The many faces of a translation

Pierre McKenzie $\boxed{\text{Montréal}}$
Thomas Schwentick $\boxed{\text{Mainz}}$
Denis Thérien $\boxed{\text{McGill}}$
Heribert Vollmer $\boxed{\text{Würzburg}}$

$$\exists x \; \underbrace{\forall y \; [(y \leq x) \wedge P_b\, x]}_{\varphi(x)}$$

Alphabet $\Sigma = \{a, b, c\}$

x, y positions within $w_1 w_2 \ldots w_n \in \Sigma^*$

$P_b\, x$ means: letter at position x is b

$\exists x \varphi(x)$ defines a language over Σ

FO languages:

- Variables: $x, y, x, x_5, \ldots$
- Logical symbols: $\wedge, \vee, \neg, (,), \ldots$
- Predicates: $<, =, \leq, P_a, P_b, P_c$
- Quantifiers: $\exists, \forall$

Equivalently [Sc65,McPa71]:

- accepted by DFAs such that
  $\exists n\ \forall w \in \Sigma^* \quad \delta_{w^n}(\cdot) = \delta_{w^{n+1}}(\cdot)$
- star-free regular expressions
- recognized by aperiodic monoids.

FO languages $\subset$ regular languages $=$ DFAs $=$ NFAs $=$ 2wayDFAs $=$ finite monoids $=$ monadic 2nd order $=$ regular expressions...

FO with one free variable:

$$\phi(x) : \Sigma^* \to \{0,1\}^*$$
$$w_1 \ldots w_i \ldots w_n \mapsto d_1 \ldots d_i \ldots d_n$$

where $d_i = \begin{cases} 1 & \text{if } \phi(x) \text{ true when } x = i \\ 0 & \text{if } \phi(x) \text{ false when } x = i \end{cases}$

$\phi(x)$ from earlier example defines

$$\{a,b,c\}^* \to \{0,1\}^*$$
$$w_1 \ldots w_n \mapsto \begin{cases} 00\ldots01 & \text{if } w_n = b \\ 00\ldots00 & \text{if } w_n \in \{a,c\} \end{cases}$$

FO formulas with one free variable define FO-translations.

FO-translations play a role in:

Logic and descriptive complexity

- as a special case of first-order interpretations

Complexity theory

- as a tight reducibility, further refined to quantifier-free projections
- allows "fine-tuning" the uniformity in low level reductions

Automata theory

- as will be seen here.

Aperiodic DFAs capture FO languages. Are FO-translations therefore captured by aperiodic deterministic finite transducers?

NO!!!

$\Psi(z) = \exists x \phi(x)$ with
our old $\phi(x)$ provides counter-example:

$$\{a,b,c\}^* \to \{0,1\}^*$$
$$w_1 \dots w_n \mapsto \begin{cases} 11\dots11 & \text{if } w_n = b \\ 00\dots00 & \text{if } w_n \in \{a,c\} \end{cases}$$

[LMSV98]: FO-translations are indeed captured by finite transducers, but these need to be aperiodic nondeterministic and single-valued.

Huh???

aperiodic nondeterministic:

the subset construction yields an ordinary aperiodic DFA

single-valued:

on any w, a unique consistent path exists from an initial state (there may be many) to a final state.

The new results

$\mathcal{V}$-translations
and
$\mathcal{V}$-SNFT

Eilenberg's $\mathcal{V}$-bimachines

$\mathcal{V}$-2way

Main theorem: When $\mathcal{V}$ is the variety of aperiodic monoids or when $\mathcal{V}$ is a group variety, these 4 models compute the same functions $\Sigma^* \to \Delta^*$.

Significance of the new results

World is nice after all...

- four animals, one cage

$\mathcal{V}$-translations a fundamental notion

- link with Eilenberg's $\mathcal{V}$-bimachines adds relevance

$\mathcal{V}$-2wayDFTs

- pave the way for an algebraic treatment of 2-way DFAs
- (2-way DFAs are able to describe "real-world" complexity classes)

What are $\mathcal{V}$-translations?

Allow generalized quantifiers Q_L, as eg in [Li66, STT88, BIS90, St94].

L = language accepted by a $\mathcal{V}$-DFA
$\psi(y)$ = quantifier-free formula

$Q_L y\, \psi(y)$ is true on $w \in \Sigma^*$
iff
$\psi(y)$ maps w to a word in L.

DEFINITION: $\mathcal{V}$-translations are those expressible as boolean combination of $\mathcal{V}$-formulas having a free variable x.

Ex: old $\phi(x)$ was an **AND**-translation.

$\mathcal{V}$-single-valued nondet transducers

clear generalization of aperiodic SNFTs

Eilenberg's $\mathcal{V}$-bimachines

- a $\mathcal{V}$-DFA A_1 reads w from left
- a $\mathcal{V}$-DFA A_2 reads w from right
- the ith output depends on

$$(A_1 \text{ on } w_1 w_2 \ldots w_{i-1},$$
$$w_i,$$
$$A_2 \text{ on } w_n w_{n-1} \ldots w_{i+1})$$

$\mathcal{V}$-2wayDFTs? The difficulties:

- no theory of 2-way DFAs
- what state transformations?
- head motion part of the algebra?
- what should the output be?

$\mathcal{V}$-2way-deterministic transducers

Set of states $S = \mathcal{L} \uplus \mathcal{R}$

- $\mathcal{L}$ set of states entered from left
- $\mathcal{R}$ set of states entered from right

Initial configuration

$$\begin{array}{l} \triangleright w_1 w_2 \ldots w_n \triangleleft \\ \quad\uparrow \\ s_0 \in \mathcal{L} \end{array}$$

Transition function

$\delta : (S \times \Sigma) \cup (\mathcal{L} \times \{\triangleleft\}) \cup (\mathcal{R} \times \{\triangleright\}) \longrightarrow S$

Output function

- $\lambda : S \times \Sigma \to$ output alphabet
- ith output symbol produced upon visiting w_i for the last time

$\mathcal{V}$-2wayDFTs: continued

And the $\mathcal{V}$-condition?

The monoid $\langle\{\delta_w : w \in \Sigma^*\}\rangle$ must belong to the variety $\mathcal{V}$, where

$$\delta_w : \mathcal{L} \uplus \mathcal{R} \to \mathcal{L} \uplus \mathcal{R}$$

and $\delta_w(s) =$

<u>case $s \in \mathcal{L}$</u>: the state reached from

$$\begin{array}{l} w_1 w_2 \ldots w_n \\ \uparrow \\ s \end{array}$$

upon exiting $w_1 w_2 \ldots w_n$

<u>case $s \in \mathcal{R}$</u>: the state reached from

$$\begin{array}{r} w_1 w_2 \ldots w_n \\ \uparrow \\ s \end{array}$$

upon exiting $w_1 w_2 \ldots w_n$

Sample proof ideas

How can a 2-way automaton A simulate Eilenberg's bimachine (A_1, A_2)?

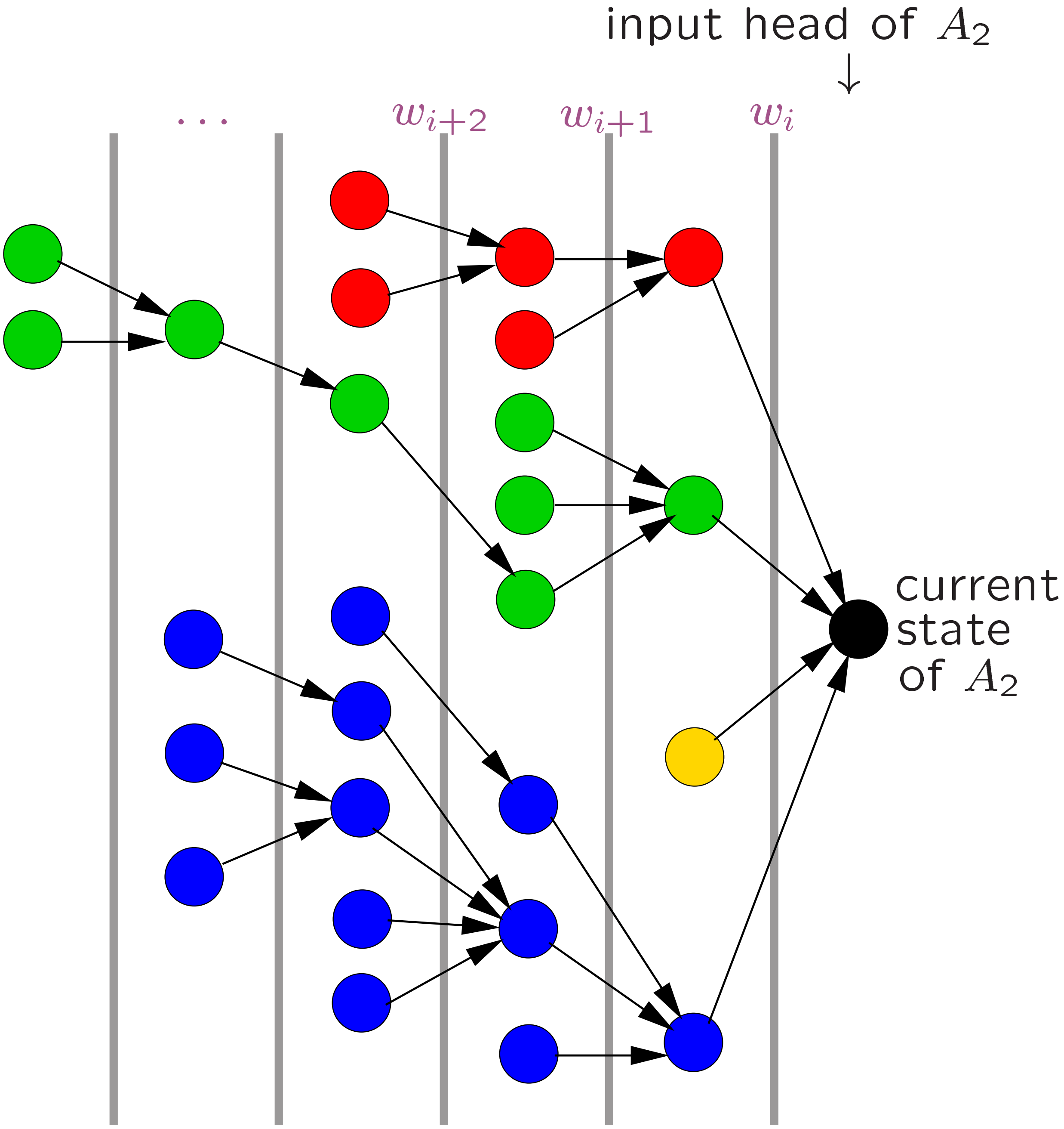
The problem:

- A_1 reads $w_1 w_2 \dots w_{i-1}$
- A_2 reads $w_n w_{n-1} \dots w_i$
- A sees $w_1 w_2 \dots w_n$: how can A remember i when sweeping back and forth to simulate A_2?

The solution [HU67]:

Given the state of A_2 on $w_n w_{n-1} \dots w_i$, A can find state of A_2 on $w_n w_{n-1} \dots w_{i+1}$

how A reverses A_2: continued



The (messy) magic

The backing-up of A_2 can be implemented in such a way as to preserve

- groups (no surprise)
- pure aperiodic behavior
- (seemingly) many other varieties.

Converse:
$\mathcal{V}$-2wayDFTs by $\mathcal{V}$-bimachines

Direct. The states of A_1 and A_2 are the maps $\langle\{\delta_w : w \in \Sigma^*\}\rangle$ of A.

Glimpse: formulas

$\mathcal{V}$-translation to $\mathcal{V}$-bimachine

Formula $Q_L y \psi(x, y)$

For each possible w_i consider three formulas $\psi^{<}_{w_i}(y)$, $\psi^{=}_{w_i}(y)$ and $\psi^{>}_{w_i}(y)$ incorporating $x < y$, $x = y$ and $x > y$.

Output at $x = i$ depends on

$\psi^{<}_{w_i}(y)$ on $w_1 w_2 \ldots w_{i-1}$
$\psi^{=}_{w_i}(y)$ on w_i
$\psi^{>}_{w_i}(y)$ on $w_{i+1} w_{i+2} \ldots w_n$

processed by two automata from $\mathcal{V}$, one for the prefix and one for the suffix.

Needs closure of $\mathcal{V}$ under reversal.

Conclusion

Order in the kingdom

Theory of two-way DFAs

- Is this the right 2-way DFA?
- If so push it.
- If not fix it.

Proceedings version misprinted

- All 79 "=" signs disappeared! Find them all on the Wuerzburg Web page of Heribert Vollmer.