:

- A formula φ is valid iff $L(\mathcal{A}_{\varphi}) = \Sigma^*$.
- A formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) \neq \emptyset$.
- A formula φ is unsatisfiable iff $L(\mathcal{A}_{\varphi}) = \emptyset$.

A key observation for ground formulae:

- The symbols degenerate to the empty one: $\Sigma = \{[]\}$.
- A ground formula φ is valid iff it is satisfiable iff $L(\mathcal{A}_{\varphi}) = []^*$.
- A ground formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) = []^*$ iff $I_m \cap F_m \neq \emptyset$.

Correspondence between formulae and automata:

- A formula φ is valid iff $L(\mathcal{A}_{\varphi}) = \Sigma^*$.
- A formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) \neq \emptyset$.
- A formula φ is unsatisfiable iff $L(\mathcal{A}_{\varphi}) = \emptyset$.

A key observation for ground formulae:

- The symbols degenerate to the empty one: $\Sigma = \{[]\}$.
- A ground formula φ is valid iff it is satisfiable iff $L(\mathcal{A}_{\varphi}) = []^*$.
- A ground formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) = []^*$ iff $I_m \cap F_m \neq \emptyset$.



T. Fiedor, L. Holík, O. Lengál, T. Vojnar

Correspondence between formulae and automata:

- A formula φ is valid iff $L(\mathcal{A}_{\varphi}) = \Sigma^*$.
- A formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) \neq \emptyset$.
- A formula φ is unsatisfiable iff $L(\mathcal{A}_{\varphi}) = \emptyset$.

A key observation for ground formulae:

- The symbols degenerate to the empty one: $\Sigma = \{[]\}$.
- A ground formula φ is valid iff it is satisfiable iff $L(\mathcal{A}_{\varphi}) = []^*$.
- A ground formula φ is satisfiable iff $L(\mathcal{A}_{\varphi}) = []^*$ iff $I_m \cap F_m \neq \emptyset$.



T. Fiedor, L. Holík, O. Lengál, T. Vojnar

Constructing the whole automaton for φ_m is not necessary!

- We construct the sets of initial/final states only,
- and test whether they intersect.

Constructing the whole automaton for φ_m **is not necessary**!

- · We construct the sets of initial/final states only,
- and test whether they intersect.
- Construction of initial states is straightforward; starting from *l*₀:

Constructing the whole automaton for φ_m **is not necessary**!

- · We construct the sets of initial/final states only,
- and test whether they intersect.
- Construction of initial states is straightforward; starting from I₀:
 - $I_1 = \{I_0\}$

Constructing the whole automaton for φ_m is not necessary!

- · We construct the sets of initial/final states only,
- and test whether they intersect.
- Construction of initial states is straightforward; starting from I₀:

•
$$I_1 = \{I_0\}$$

• $I_2 = \{I_1\} = \{\{I_0\}\}$

Constructing the whole automaton for φ_m is not necessary!

- We construct the sets of initial/final states only,
- and test whether they intersect.
- Construction of initial states is straightforward; starting from I₀:

•
$$I_1 = \{I_0\}$$

• $I_2 = \{I_1\} = \{\{I_0\}\}$
:
• $I_m = \{I_{m-1}\} = \underbrace{\{\{\dots, \{I_0\}, \dots\}\}}_{m}$

No space for the subset construction to construct anything else!

Constructing the whole automaton for φ_m is not necessary!

- · We construct the sets of initial/final states only,
- and test whether they intersect.
- Construction of initial states is straightforward; starting from *l*₀:

•
$$I_1 = \{I_0\}$$

• $I_2 = \{I_1\} = \{\{I_0\}\}$

$$I_m = \{I_{m-1}\} = \{\{\dots, \{I_0\}, \dots\}\}$$

No space for the subset construction to construct anything else!

Final states are more tricky:

- · a need to saturate after projection as described previously,
- a lot of space for constructing different sets of sets of ... of states,
- a need of switching the acceptance mode.

- Given a formula $\varphi = \neg \exists \mathcal{X}_m \neg \ldots \neg \exists \mathcal{X}_2 \neg \exists \mathcal{X}_1 : \varphi_0(\mathbb{X})$ in $\exists \mathsf{PNF}$,
 - start by constructing the base automaton for φ_0 .

- start by constructing the base automaton for φ_0 .
- Given final states F_i of level i,
 - compute the set F[∃]_i of their predecessors over 0 (pre₀) after projecting X_{i+1},

- start by constructing the base automaton for φ_0 .
- Given final states F_i of level i,
 - compute the set F[∃]_i of their predecessors over 0 (pre₀) after projecting X_{i+1},
 - after subset construction, any set of states of level *i* containing a state from F[∃]_i is final,

- start by constructing the base automaton for φ_0 .
- Given final states F_i of level i,
 - compute the set F[∃]_i of their predecessors over 0 (pre₀) after projecting X_{i+1},
 - after subset construction, any set of states of level *i* containing a state from F[∃]_i is final,
 - after negation, such sets become non-final.

- start by constructing the base automaton for φ_0 .
- Given final states F_i of level i,
 - compute the set F[∃]_i of their predecessors over 0 (pre₀) after projecting X_{i+1},
 - after subset construction, any set of states of level *i* containing a state from F[∃]_i is final,
 - after negation, such sets become non-final.
 - We stay with non-final states.

- start by constructing the base automaton for φ_0 .
- Given final states *F_i* of level *i*,
 - compute the set F[∃]_i of their predecessors over 0 (pre₀) after projecting X_{i+1},
 - after subset construction, any set of states of level *i* containing a state from F[∃]_i is final,
 - after negation, such sets become non-final.
 - We stay with non-final states.
 - Hence, non-final states $N_{i+1} = \uparrow \{\{q\} \mid q \in F_i^{\exists}\},\$
 - ▶ the upward closed set with the set of generators (minimal elements) $\{\{q\} \mid q \in F_i^{\exists}\}.$

- start by constructing the base automaton for φ_0 .
- Given final states *F_i* of level *i*,
 - compute the set F[∃]_i of their predecessors over 0 (pre₀) after projecting X_{i+1},
 - after subset construction, any set of states of level *i* containing a state from F[∃]_i is final,
 - after negation, such sets become non-final.
 - We stay with non-final states.
 - Hence, non-final states $N_{i+1} = \uparrow \{\{q\} \mid q \in F_i^{\exists}\},\$
 - ▶ the upward closed set with the set of generators (minimal elements) $\{\{q\} \mid q \in F_i^{\exists}\}.$
 - i.e., $N_{i+1} = \uparrow \coprod \{F_i^{\exists}\}$ choice (unordered Cartesian product).

- start by constructing the base automaton for φ_0 .
- Given final states *F_i* of level *i*,
 - compute the set F[∃]_i of their predecessors over 0 (pre₀) after projecting X_{i+1},
 - after subset construction, any set of states of level *i* containing a state from F[∃]_i is final,
 - after negation, such sets become non-final.
 - We stay with non-final states.
 - Hence, non-final states $N_{i+1} = \uparrow \{\{q\} \mid q \in F_i^{\exists}\},\$
 - ► the upward closed set with the set of generators (minimal elements) $\{\{q\} \mid q \in F_i^{\exists}\}.$
 - i.e., $N_{i+1} = \uparrow \coprod \{F_i^{\exists}\}$ choice (unordered Cartesian product).
 - ▶ Let $Q = \{Q_1, ..., Q_n\}, \coprod Q = \{\{q_1, ..., q_n\} \mid (q_1, ..., q_n) \in \prod_{i=1}^n Q_i\}.$

Example: projection of *X* and negation on the FA:



Example: projection of *X* and negation on the FA:



- After projection:
 - $F_0^{\exists} = \{2,3\}.$

Example: projection of *X* and negation on the FA:



After projection:

• $F_0^{\exists} = \{2, 3\}.$

After negation:

• $N_1 = \{\{2\}, \{3\}, \{2,0\}, \{3,0\}, \dots, \{2,3,0\}, \{2,3,1\}, \dots, \{0,1,2,3\}\},\$

Example: projection of *X* and negation on the FA:



After projection:

• $F_0^{\exists} = \{2, 3\}.$

After negation:

- $N_1 = \{\{2\}, \{3\}, \{2,0\}, \{3,0\}, \dots, \{2,3,0\}, \{2,3,1\}, \dots, \{0,1,2,3\}\},\$
- $N_1 = \uparrow \{\{2\}, \{3\}\} = \uparrow \coprod \{\{2, 3\}\}.$

Given non-final states N_i of level *i*,

- compute the set N_i^{\exists} of their controllable predecessors over 0 (*cpre*₀) after projecting \mathcal{X}_{i+1} ,
 - only states that cannot get to a final state stay non-final,

Given non-final states N_i of level i,

- compute the set N_i^{\exists} of their controllable predecessors over 0 (*cpre*₀) after projecting \mathcal{X}_{i+1} ,
 - only states that cannot get to a final state stay non-final,
- after subset construction, any set of states of level *i* consisting of non-final states of N_i[∃] is non-final,

- Given non-final states N_i of level i,
 - compute the set N_i^{\exists} of their controllable predecessors over 0 (*cpre*₀) after projecting \mathcal{X}_{i+1} ,
 - only states that cannot get to a final state stay non-final,
 - after subset construction, any set of states of level *i* consisting of non-final states of N_i[∃] is non-final,
 - after negation, any such set becomes final.
Introduction to the Computation of Final States (3/3)

- Given non-final states N_i of level *i*,
 - compute the set N_i^{\exists} of their controllable predecessors over 0 (*cpre*₀) after projecting \mathcal{X}_{i+1} ,
 - only states that cannot get to a final state stay non-final,
 - after subset construction, any set of states of level *i* consisting of non-final states of N_i[∃] is non-final,
 - after negation, any such set becomes final.
 - Hence, final states $F_{i+1} = \bigcup \{N_i^{\exists}\}$ (downward closure).

Introduction to the Computation of Final States (3/3)

- Given non-final states N_i of level *i*,
 - compute the set N_i^{\exists} of their controllable predecessors over 0 (*cpre*₀) after projecting \mathcal{X}_{i+1} ,
 - only states that cannot get to a final state stay non-final,
 - after subset construction, any set of states of level *i* consisting of non-final states of N_i[∃] is non-final,
 - after negation, any such set becomes final.
 - Hence, final states $F_{i+1} = \bigcup \{N_i^{\exists}\}$ (downward closure).
- Continue with iterating the computation of non-final states from final, final from non-final, ...

Introduction to the Computation of Final States (3/3)

- Given non-final states N_i of level *i*,
 - compute the set N_i^{\exists} of their controllable predecessors over 0 (*cpre*₀) after projecting \mathcal{X}_{i+1} ,
 - only states that cannot get to a final state stay non-final,
 - after subset construction, any set of states of level *i* consisting of non-final states of N_i[∃] is non-final,
 - after negation, any such set becomes final.
 - Hence, final states $F_{i+1} = \bigcup \{N_i^{\exists}\}$ (downward closure).
- Continue with iterating the computation of non-final states from final, final from non-final, ...
- Do not enumerate the sets F_i/N_i :
 - use symbolic encoding via expressions with the $\uparrow \coprod / \downarrow$ operators.
 - A form of antichain reduction: keeping minimal/maximal elements.

• Given $\varphi = \neg \exists \mathcal{X}_m \neg \ldots \neg \exists \mathcal{X}_2 \neg \exists \mathcal{X}_1 : \varphi_0(\mathbb{X}).$

- Given $\varphi = \neg \exists \mathcal{X}_m \neg \ldots \neg \exists \mathcal{X}_2 \neg \exists \mathcal{X}_1 : \varphi_0(\mathbb{X}).$
- **1** Add final states after $\exists: F_0^{\exists} = \{\mu Z.F \cup \text{pre}_0(Z)\}.$

• Given
$$\varphi = \neg \exists \mathcal{X}_m \neg \ldots \neg \exists \mathcal{X}_2 \neg \exists \mathcal{X}_1 : \varphi_0(\mathbb{X}).$$

- **1** Add final states after $\exists: F_0^{\exists} = \{\mu Z.F \cup \text{pre}_0(Z)\}.$
- 2 Negate the final states: $N_1 = \uparrow \coprod \{F_0^{\exists}\}.$

- Given $\varphi = \neg \exists \mathcal{X}_m \neg \ldots \neg \exists \mathcal{X}_2 \neg \exists \mathcal{X}_1 : \varphi_0(\mathbb{X}).$
- **1** Add final states after $\exists: F_0^{\exists} = \{\mu Z.F \cup \text{pre}_0(Z)\}.$
- **2** Negate the final states: $N_1 = \uparrow \coprod \{F_0^{\exists}\}.$
- 3 Prune non-final states after $\exists: N_1^{\exists} = \{\nu Z.N_1 \cap \operatorname{cpre}_0(Z)\}.$

- Given $\varphi = \neg \exists \mathcal{X}_m \neg \ldots \neg \exists \mathcal{X}_2 \neg \exists \mathcal{X}_1 : \varphi_0(\mathbb{X}).$
- **1** Add final states after $\exists: F_0^{\exists} = \{\mu Z.F \cup \text{pre}_0(Z)\}.$
- **2** Negate the final states: $N_1 = \uparrow \coprod \{F_0^{\exists}\}.$
- **3** Prune non-final states after $\exists: N_1^{\exists} = \{\nu Z.N_1 \cap \operatorname{cpre}_0(Z)\}.$
- 4 Negate the non-final states: $F_2 = \downarrow \{N_1^{\exists}\}$.

- Given $\varphi = \neg \exists \mathcal{X}_m \neg \ldots \neg \exists \mathcal{X}_2 \neg \exists \mathcal{X}_1 : \varphi_0(\mathbb{X}).$
- **1** Add final states after $\exists: F_0^{\exists} = \{\mu Z.F \cup pre_0(Z)\}.$
- **2** Negate the final states: $N_1 = \uparrow \coprod \{F_0^{\exists}\}.$
- **3** Prune non-final states after $\exists: N_1^{\exists} = \{\nu Z.N_1 \cap \operatorname{cpre}_0(Z)\}.$
- 4 Negate the non-final states: $F_2 = \downarrow \{N_1^{\exists}\}$.
- 5 Keep alternating between computing final and non-final states until F_m as follows:

•
$$F_{i+1} = \downarrow \{\nu Z.N_i \cap \operatorname{cpre}_0(Z)\},\$$

• $N_{i+1} = \uparrow \coprod \{ \mu Z.F_i \cup \operatorname{pre}_0(Z) \}.$

 (Controllable) predecessors of symbolic states can be computed without enumerating their elements.

- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:

- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



Likewise for the predecessors of a downward closed set.

- (Controllable) predecessors of symbolic states can be computed without enumerating their elements.
- E.g., note that $cpre_0(\uparrow \{Q\}) = \uparrow \coprod \{pre_0(q) \mid q \in Q\}$:



- Likewise for the predecessors of a downward closed set.
- Can be adapted for symbolic states with the needed structure.









T. Fiedor, L. Holík, O. Lengál, T. Vojnar

WS1S through NFA



T. Fiedor, L. Holík, O. Lengál, T. Vojnar



T. Fiedor, L. Holík, O. Lengál, T. Vojnar

State Space Pruning

- Sets of states on the various levels of the subset construction encoded as up(down)ward closed sets given by their generators.
 - First source of reduction.
 - Can be viewed as having an antichain flavour: generators and the bigger/smaller states are comparable.

State Space Pruning

- Sets of states on the various levels of the subset construction encoded as up(down)ward closed sets given by their generators.
 - First source of reduction.
 - Can be viewed as having an antichain flavour: generators and the bigger/smaller states are comparable.
- Further, we prune the generators subsumed by other generators:
 - the subsumption relation is computed on nested structure of symbolic representation of lower levels as follows.

$$\downarrow \mathbb{X} \subseteq \downarrow \mathbb{Y} \qquad \Longleftrightarrow \qquad \forall X \in \mathbb{X} . \exists Y \in \mathbb{Y} . X \subseteq Y$$

$$\uparrow \coprod \mathbb{X} \subseteq \uparrow \coprod \mathbb{Y} \qquad \Longleftrightarrow \qquad \forall Y \in \mathbb{Y} . \exists X \in \mathbb{X} . X \subseteq Y$$

Can also be done on the symbolic representation.

- Can also be done on the symbolic representation.
- Depending on whether the number of alternations is even or odd, test:

- Can also be done on the symbolic representation.
- Depending on whether the number of alternations is even or odd, test:
 - · whether initial states intersect the final ones,

• $I_m \cap F_m \neq \emptyset \Leftrightarrow I_{m-1} \in F_m$, or

- Can also be done on the symbolic representation.
- Depending on whether the number of alternations is even or odd, test:
 - · whether initial states intersect the final ones,

•
$$I_m \cap F_m \neq \emptyset \iff I_{m-1} \in F_m$$
, or

- · or that initial states are not among the non-final ones,
 - reduces to an "and/or" search:

$$\{x\} \in \downarrow \mathbb{Y} \qquad \Longleftrightarrow \qquad \exists Y \in \mathbb{Y} : x \in Y \\ \{x\} \in \uparrow \coprod \mathbb{Y} \qquad \Longleftrightarrow \qquad \forall Y \in \mathbb{Y} : x \in Y$$

Implementations

dWiNA (deciding WS1S using Non-deterministic Automata):

- our prototype implementation,
- antichain-based approach, with non-deterministic automata,
- uses library VATA for manipulation with the automata:
 - uses degenerated tree automata.
- MONA:
 - (old but) state-of-the-art tool,
 - · classic approach, with deterministic automata,
 - implemented range of optimizations like:
 - automata minimization,
 - automata caching,
 - using a DAG representation for formulae,
 - and many others.

Experiments on Formulae from Verification

Compared with MONA:

- on formulae from verification benchmarks,
 - taken from the STRAND tool (STRucture ANd Data),
 - encoding loop invariants of heap-manipulating programs,
- in the general and $\exists PNF$ form.

		M	ANC	dWiNA		
	Time [s]		Space [states]		Time [s]	Space [states]
benchmark	general	∃PNF	general	∃PNF	Prefix	Prefix
list-insert-after-loop	0.01	0.01	167	686	0.01	28
list-insert-before-head	0.01	0.01	43	152	0.01	38
list-insert-before-loop	0.01	0.01	103	1021	0.01	38
list-insert-in-loop	0.01	0.01	463	5015	0.01	59
list-reverse-after-loop	0.01	0.01	179	1 326	0.01	100
list-reverse-in-loop	0.02	0.47	1311	70278	0.02	260
bubblesort-else	0.01	0.45	1 285	12071	0.01	14
bubblesort-if-else	0.02	2.17	4 260	116760	0.23	234
bubblesort-if-if	0.12	5.29	8 390	233 372	1.14	28

Experiments with Generated Formulae

- Compared with MONA:
 - on generated formulae,
 - parametric, various lengths of prefix, number of alternations,
 - base formulae encode various set problems (transitivity, etc.),
 - in the ∃PNF form.

An example of a generated formula:

$$\exists Y : \neg \exists X_1 \neg \ldots \neg \exists X_k, \ldots, X_n : \bigwedge_{1 \leq i < n} (X_i \subseteq Y \land X_i \subset X_{i+1}) \Rightarrow X_{i+1} \subseteq Y.$$

		M	ANC	dWiNA		
	Time [s]		Space [states]		Time [s]	Space [states]
benchmark	general	∃PNF	general	∃PNF	Prefix	Prefix
1 alternation	-	0.11	-	10718	0.01	39
2 alternations	-	0.20	-	25517	0.01	44
3 alternations	-	0.57	-	60 924	0.01	50
4 alternations	-	1.79	-	145765	0.02	58
5 alternations	-	4.98	-	349314	0.02	70
6 alternations	-	то	-	то	0.47	90
Future Work

- Extension to WS2S.
- Generalization of the symbolic tree representation:
 - to process logical connectives,
 - to handle general (non-∃PNF) formulae.

Syntactical optimizations:

- using Direct Acyclic Graph (DAG) for representation of formulae,
- anti-prenexing,
- smarter conversion to ∃PNF, ...