What we are:

Logic WMSO over finite words

$w = a_0 \dots a_{n-1}$

Important: Positions $\{0, \dots, n-1\}$

Predicates:

- $P_a(x)$ "Position x carries letter a"
- $x < y$ "Position x is before position y"

Second-order variables:

- X "Set of positions (finite set)"
- $X(x)$ "position x is in the set X"

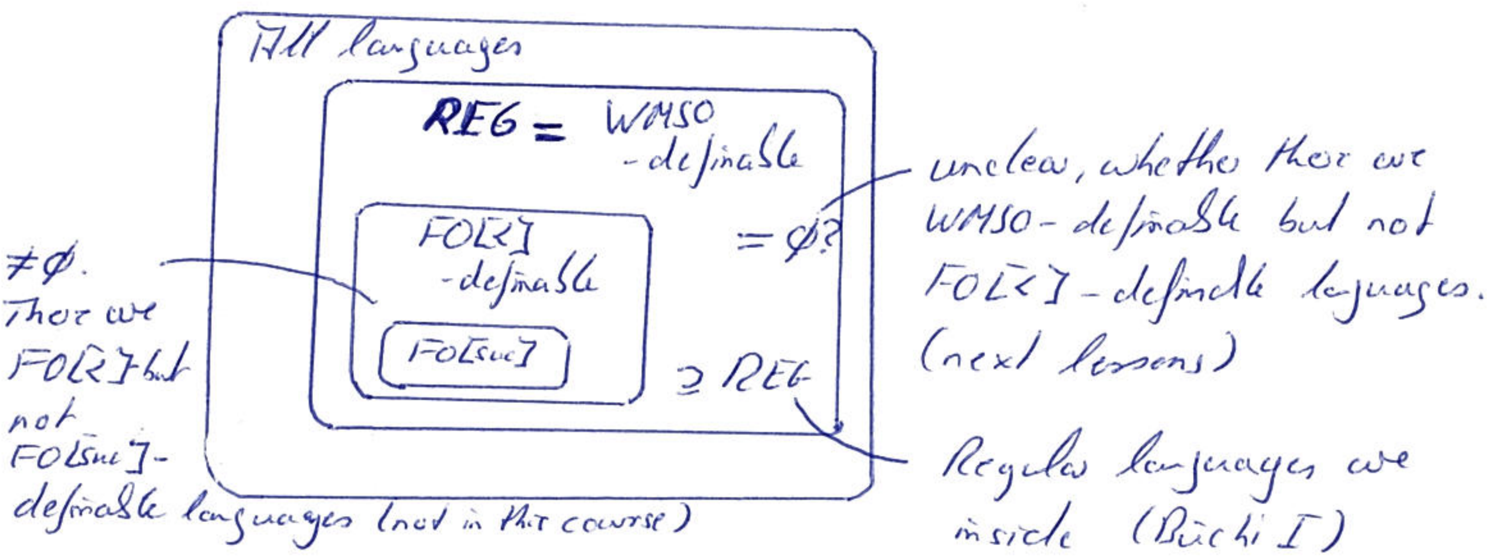
WMSO-formulas define languages:

$\exists x : \exists y : P_a(x) \wedge P_b(y) \wedge x < y$ over $\Sigma = \{a, b\}$

defines

$\{a,b\}^* . a . \{a,b\}^* . b . \{a,b\}^*$.

We discussed



This lecture: = REG

2.2 Büchi's theorem

WMSO-definability = regularity

2.2.1 From automata to logic

Theorem (Büchi I):

For every regular language L, we can effectively construct a WMSO-formula φ_L with $L = L(\varphi_L)$.

Construction:

Let $L = L(A)$ with $A = (Q, q_0, \rightarrow, Q_F)$

where $Q = \{q_0, \ldots, q_n\}$

Then we define

$$\varphi_{L(A)} := \exists X_0 : \ldots \exists X_n : (1) \wedge (2) \wedge (3) \wedge (4)$$

$$(\wedge 5 \text{ if } \varepsilon \notin L(A))$$

with

(1) $\bigwedge_{0 \leq i < j \leq n} \forall x : \neg (X_i(x) \wedge X_j(x))$

(2) $\forall x : first(x) \rightarrow X_0(x)$

(3) $\forall x : \forall y : succ(x,y) \rightarrow \bigvee_{q_i \xrightarrow{a} q_j} (X_i(x) \wedge P_a(x) \wedge X_j(y))$

(4) $\forall x : last(x) \rightarrow \bigvee_{q_i \xrightarrow{a} q_j \in Q_F} (X_i(x) \wedge P_a(x))$

(5) $\exists x : x = x$

Intuitively:

(1) Every letter stems from a single state (no branching in the word)

(2) Run starts in q_0

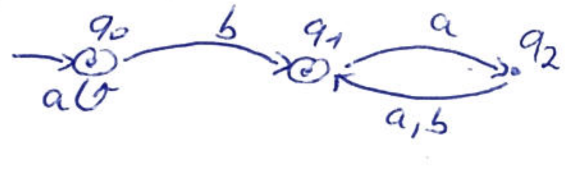
(3) Successor state respects transition relation

(4) Last letter leads to a final state

(5) There is a letter.

Example:

A



Then we have

$$\varphi_{L(A)} = \exists X_0 : \exists X_1 : \exists X_2 :$$

$$
\begin{aligned}
(1) \quad & \forall x : \neg (X_0(x) \wedge X_1(x)) \\
& \wedge \forall x : \neg (X_0(x) \wedge X_2(x)) \\
& \wedge \forall x : \neg (X_1(x) \wedge X_2(x)) \\
\wedge & \\
(2) \quad & \forall x : first(x) \to X_0(x) \\
\wedge & \\
(3) \quad & \forall x : \forall y : suc(x,y) \to (3a) \vee (3b) \vee (3c) \vee (3d) \vee (3e)
\end{aligned}
$$

with

$$
\begin{aligned}
(3a) \quad & X_0(x) \wedge P_a(x) \wedge X_0(y) \\
(3b) \quad & X_0(x) \wedge P_b(x) \wedge X_1(y) \\
(3c) \quad & X_1(x) \wedge P_c(x) \wedge X_2(y) \\
(3d) \quad & X_2(x) \wedge P_a(x) \wedge X_1(y) \\
(3e) \quad & X_2(x) \wedge P_b(x) \wedge X_1(y)
\end{aligned}
$$

$$\wedge \; (4) \quad \forall x : last(x) \to (4a) \vee (4b) \vee (4c) \vee (4d)$$

with

$$
\begin{aligned}
(4a) \quad & X_0(x) \wedge P_a(x) \\
(4b) \quad & X_0(x) \wedge P_b(x) \\
(4c) \quad & X_2(x) \wedge P_a(x) \\
(4d) \quad & X_2(x) \wedge P_b(x).
\end{aligned}
$$

Consider run

$$q_0 \xrightarrow{b} q_1 \xrightarrow{a} q_2 \xrightarrow{a} q_1 \xrightarrow{a} q_2 \xrightarrow{b} q_1 \qquad (\text{run on } b.a.a.a.b)$$

$$0 \qquad 1 \qquad 2 \qquad 3 \qquad 4 \qquad (\text{positions}).$$

Let $X_0 = \{0\}$

$X_1 = \{1, 3\}$ (X_i contains positions of letters

$X_2 = \{2, 4\}$. that stem from q_i)

The sets satisfy (1)–(4) above:

(1) as they are disjoint

(2) as $0 \in X_0$.

(3) as an example, take $succ(0,1)$.

Choose transition (3b). We have

$X_0(0)$ and $P_b(0)$ and $X_1(1)$.

(4) let $x = 4$. Then (4d) holds:

$X_2(4) \wedge P_b(4)$ (and $q_2 \xrightarrow{b} q_1$ with $q_1 \in Q_F$).

<u>Proof sketch (for correctness of construction):</u>

Show $L(\varphi_{L(A)}) = L(A)$.

$S_w \models \varphi_{L(A)}$

$\Leftrightarrow$ there are $X_0 \ldots X_n \subseteq D_w$ so that

(1) to (4) (and potentially (5)) hold.

$\overset{?}{\Leftrightarrow}$ there is an accepting run of A on w.

$\Leftrightarrow$ $w \in L(A)$.

For ?:

"$\Leftarrow$" Take $X_0 \ldots X_n$ as in the example.

Check that this interpretation satisfies (1) to (4) (and (5)).

"$\Rightarrow$" Take single initial state from (2).

Find successors with (3).

They are unique with (3).

Accept by (4). $\square$

<u>2.2.2 From formulas to automata</u>

<u>Goal:</u> Represent models of WMSO-formulas by NFA.

<u>Approach:</u> Proceed by induction on structure of φ.

<u>Problem:</u> $\exists X: \varphi(X)$ closed but

$\varphi(X)$ itself contains X free.

Idea: · Let V subset of (first- and second-order) variables (those free in the formula of interest)

· Encode interpretations $I : V \to D_w \cup \mathcal{P}(D_w)$ of free variables into suitable alphabet Σ_V

· Assign (inductively) to every formula $\varphi(X)$ (with X free) an NFA A_φ over Σ_V so that

$$S_w, I \models \varphi \quad \text{iff.} \quad w_I \in \mathcal{L}(A_\varphi).$$

Theorem (Büchi II):

For every WMSO-sentence φ, we can effectively construct an NFA A_φ with $L(A_\varphi) = L(\varphi)$.

Key technique: alphabet extension

Let $\varphi(V)$ a formula with free variables V

· Alphabet Σ_V assigns to each variable $x, X \in V$ a truth value

$$\Sigma_V := \Sigma \times \{0,1\}^V$$

Identify S_w, I with $w_I \in \Sigma_V^*$ where

· $$\underbrace{(\underbrace{w_I(h)}_{\substack{h\text{-th position}\\ \text{in } w_I}})(x)}_{\text{entry for } x \text{ (in function in } \{0,1\}^V)} := \begin{cases} 1 & \text{if } I(x) = h \\ 0 & \text{otherwise} \end{cases}$$

· $$(w_I(h))(X) := \begin{cases} 1 & \text{if } I(X) \ni h \\ 0 & \text{otherwise} \end{cases}$$

Example:

$S_{ab.aab}, I$ with

$I(X) :=$ even positions

$I(Y) :=$ positions that are prime.

Translates into

$$w_I = \begin{array}{ccccc|l} \begin{pmatrix} a \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix} & \begin{pmatrix} a \\ 1 \\ 1 \end{pmatrix} & \begin{pmatrix} a \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} b \\ 1 \\ 0 \end{pmatrix} & \begin{matrix} \Sigma \\ I(X) \\ I(Y) \end{matrix} \\ \hline 0 & 1 & 2 & 3 & 4 & \text{positions} \end{array}$$

<u>Note</u>: For first order variables x, we have <u>at most one</u> extended letter with 1 at x.

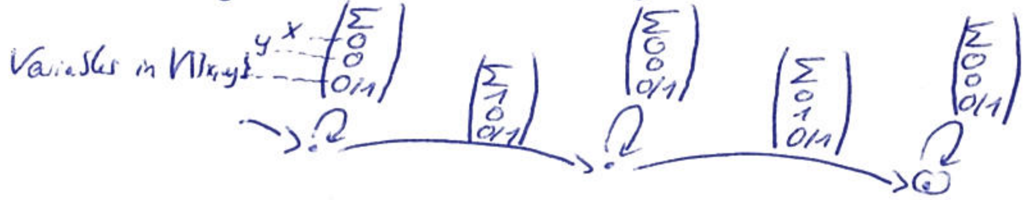
<u>Construction</u>:

- Let V be so that it contains the free variables of φ
- Inductively construct $\mathcal{A}_\varphi$ over $\Sigma_V = \Sigma \times \{0,1\}^V$ so that

$$w_I \in \mathcal{L}(\mathcal{A}_\varphi) \quad \text{iff} \quad (w, I) \models \varphi.$$

<u>Base case</u>:

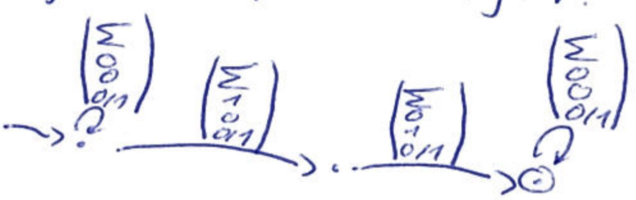
- $\mathcal{A}_{x<y}$ over Σ_V with $x, y \in V$:



↳ Here, Σ indicates set of transitions, one for each letter in Σ.

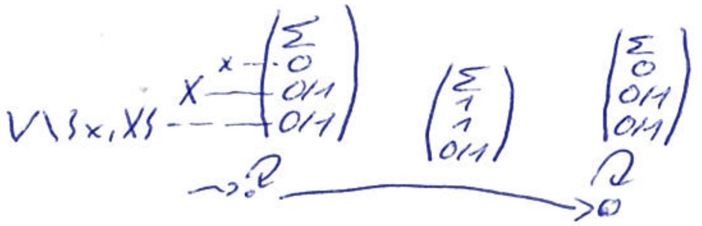
↳ Last entry is function in $\{0,1\}^{V \setminus \{x,y\}}$ that assigns an arbitrary value to each variable in $V \setminus \{x,y\}$. (indicated by 0/1)

- $\mathcal{A}_{succ(x,y)}$ over Σ_V with $x, y \in V$:



- $\mathcal{A}_{P_a(x)}$ over Σ_V with $x \in V$:

- $\mathcal{A}_{X(x)}$ over Σ_V with $x, X \in V$



For the induction step, assume we already constructed $\mathcal{A}_\varphi$ and $\mathcal{A}_\psi$ over V' for φ and ψ. So

$$w_I \in \mathcal{L}(\mathcal{A}_{\varphi\,(\psi)}) \quad \text{iff} \quad S_w, I \models \varphi_{(\psi)}$$

We construct automata for $\varphi \vee \psi$, $\neg\varphi$, $\exists X: \varphi$, $\exists x: \varphi$ over V (so that V contains the free variables).

<u>Case $\varphi \vee \psi$:</u> Take $\underbrace{\mathcal{A}_\varphi \cup \mathcal{A}_\psi}$ over Σ_V

Automaton for union of languages.

<u>Case $\neg\varphi$:</u> Take $\underbrace{\overline{\mathcal{A}_\varphi}}$ over Σ_V

Automaton for complement of language $\mathcal{A}_\varphi$

↳ Determinize $\mathcal{A}_\varphi$

↳ Invert final states.

<u>Case $\exists X: \varphi$</u>

<u>Intuition:</u> Guess content of X nondeterministically

<u>Technically:</u>

- Compute $\mathcal{A}_\varphi$ over $\Sigma_{V'}$ with $V' = V \cup \{X\}$
- <u>Project away</u> component X of the extended letters

↳ This yields $\mathcal{A}_{\exists X:\varphi}$ <u>over Σ_V</u>

Let $\mathcal{A}_\varphi$ over $V \cup \{X\}$ be $(Q, q_0, \rightarrow, Q_F)$.

Define

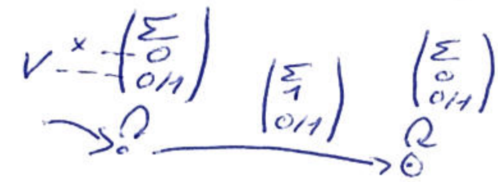
$$\mathcal{A}_{\exists X:\varphi} = (Q, q_0, \rightarrow', Q_F) \text{ with}$$

$$q \xrightarrow{a'}{}' q' \text{ if } q \xrightarrow{a} q' \text{ and } a' = a|_V.$$

Case $\exists x: \varphi$: · Compute A_φ over $\Sigma_{V'}$ with $V' = V \cup \{x\}$

· Intersect $\mathcal{L}(A_\varphi)$ with language of

$A_{\exists x}$



"Ensures x is guessed precisely once"

· Afterwards project away x to get $A_{\exists x:\varphi}$ over Σ_V.

– Note that

$A_{\exists X:\varphi}$ and $A_{\exists x:\varphi}$ may be nondeterministic, although A_φ was deterministic.

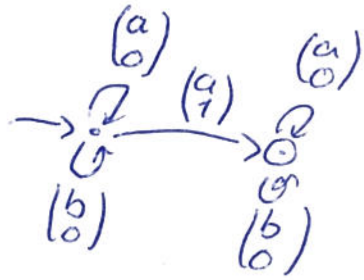
Example:

Let $\Sigma = \{a, b\}$.

Consider $\varphi = \exists x: P_a(x)$ (defines $\{a,b\}^* . a . \{a,b\}^*$)

Then

$A_{P_a(x)}$ over $\Sigma_{\{x\}}$:



After projection:

$A_{\exists x:P_a(x)}$ over $\Sigma_\emptyset$: