



Automata and Logic

Aart Middeldorp and Samuel Frontull

Definitions

► **nondeterministic finite automaton (NFA)** is quintuple $N = (Q, \Sigma, \Delta, S, F)$ with

- ① Q : finite set of states
- ② Σ : input alphabet
- ③ $\Delta: Q \times \Sigma \rightarrow 2^Q$: transition function
- ④ $S \subseteq Q$: set of start states
- ⑤ $F \subseteq Q$: final (accept) states

► $\hat{\Delta}: 2^Q \times \Sigma^* \rightarrow 2^Q$ is inductively defined by

$$\hat{\Delta}(A, \epsilon) = A$$

$$\hat{\Delta}(A, xa) = \bigcup_{q \in \hat{\Delta}(A, x)} \Delta(q, a)$$

► $x \in \Sigma^*$ is accepted by N if $\hat{\Delta}(S, x) \cap F \neq \emptyset$

Theorem

every set accepted by NFA is regular

Outline

1. Summary of Previous Lecture
2. Regular Expressions
3. Intermezzo
4. Homomorphisms
5. Decision Problems
6. Further Reading

Definitions

► **NFA with ϵ -transitions (NFA_ϵ)** is sextuple $N = (Q, \Sigma, \epsilon, \Delta, S, F)$ such that

- ① $\epsilon \notin \Sigma$
- ② $M_\epsilon = (Q, \Sigma \cup \{\epsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\epsilon\}$

► **ϵ -closure** of set $A \subseteq Q$ is defined as $C_\epsilon(A) = \bigcup \{\hat{\Delta}_{N_\epsilon}(A, x) \mid x \in \{\epsilon\}^*\}$

► $\hat{\Delta}_N: 2^Q \times \Sigma^* \rightarrow 2^Q$ is inductively defined by

$$\hat{\Delta}_N(A, \epsilon) = C_\epsilon(A)$$

$$\hat{\Delta}_N(A, xa) = \bigcup \{C_\epsilon(\Delta(q, a)) \mid q \in \hat{\Delta}_N(A, x)\}$$

Theorem

- every set accepted by NFA_ϵ is regular
- regular sets are **effectively** closed under **union**, **concatenation**, and **asterate**

Automata

- ▶ (deterministic, nondeterministic, alternating) **finite automata**
- ▶ **regular expressions**
- ▶ (alternating) Büchi automata

Logic

- ▶ (weak) monadic second-order logic
- ▶ Presburger arithmetic
- ▶ linear-time temporal logic

Definitions

- ▶ **regular expression** α over alphabet Σ :

$a \in \Sigma$ ϵ $\emptyset$ $\beta + \gamma$ $\beta\gamma$ β^*

- ▶ set of strings $L(\alpha) \subseteq \Sigma^*$ matched by regular expression α :

$L(a) = \{a\}$ $L(\beta + \gamma) = L(\beta) \cup L(\gamma)$
 $L(\epsilon) = \{\epsilon\}$ $L(\beta\gamma) = L(\beta)L(\gamma)$
 $L(\emptyset) = \emptyset$ $L(\beta^*) = L(\beta)^*$

Example

regular expression $(a + b)^*b$ matches all strings over $\Sigma = \{a, b\}$ that end with b

Definition

regular expressions α and β are **equivalent** ($\alpha \equiv \beta$) if $L(\alpha) = L(\beta)$

Outline

1. Summary of Previous Lecture
2. Regular Expressions
3. Intermezzo
4. Homomorphisms
5. Decision Problems
6. Further Reading

Theorem

finite automata and regular expressions are **equivalent**:

for all $A \subseteq \Sigma^*$ A is regular $\iff A = L(\alpha)$ for some regular expression α

Proof ($\Leftarrow$)

induction on regular expression α

α	$L(\alpha)$	finite automaton	α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$	$\longrightarrow \bigcirc \xrightarrow{a} \bigodot$	$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
ϵ	$\{\epsilon\}$	$\longrightarrow \bigodot$	$\beta\gamma$	$L(\beta)L(\gamma)$
$\emptyset$	$\emptyset$	$\longrightarrow \bigcirc$	β^*	$L(\beta)^*$

$L(\beta)$ and $L(\gamma)$ are regular according to induction hypothesis

$\implies L(\alpha)$ is regular according to closure properties of regular sets

Lemma (Arden's Lemma)

if $A, B, X \subseteq \Sigma^*$ such that $X = AX \cup B$ and $\epsilon \notin A$ then $X = A^*B$

Proof

$X \subseteq A^*B$

- ▶ let $x \in X = AX \cup B$
- ▶ induction on $|x|$
 - ▶ $x \in AX \implies x = ay$ for some $a \in A$ and $y \in X \implies y \in A^*B \implies x \in A^*B$
 - ▶ $x \in B \implies x \in A^*B$ because $\epsilon \in A^*$

$X \supseteq A^*B$

- ▶ $x \in A^*B \implies x = x_1 \cdots x_k y$ for some $x_1, \dots, x_k \in A$ and $y \in B$
- ▶ induction on k
 - ▶ $k = 0 \implies x = y \in B \subseteq X$
 - ▶ $k > 0 \implies x_2 \cdots x_k y \in X \implies x \in AX \subseteq X$

Theorem

finite automata and regular expressions are **equivalent**:

for all $A \subseteq \Sigma^*$ A is regular $\iff A = L(\alpha)$ for some regular expression α

Proof ($\implies$)

given NFA $N = (Q, \Sigma, \Delta, S, F)$ with $Q = \{1, \dots, n\}$ and $S = \{1\}$

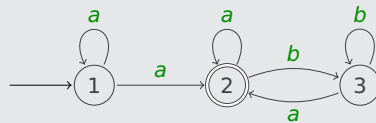
- ▶ define system of equations

$$X_i = \left(\bigcup_{a \in \Sigma} \bigcup_{j \in \Delta(i, a)} \{a\} X_j \right) \cup \begin{cases} \{\epsilon\} & \text{if } i \in F \\ \emptyset & \text{otherwise} \end{cases}$$

with unknowns $X_1, \dots, X_n$

- ▶ transform X_1 into regular expression by successive substitution and Arden's lemma

Example



$X_1 = aX_1 + aX_2$	$X_2 = aX_2 + bX_3 + \epsilon$	$X_3 = aX_2 + bX_3$	
$X_1 = a^*aX_2$	$X_2 = a^*(bX_3 + \epsilon)$	$X_3 = b^*aX_2$	(Arden's lemma)
$X_1 = a^*aX_2$	$X_2 = a^*(bb^*aX_2 + \epsilon)$		(substitute)
$X_1 = a^*aX_2$	$X_2 = a^*bb^*aX_2 + a^*$		(distribute)
$X_1 = a^*aX_2$	$X_2 = (a^*bb^*a)^*a^*$		(Arden's lemma)
$X_1 = a^*a(a^*bb^*a)^*a^*$			(substitute)

Outline

1. Summary of Previous Lecture
2. Regular Expressions
3. Intermezzo
4. Homomorphisms
5. Decision Problems
6. Further Reading

Question

Which of the following strings belong to $L((a + ba + bab)^*)$?

- A** ϵ
- B** $ababa$
- C** all strings over $\{a, b\}$ that start with a
- D** all strings over $\{a, b\}$ that do not contain two consecutive b 's



Theorem

regular sets are effectively closed under **homomorphic image** and **preimage**

Definitions

► **homomorphism** is mapping $h: \Sigma^* \rightarrow \Gamma^*$ such that

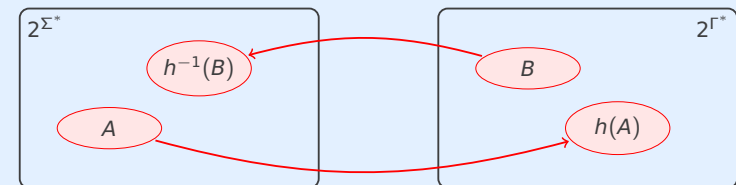
$$h(\epsilon) = \epsilon \quad h(xy) = h(x)h(y)$$

so homomorphism is completely determined by its effect on Σ

- if $A \subseteq \Sigma^*$ then $h(A) = \{h(x) \mid x \in A\} \subseteq \Gamma^*$ "image of A under h "
- if $B \subseteq \Gamma^*$ then $h^{-1}(B) = \{x \mid h(x) \in B\} \subseteq \Sigma^*$ "preimage of B under h "

Outline

1. Summary of Previous Lecture
2. Regular Expressions
3. Intermezzo
- 4. Homomorphisms**
5. Decision Problems
6. Further Reading



- homomorphism $h: \Sigma^* \rightarrow \Gamma^*$
- $h^{-1}(h(A)) \supseteq A$
- $h(h^{-1}(B)) \subseteq B$

Example

$\Sigma = \Gamma = \{0, 1\}$ $h(0) = 11$ $h(1) = 1$ $A = B = \{0\}$

- $h^{-1}(h(A)) = h^{-1}(\{11\}) = \{0, 11\} \supsetneq A$
- $h(h^{-1}(B)) = h(\emptyset) = \emptyset \subsetneq B$

Example

$A \subseteq \{0, 1\}^*$ is regular $\implies \{xy \mid x1y \in A\}$ is regular

- ▶ $\Sigma = \{0, 1\}$ and $\Gamma = \{0, 1, 2\}$
- ▶ define homomorphisms $h, i: \Gamma^* \rightarrow \Sigma^*$ by

$$h(0) = 0 \quad h(1) = h(2) = 1 \quad i(0) = 0 \quad i(1) = 1 \quad i(2) = \epsilon$$

- ▶ $h^{-1}(A) = \{x \in \Gamma^* \mid h(x) \in A\}$
- ▶ $h^{-1}(A) \cap L((0+1)^*2(0+1)^*) = \{x2y \mid x1y \in A\}$
- ▶ $\{xy \mid x1y \in A\} = i(h^{-1}(A) \cap L((0+1)^*2(0+1)^*))$ is regular

Theorem

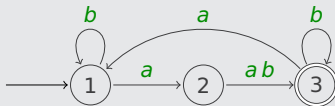
regular sets are effectively closed under homomorphic image and **preimage**

Proof

- ▶ NFA $M = (Q, \Gamma, \Delta, S, F)$
- ▶ homomorphism $h: \Sigma^* \rightarrow \Gamma^*$
- ▶ $h^{-1}(L(M)) = L(M')$ for NFA $M' = (Q, \Sigma, \Delta', S, F)$ with $\Delta'(q, a) = \widehat{\Delta}(\{q\}, h(a))$
- ▶ claim: $\widehat{\Delta'}(A, x) = \widehat{\Delta}(A, h(x))$ for all $A \subseteq Q$ and $x \in \Sigma^*$
- proof of claim: easy induction on $|x|$

Example

- ▶ DFA M



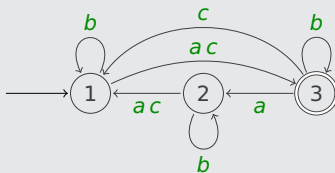
- ▶ homomorphism $h: \{a, b, c\}^* \rightarrow \{a, b\}^*$

$$h(a) = aa$$

$$h(b) = \epsilon$$

$$h(c) = bab$$

- ▶ DFA M'



$$\delta'(3, c) = \widehat{\delta}(3, bab) = 1$$

Theorem

regular sets are effectively closed under **homomorphic image** and preimage

Proof

- ▶ regular expression α over Σ
- ▶ homomorphism $h: \Sigma^* \rightarrow \Gamma^*$
- ▶ $h(L(\alpha)) = L(\alpha')$ for regular expression α' defined inductively:

$$a' = h(a) \text{ for } a \in \Sigma$$

$$\epsilon' = \epsilon$$

$$\emptyset' = \emptyset$$

$$(\beta + \gamma)' = \beta' + \gamma'$$

$$(\beta\gamma)' = \beta'\gamma'$$

$$\beta^{*'} = \beta'^*$$

Definitions

- ▶ **Hamming distance** $H(x, y)$ is number of places where bit strings x and y differ
- ▶ if $|x| \neq |y|$ then $H(x, y) = \infty$
- ▶ $N_k(A) = \{x \in \{0, 1\}^* \mid H(x, y) \leq k \text{ for some } y \in A\}$

Lemma

$A \subseteq \{0, 1\}^*$ is regular $\implies \forall k \in \mathbb{N} \ N_k(A)$ is regular

Proof

$D_k = \{x \in (\{0, 1\} \times \{0, 1\})^* \mid x \text{ contains at most } k \text{ pairs } (0, 1) \text{ or } (1, 0)\}$ is regular
 $= \{x \in (\{0, 1\} \times \{0, 1\})^* \mid H(\text{fst}(x), \text{snd}(x)) \leq k\}$

$N_k(A) = \text{fst}(\text{snd}^{-1}(A) \cap D_k)$

Example

▶ $A = \{0011\} \quad k = 2$

▶ $N_k(A)$ consists of

0	0	1	1	1	0	1	1	0	1	1	1	0	0	0	1	0	0	1	0	1	1	1	1
1	0	0	1	1	0	1	0	0	1	0	1	0	1	1	0	0	0	0	0				

▶ $\text{fst}(\text{snd}^{-1}(A) \cap D_k)$ consists of

0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	1
0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
0	1	1	0	0	1	1	1	1	0	0	0	1	0	0	1	1	0	1	0	1	0	1	1
0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1								
0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1				

Outline

1. Summary of Previous Lecture
2. Regular Expressions
3. Intermezzo
4. Homomorphisms
5. Decision Problems
6. Further Reading

Remark

most decision problems concerning regular sets are decidable

Theorem

problems

instance: DFA M and string x

question: $x \in L(M)$?

instance: DFA M

question: $L(M) = \emptyset$?

instance: DFAs M and N

question: $L(M) = L(N)$?

are decidable

Remark

representation of regular sets (DFA, NFA, regular expression) may affect complexity of decision problems

Outline

- 1. Summary of Previous Lecture
- 2. Regular Expressions
- 3. Intermezzo
- 4. Homomorphisms
- 5. Decision Problems
- 6. Further Reading

Kozen

- ▶ Lecture 7–10

Important Concepts

- ▶ Arden's lemma
- ▶ homomorphic image
- ▶ regular expression
- ▶ homomorphism
- ▶ homomorphic preimage

homework for October 24