

Verified Decision Procedures

for

Monadic Second-Order Logic on Strings

Functional Pearl

Dmitriy Traytel Tobias Nipkow



Technische Universität München



Overview

MSO

Overview

MSO

$$\mathcal{L}_{\text{MSO}}(\varphi) = \mathcal{L}_{\text{MSO}}(\psi)?$$

Overview

Finite Automata

MSO

$$\mathcal{L}_{\text{MSO}}(\varphi) = \mathcal{L}_{\text{MSO}}(\psi)?$$

Overview

Finite Automata



MONA (> 40 kLOC of C/C++)

MSO

$$\mathcal{L}_{\text{MSO}}(\varphi) = \mathcal{L}_{\text{MSO}}(\psi)?$$

Overview

Finite Automata



MONA (> 40 kLOC of C/C++)

MSO

$$\mathcal{L}_{\text{MSO}}(\varphi) = \mathcal{L}_{\text{MSO}}(\psi)?$$

Regular Expressions

Overview

Finite Automata



MONA (> 40 kLOC of C/C++)

MSO

$$\mathcal{L}_{\text{MSO}}(\varphi) = \mathcal{L}_{\text{MSO}}(\psi)?$$

Regular Expressions

$$\mathcal{L}(\alpha) = \mathcal{L}(\beta)?$$

Overview

Finite Automata

↑
MONA (> 40 kLOC of C/C++)

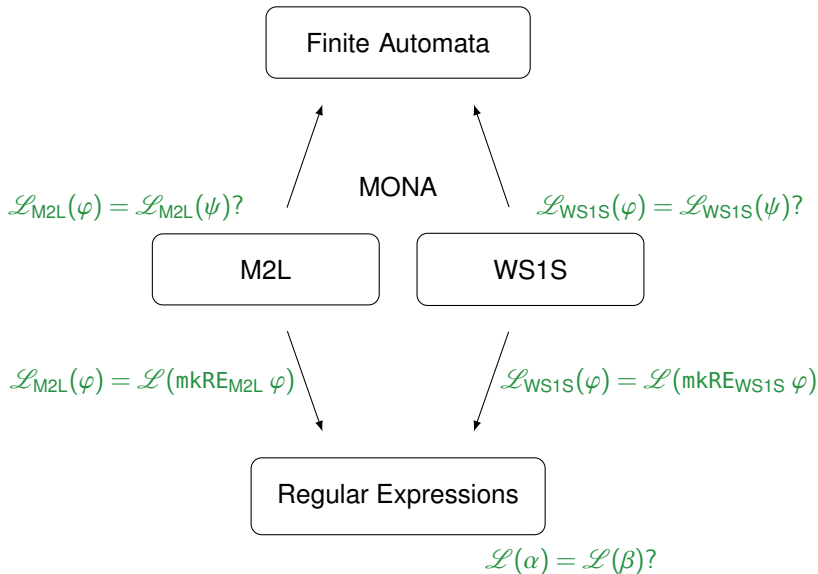
MSO

↓
 $\mathcal{L}_{\text{MSO}}(\varphi) = \mathcal{L}(\text{mkRE } \varphi)$
 $\mathcal{L}_{\text{MSO}}(\varphi) = \mathcal{L}_{\text{MSO}}(\psi)?$

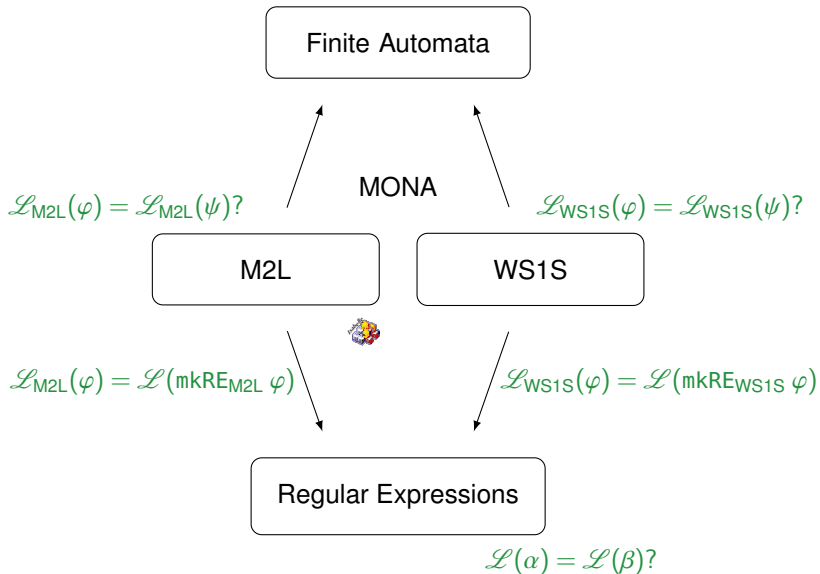
Regular Expressions

$\mathcal{L}(\alpha) = \mathcal{L}(\beta)?$

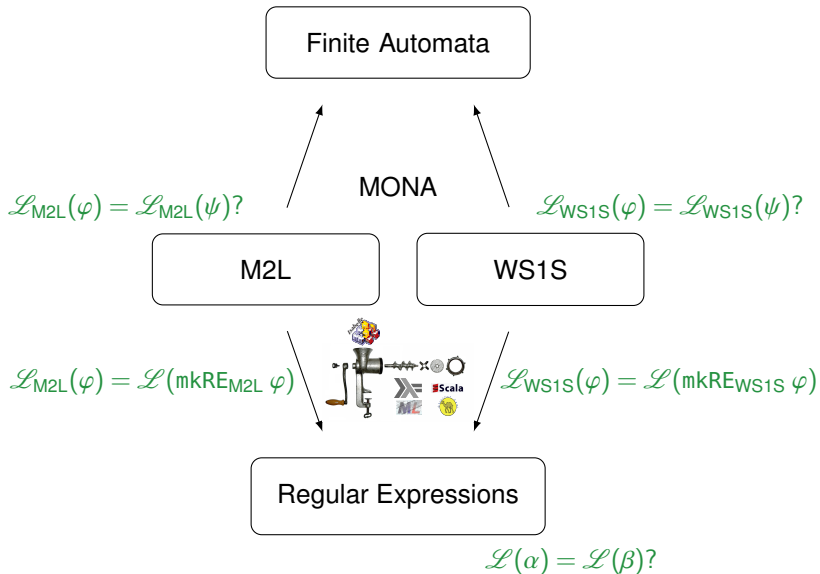
Overview



Overview



Overview



Outline

Regular Expressions Equivalence

MSO

Outline

Regular Expressions Equivalence

MSO

Regular Expressions

$$\mathcal{L}(\emptyset) = \{\}$$

$$\mathcal{L}(\varepsilon) = \{\emptyset\}$$

$$\mathcal{L}(a) = \{[a]\} \quad a \in \Sigma$$

$$\mathcal{L}(\alpha + \beta) = \mathcal{L}(\alpha) \cup \mathcal{L}(\beta)$$

$$\mathcal{L}(\alpha \cdot \beta) = \mathcal{L}(\alpha) \cdot \mathcal{L}(\beta)$$

$$\mathcal{L}(\alpha^*) = \mathcal{L}(\alpha)^*$$

Extended Regular Expressions

$$\mathcal{L}(\emptyset) = \{\}$$

$$\mathcal{L}(\varepsilon) = \{\emptyset\}$$

$$\mathcal{L}(a) = \{[a]\} \quad a \in \Sigma$$

$$\mathcal{L}(\alpha + \beta) = \mathcal{L}(\alpha) \cup \mathcal{L}(\beta)$$

$$\mathcal{L}(\alpha \cdot \beta) = \mathcal{L}(\alpha) \cdot \mathcal{L}(\beta)$$

$$\mathcal{L}(\alpha^*) = \mathcal{L}(\alpha)^*$$

$$\mathcal{L}(\alpha \cap \beta) = \mathcal{L}(\alpha) \cap \mathcal{L}(\beta)$$

$$\mathcal{L}(\neg \alpha) = \Sigma^* \setminus \mathcal{L}(\alpha)$$

Π -Extended Regular Expressions

$$\mathcal{L}(\emptyset) = \{\}$$

$$\mathcal{L}(\varepsilon) = \{\emptyset\}$$

$$\mathcal{L}(a) = \{[a]\} \quad a \in \Sigma$$

$$\mathcal{L}(\alpha + \beta) = \mathcal{L}(\alpha) \cup \mathcal{L}(\beta)$$

$$\mathcal{L}(\alpha \cdot \beta) = \mathcal{L}(\alpha) \cdot \mathcal{L}(\beta)$$

$$\mathcal{L}(\alpha^*) = \mathcal{L}(\alpha)^*$$

$$\mathcal{L}(\alpha \cap \beta) = \mathcal{L}(\alpha) \cap \mathcal{L}(\beta)$$

$$\mathcal{L}(\neg \alpha) = \Sigma^* \setminus \mathcal{L}(\alpha)$$

$$\mathcal{L}(\Pi \alpha) = \{ w \mid w \in \mathcal{L}(\alpha) \}$$

Π -Extended Regular Expressions

$$\mathcal{L}_n(\emptyset) = \{\}$$

$$\mathcal{L}_n(\varepsilon) = \{\emptyset\}$$

$$\mathcal{L}_n(a) = \{[a]\} \quad a \in \Sigma_n$$

$$\mathcal{L}_n(\alpha + \beta) = \mathcal{L}_n(\alpha) \cup \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\alpha \cdot \beta) = \mathcal{L}_n(\alpha) \cdot \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\alpha^*) = \mathcal{L}_n(\alpha)^*$$

$$\mathcal{L}_n(\alpha \cap \beta) = \mathcal{L}_n(\alpha) \cap \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\neg \alpha) = \Sigma_n^* \setminus \mathcal{L}_n(\alpha)$$

$$\mathcal{L}_n(\Pi \alpha) = \{ w \mid w \in \mathcal{L}_{n+1}(\alpha) \}$$

Π -Extended Regular Expressions

$$\mathcal{L}_n(\emptyset) = \{\}$$

$$\mathcal{L}_n(\varepsilon) = \{\square\}$$

$$\mathcal{L}_n(a) = \{[a]\} \quad a \in \Sigma_n$$

$$\mathcal{L}_n(\alpha + \beta) = \mathcal{L}_n(\alpha) \cup \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\alpha \cdot \beta) = \mathcal{L}_n(\alpha) \cdot \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\alpha^*) = \mathcal{L}_n(\alpha)^*$$

$$\mathcal{L}_n(\alpha \cap \beta) = \mathcal{L}_n(\alpha) \cap \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\neg \alpha) = \Sigma_n^* \setminus \mathcal{L}_n(\alpha)$$

$$\mathcal{L}_n(\Pi \alpha) = \{ w \mid w \in \mathcal{L}_{n+1}(\alpha) \}$$

Example $\Sigma_n = \{\top, \perp\}^n$ $\left[\begin{array}{c|c|c} \top & \perp & \perp \\ \perp & \top & \top \\ \perp & \perp & \top \end{array} \right] \in \Sigma_3^*$

Π -Extended Regular Expressions

$$\mathcal{L}_n(\emptyset) = \{\}$$

$$\mathcal{L}_n(\varepsilon) = \{\square\}$$

$$\mathcal{L}_n(a) = \{[a]\} \quad a \in \Sigma_n$$

$$\mathcal{L}_n(\alpha + \beta) = \mathcal{L}_n(\alpha) \cup \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\alpha \cdot \beta) = \mathcal{L}_n(\alpha) \cdot \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\alpha^*) = \mathcal{L}_n(\alpha)^*$$

$$\mathcal{L}_n(\alpha \cap \beta) = \mathcal{L}_n(\alpha) \cap \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\neg \alpha) = \Sigma_n^* \setminus \mathcal{L}_n(\alpha)$$

$$\mathcal{L}_n(\Pi \alpha) = \{ \mathbf{w} \mid \mathbf{w} \in \mathcal{L}_{n+1}(\alpha) \}$$

Example $\Sigma_n = \{\top, \perp\}^n$

$$\left[\begin{array}{c|c|c} \top & \perp & \perp \\ \perp & \top & \top \\ \perp & \perp & \top \end{array} \right] \in \Sigma_2^*$$

Π -Extended Regular Expressions

$$\mathcal{L}_n(\emptyset) = \{\}$$

$$\mathcal{L}_n(\varepsilon) = \{\square\}$$

$$\mathcal{L}_n(a) = \{[a]\} \quad a \in \Sigma_n$$

$$\mathcal{L}_n(\alpha + \beta) = \mathcal{L}_n(\alpha) \cup \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\alpha \cdot \beta) = \mathcal{L}_n(\alpha) \cdot \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\alpha^*) = \mathcal{L}_n(\alpha)^*$$

$$\mathcal{L}_n(\alpha \cap \beta) = \mathcal{L}_n(\alpha) \cap \mathcal{L}_n(\beta)$$

$$\mathcal{L}_n(\neg \alpha) = \Sigma_n^* \setminus \mathcal{L}_n(\alpha)$$

$$\mathcal{L}_n(\Pi \alpha) = \{\text{map } \pi w \mid w \in \mathcal{L}_{n+1}(\alpha)\}$$

$$\pi: \Sigma_{n+1} \rightarrow \Sigma_n$$

Example $\Sigma_n = \{\top, \perp\}^n$

$$\left[\begin{array}{c|c|c} \top & \perp & \perp \\ \perp & \top & \top \\ \perp & \perp & \top \end{array} \right] \in \Sigma_2^* \quad \begin{array}{l} \pi = \text{tail} \\ \pi^{-1}a = \{\top a, \perp a\} \end{array}$$

Derivatives of Regular Expressions

Characteristic property $\mathcal{L}_n(\mathcal{D}_a(\alpha)) = \{w \mid aw \in \mathcal{L}_n(\alpha)\}$

Derivatives of Regular Expressions

Characteristic property $\mathcal{L}_n(\mathcal{D}_a(\alpha)) = \{w \mid aw \in \mathcal{L}_n(\alpha)\}$

$$\mathcal{D}_a(\emptyset) = \emptyset$$

$$\mathcal{D}_a(\varepsilon) = \emptyset$$

$$\mathcal{D}_a(b) = \text{if } a = b \text{ then } \varepsilon \text{ else } \emptyset$$

$$\mathcal{D}_a(\alpha + \beta) = \mathcal{D}_a(\alpha) + \mathcal{D}_a(\beta)$$

$$\mathcal{D}_a(\alpha \cdot \beta) = \text{if } \varepsilon \in \mathcal{L}(\alpha) \text{ then } \mathcal{D}_a(\alpha) \cdot \beta + \mathcal{D}_a(\beta) \text{ else } \mathcal{D}_a(\alpha) \cdot \beta$$

$$\mathcal{D}_a(\alpha^*) = \mathcal{D}_a(\alpha) \cdot \alpha^*$$

$$\mathcal{D}_a(\alpha \cap \beta) = \mathcal{D}_a(\alpha) \cap \mathcal{D}_a(\beta)$$

$$\mathcal{D}_a(\neg \alpha) = \neg \mathcal{D}_a(\alpha)$$

Derivatives of Regular Expressions

Characteristic property $\mathcal{L}_n(\mathcal{D}_a(\alpha)) = \{w \mid aw \in \mathcal{L}_n(\alpha)\}$

$$\mathcal{D}_a(\emptyset) = \emptyset$$

$$\mathcal{D}_a(\varepsilon) = \emptyset$$

$$\mathcal{D}_a(b) = \text{if } a = b \text{ then } \varepsilon \text{ else } \emptyset$$

$$\mathcal{D}_a(\alpha + \beta) = \mathcal{D}_a(\alpha) + \mathcal{D}_a(\beta)$$

$$\mathcal{D}_a(\alpha \cdot \beta) = \text{if } \varepsilon \in \mathcal{L}(\alpha) \text{ then } \mathcal{D}_a(\alpha) \cdot \beta + \mathcal{D}_a(\beta) \text{ else } \mathcal{D}_a(\alpha) \cdot \beta$$

$$\mathcal{D}_a(\alpha^*) = \mathcal{D}_a(\alpha) \cdot \alpha^*$$

$$\mathcal{D}_a(\alpha \cap \beta) = \mathcal{D}_a(\alpha) \cap \mathcal{D}_a(\beta)$$

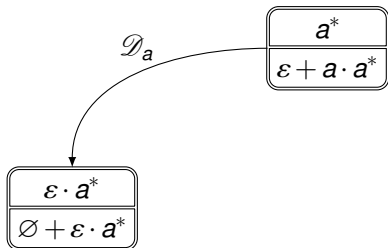
$$\mathcal{D}_a(\neg \alpha) = \neg \mathcal{D}_a(\alpha)$$

$$\mathcal{D}_a(\prod \alpha) = \prod \left(\bigoplus_{b \in \pi^{-1}a} \mathcal{D}_b(\alpha) \right)$$

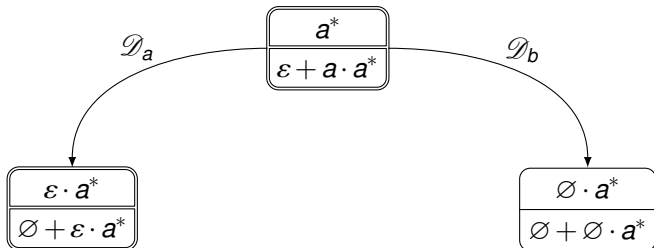
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$

a^*
$\varepsilon + a \cdot a^*$

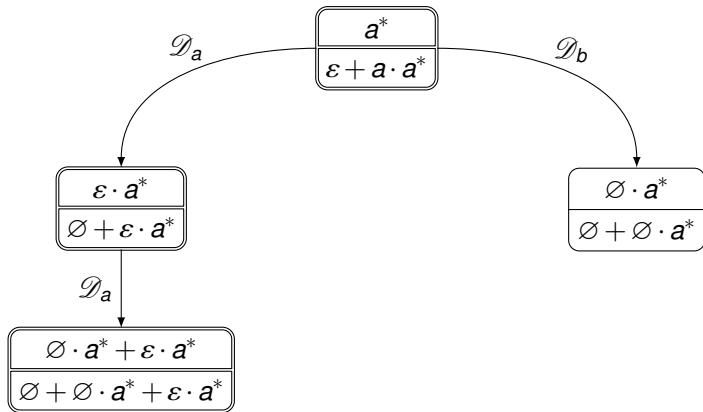
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



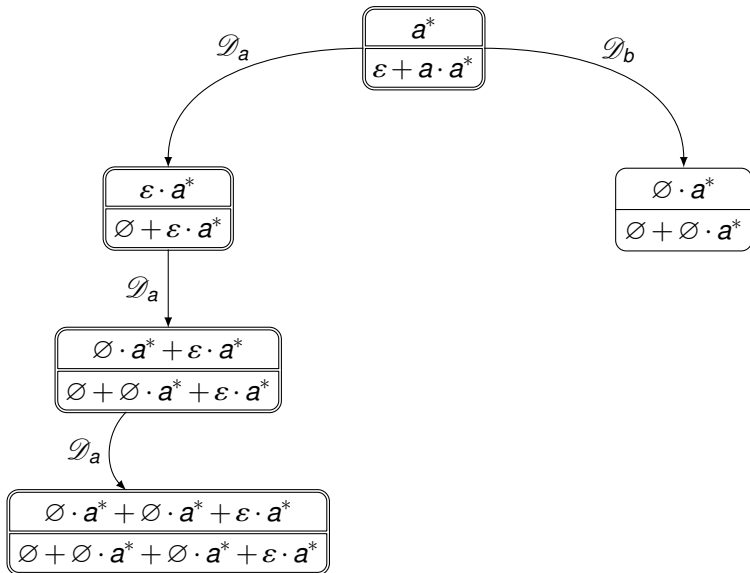
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



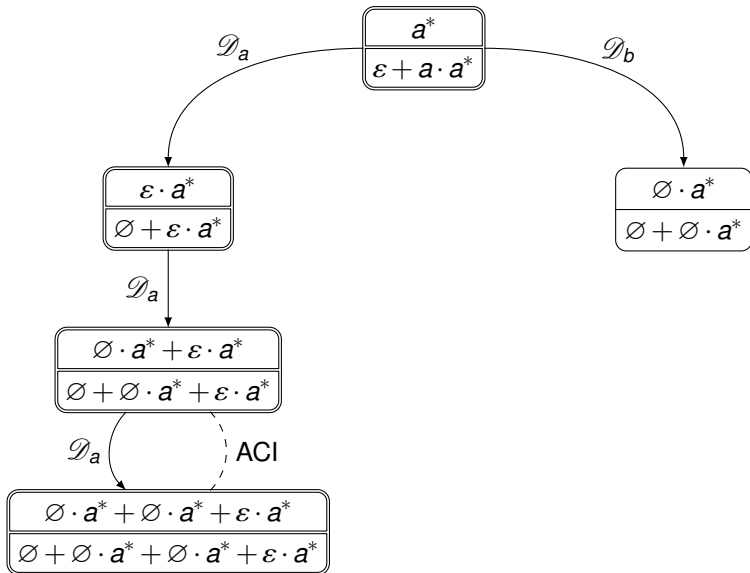
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



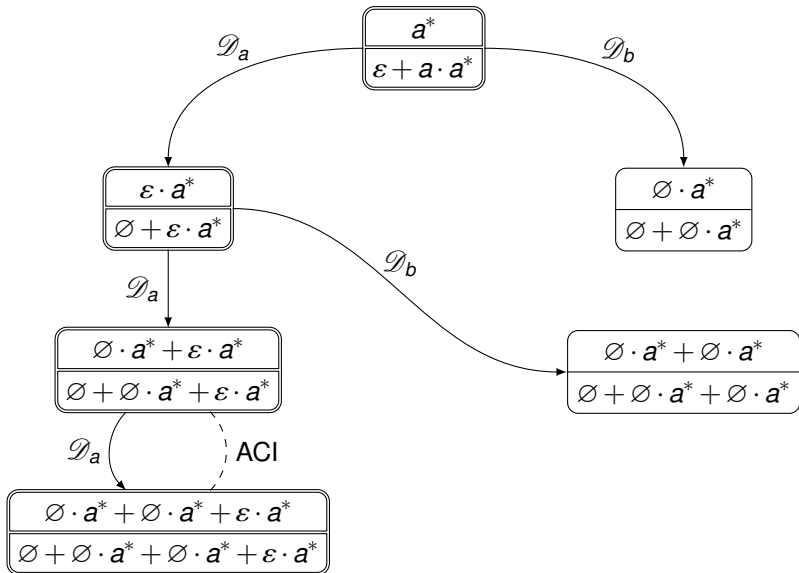
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



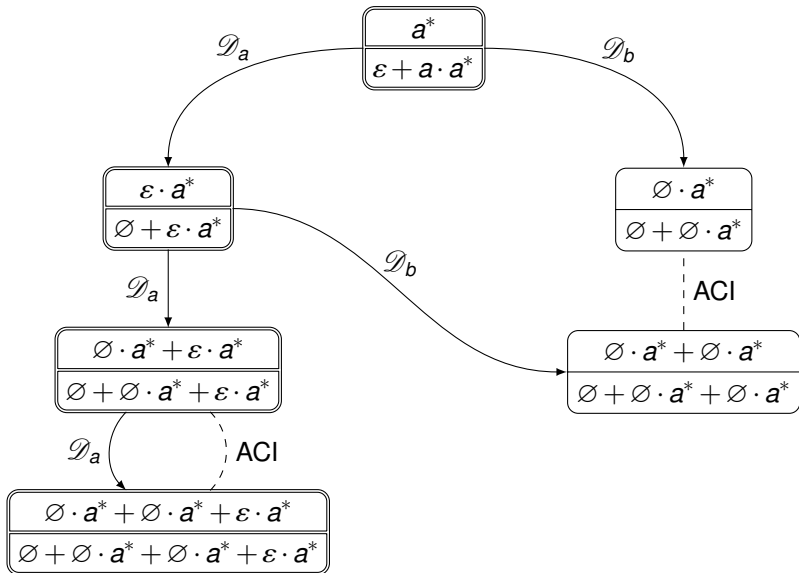
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



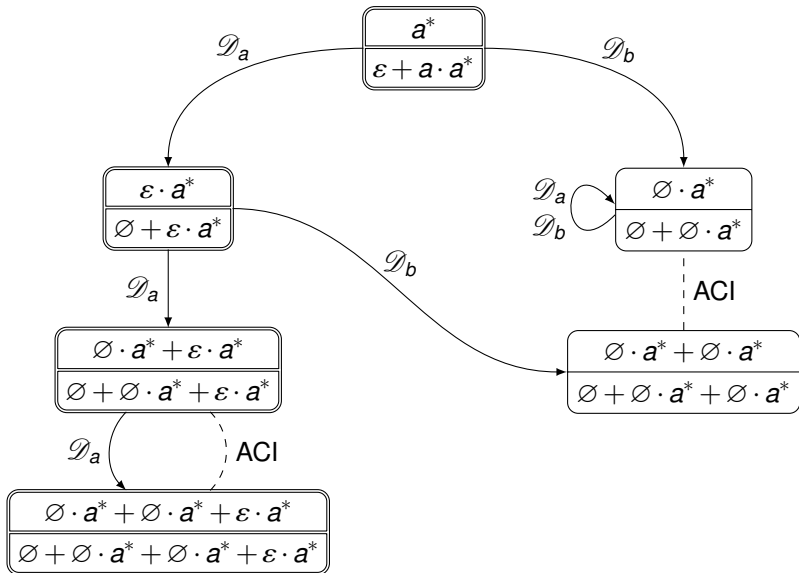
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



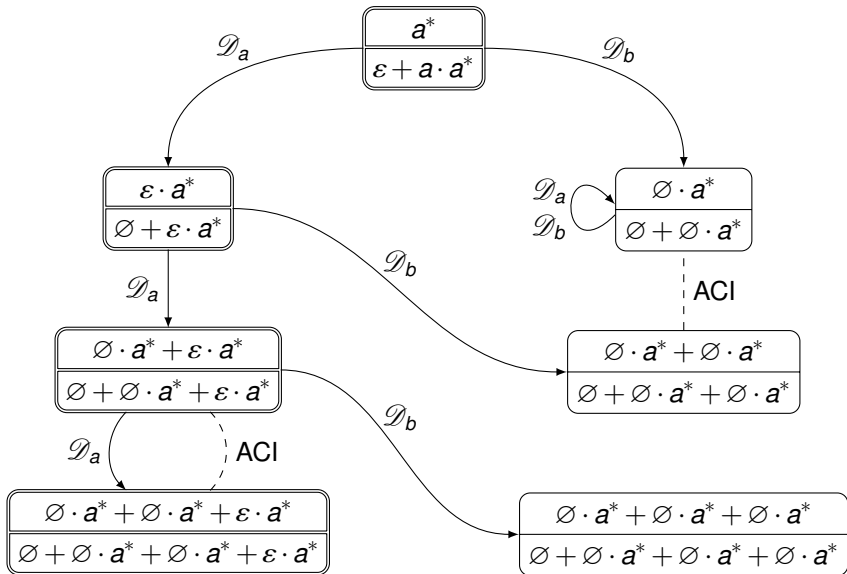
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



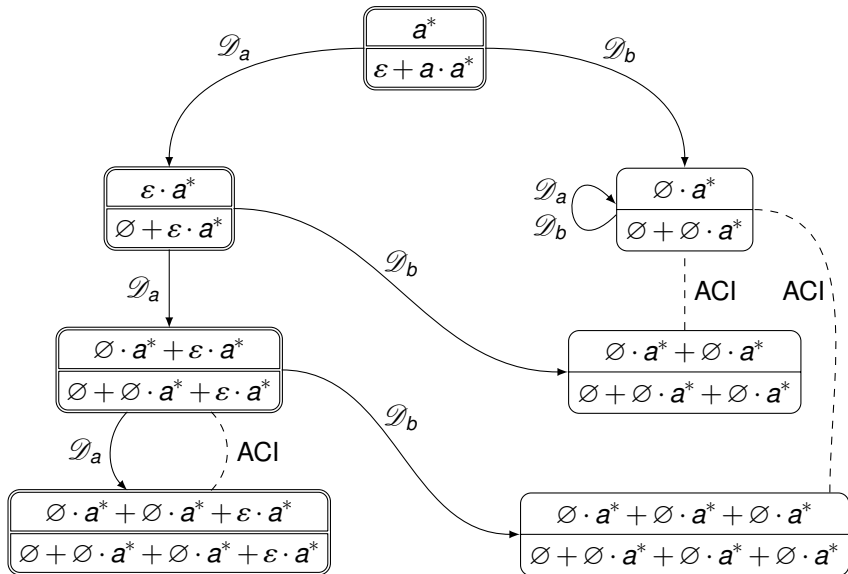
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



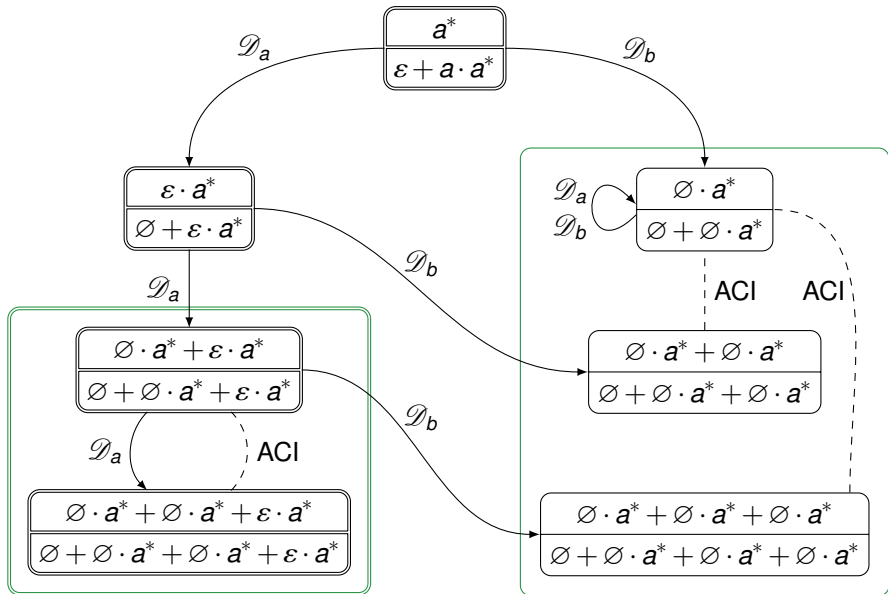
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



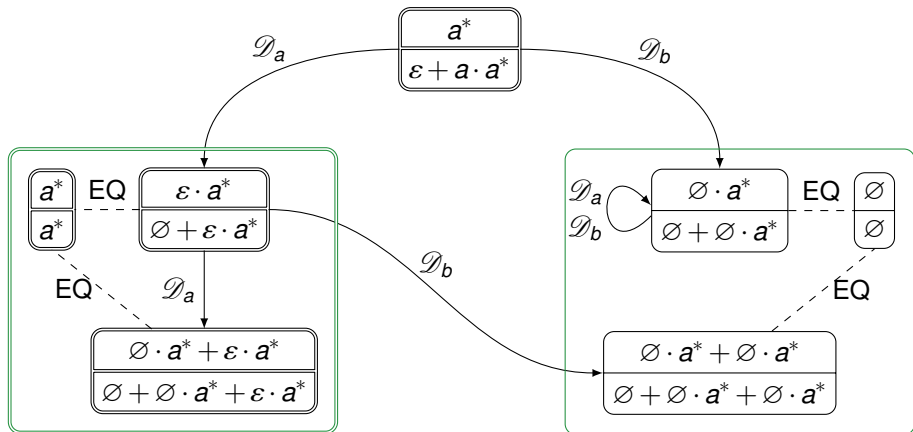
DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



DP by Example: $a^* \stackrel{?}{=} \varepsilon + a \cdot a^*$ for $\Sigma = \{a, b\}$



Related Work

- Theoretical groundwork
 - JACM 1964 Brzozowski
 - CONCUR 1998 Rutten

Related Work

- Theoretical groundwork

JACM 1964 Brzozowski

CONCUR 1998 Rutten

- FP community

JFP 2009 Owens, Reppy, and Turon

ICFP 2010 Fischer, Huch, and Wilke

ICFP 2010 Danielsson

ICFP 2011 Might, Darais, and Spiewak

Related Work

- Theoretical groundwork
 - JACM 1964 Brzozowski
 - CONCUR 1998 Rutten
- FP community
 - JFP 2009 Owens, Reppy, and Turon
 - ICFP 2010 Fischer, Huch, and Wilke
 - ICFP 2010 Danielsson
 - ICFP 2011 Might, Darais, and Spiewak
- ITP community
 - ITP 2010 Braibant and Pous
 - JAR 2011 Krauss and Nipkow
 - CPP 2011 Coquand and Siles
 - ITP 2012 Asperti
 - RAMiCS 2012 Moreira, Pereira, and de Sousa

Outline

Regular Expressions Equivalence

MSO

MSO Formulas

formula = $Q_a(x)$
| $x < y$
| $x \in X$
| \neg formula
| formula \vee formula
| formula \wedge formula
| $\exists x$. formula
| $\exists X$. formula

MSO Formulas

formula = $Q_a(x)$
| $x < y$
| $x \in X$
| \neg formula
| formula \vee formula
| formula \wedge formula
| $\exists x$. formula
| $\exists X$. formula

$(w, \mathcal{J}) \models Q_a(x) \Leftrightarrow w !! \mathcal{J} x = a$

MSO Formulas

formula = $Q_a(x)$
| $x < y$
| $x \in X$
| \neg formula
| formula \vee formula
| formula \wedge formula
| $\exists x$. formula
| $\exists X$. formula

$$(w, \mathcal{I}) \models Q_a(x) \Leftrightarrow w !! \mathcal{I} x = a$$

$$\mathcal{L}_{\text{M2L}}(\varphi) = \{\text{enc}(w, \mathcal{I}) \mid (w, \mathcal{I}) \models \varphi\}$$

Representation of Interpretations as Words

$$(w = aba, \quad \mathfrak{I} = \{x \mapsto 0, X \mapsto \{1,2\}, y \mapsto 2\})$$

Representation of Interpretations as Words

$$(w = aba, \quad \mathcal{I} = \{x \mapsto 0, X \mapsto \{1,2\}, y \mapsto 2\})$$

↓ enc

$$\begin{array}{l} x \\ X \\ y \end{array} \left[\begin{array}{c|c|c} a & b & a \\ \top & \perp & \perp \\ \perp & \top & \top \\ \perp & \perp & \top \end{array} \right]$$

$$\Sigma_n = \Sigma \times \{\top, \perp\}^n$$

Representation of Interpretations as Words

$$(w = aba, \quad \mathcal{I} = \{x \mapsto 0, X \mapsto \{1,2\}, y \mapsto 2\})$$

↓ enc

$$\begin{array}{c} x \\ X \\ y \end{array} \left[\begin{array}{c|c|c} a & b & a \\ \top & \perp & \perp \\ \perp & \top & \top \\ \perp & \perp & \top \end{array} \right]$$

$$\Sigma_n = \Sigma \times \{\top, \perp\}^n$$

$$\pi(a, bs) = (a, \text{tail } bs)$$

$$\pi^{-1}(a, bs) = \{(a, \top bs), (a, \perp bs)\}$$

From MSO Formulas to Regular Expressions

$$\text{mkRE } n(Q_a(m)) = \Sigma_n^* \cdot \begin{pmatrix} a \\ \top/\perp \\ \top \\ \top/\perp \end{pmatrix} \cdot \Sigma_n^* \cap \text{WF } n \{m\}$$

From MSO Formulas to Regular Expressions

$$\text{mkRE } n(Q_a(m)) = \Sigma_n^* \cdot \begin{pmatrix} a \\ \top/\perp \\ \top \\ \top/\perp \end{pmatrix} \cdot \Sigma_n^* \cap \text{WF } n \{m\}$$

⋮

$$\text{mkRE } n(\varphi_1 \vee \varphi_2) = (\text{mkRE } n \varphi_1 + \text{mkRE } n \varphi_2) \cap \text{WF } n(\text{FV}(\varphi_1 \vee \varphi_2))$$

From MSO Formulas to Regular Expressions

$$\text{mkRE } n (\text{Q}_a(m)) = \Sigma_n^* \cdot \begin{pmatrix} a \\ \top/\perp \\ \top \\ \top/\perp \end{pmatrix} \cdot \Sigma_n^* \cap \text{WF } n \{m\}$$

⋮

$$\text{mkRE } n (\varphi_1 \vee \varphi_2) = (\text{mkRE } n \varphi_1 + \text{mkRE } n \varphi_2) \cap \text{WF } n (\text{FV} (\varphi_1 \vee \varphi_2))$$

⋮

$$\text{mkRE } n (\exists x. \varphi) = \Pi (\text{mkRE } (n+1) \varphi)$$

$$\text{mkRE } n (\exists X. \varphi) = \Pi (\text{mkRE } (n+1) \varphi)$$

From MSO Formulas to Regular Expressions

$$\text{mkRE } n(Q_a(m)) = \Sigma_n^* \cdot \begin{pmatrix} a \\ \top/\perp \\ \top \\ \top/\perp \end{pmatrix} \cdot \Sigma_n^*$$

⋮

$$\text{mkRE } n(\varphi_1 \vee \varphi_2) = \text{mkRE } n \varphi_1 + \text{mkRE } n \varphi_2$$

⋮

$$\text{mkRE } n(\exists x. \varphi) = \Pi(\text{mkRE } (n+1) \varphi \cap \text{WF } (n+1) \{x\})$$

$$\text{mkRE } n(\exists X. \varphi) = \Pi(\text{mkRE } (n+1) \varphi)$$

From MSO Formulas to Regular Expressions

$$\text{mkRE } n(Q_a(m)) = \Sigma_n^* \cdot \begin{pmatrix} a \\ \top/\perp \\ \top \\ \top/\perp \end{pmatrix} \cdot \Sigma_n^*$$

⋮

$$\text{mkRE } n(\varphi_1 \vee \varphi_2) = \text{mkRE } n \varphi_1 + \text{mkRE } n \varphi_2$$

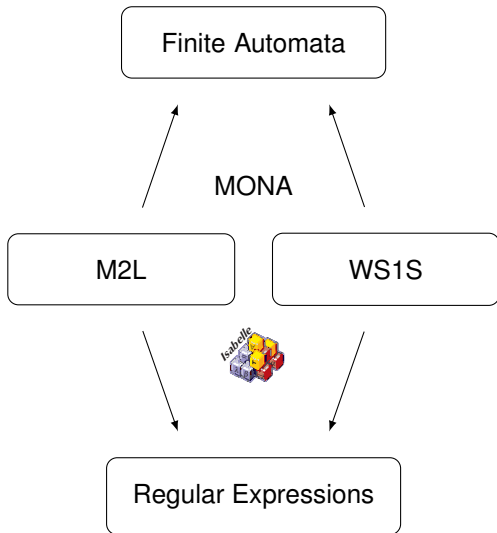
⋮

$$\text{mkRE } n(\exists x. \varphi) = \Pi (\text{mkRE } (n+1) \varphi \cap \text{WF } (n+1) \{x\})$$

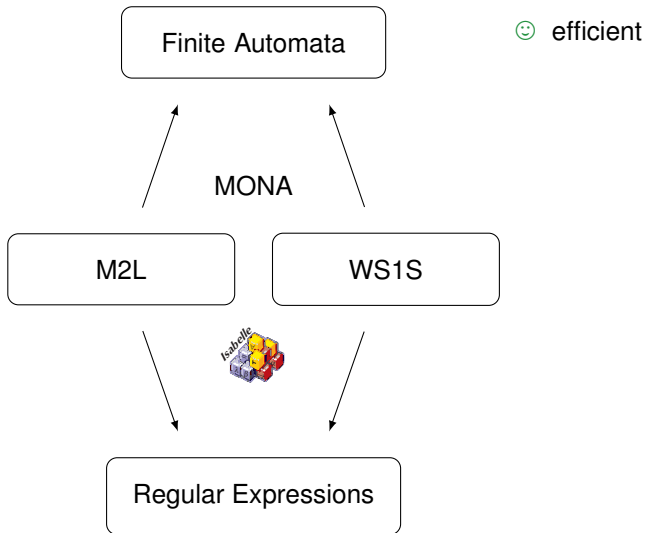
$$\text{mkRE } n(\exists X. \varphi) = \Pi (\text{mkRE } (n+1) \varphi)$$

Theorem $\mathcal{L}_{\text{M2L}}(\varphi) = \mathcal{L}_n(\text{mkRE } n \varphi \cap \text{WF } n(\text{FV } \varphi)) - \{\varepsilon\}$

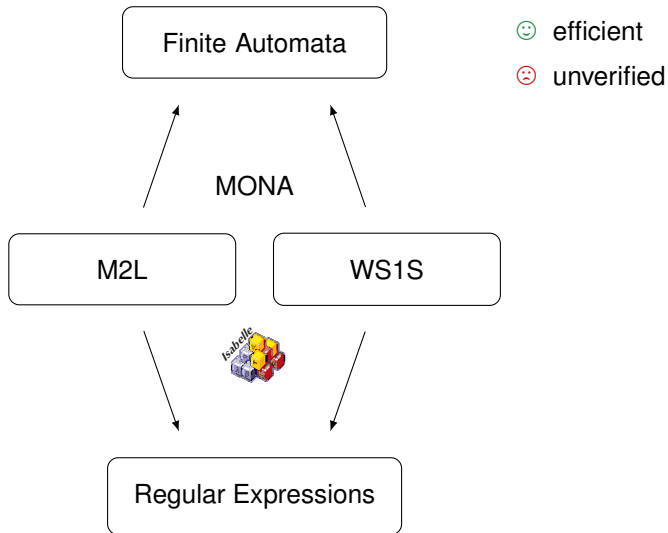
Head to Head



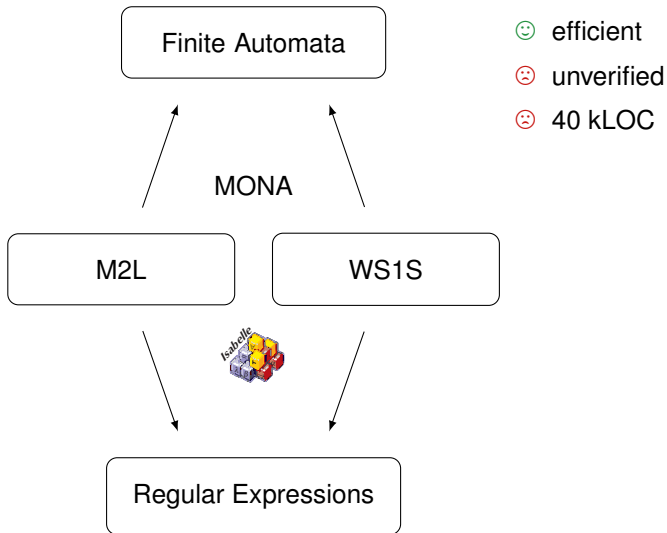
Head to Head



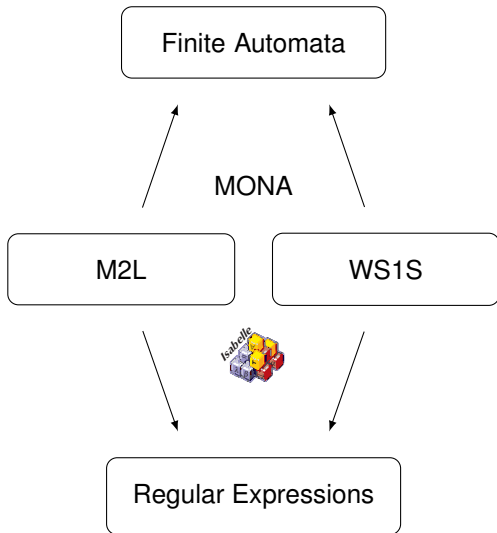
Head to Head



Head to Head

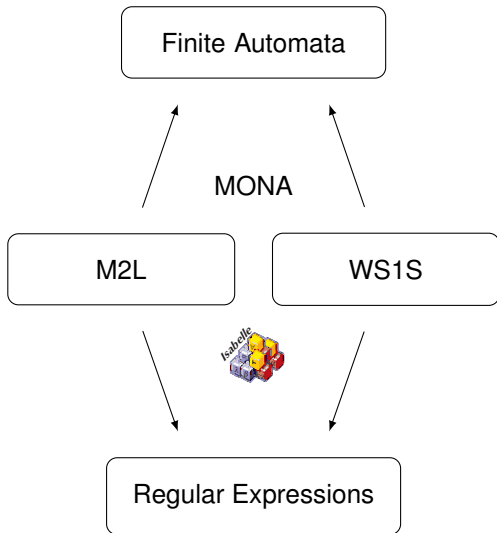


Head to Head



- 😊 efficient
- ☹️ unverified
- ☹️ 40 kLOC
- ☹️ 40 kLOC of C/C++

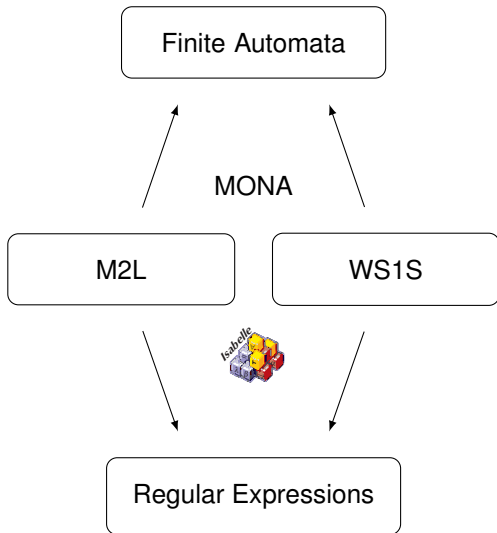
Head to Head



- 😊 efficient
- ☹️ unverified
- ☹️ 40 kLOC
- ☹️ 40 kLOC of C/C++

☹️ inefficient

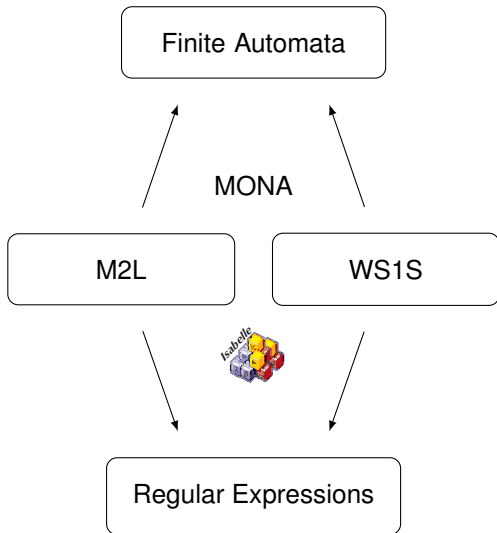
Head to Head



- 😊 efficient
- ☹️ unverified
- ☹️ 40 kLOC
- ☹️ 40 kLOC of C/C++

- ☹️ inefficient
- 😊 simple: 350 LOC of Isabelle/HOL

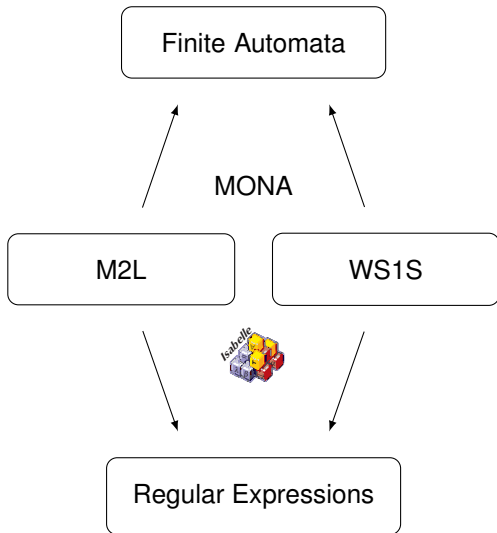
Head to Head



- 😊 efficient
- ☹️ unverified
- ☹️ 40 kLOC
- ☹️ 40 kLOC of C/C++

- ☹️ inefficient
- 😊 simple: 350 LOC of Isabelle/HOL
- 😊 functional: 2,5 kLOC of generated Haskell

Head to Head



- 😊 efficient
- ☹️ unverified
- ☹️ 40 kLOC
- ☹️ 40 kLOC of C/C++

- ☹️ inefficient
- 😊 simple: 350 LOC of Isabelle/HOL
- 😊 functional: 2,5 kLOC of generated Haskell
- 😊 sound, complete, and terminating: 5 kLQP

Head to Head

Finite Automata

😊 efficient

☹️ unverified

☹️ 40 kLOC

100 kLOC of C/C++

Thanks for listening!

pre: 350 LOC of
Isabelle/HOL

Regular Expressions

😊 functional: 2,5 kLOC of
generated Haskell

😊 sound, complete, and
terminating: 5 kLQP

Verified Decision Procedures

for

Monadic Second-Order Logic on Strings

Functional Pearl

Dmitriy Traytel Tobias Nipkow



Technische Universität München

