

CMPE 322/327 - Theory of Computation

Week 4: Pattern Matching & Regular Expressions

Burak Ekici

March 14-18, 2022

Outline

- 1 A Quick Recap
- 2 Pattern Matching
- 3 Regular Expressions
- 4 Homomorphisms

Definitions

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

Definitions

- **nondeterministic finite automaton (NFA)** is quintuple $N = (Q, \Sigma, \Delta, s, F)$ with
 - 1 Q : finite set of states

Definitions

- **nondeterministic finite automaton (NFA)** is quintuple $N = (Q, \Sigma, \Delta, s, F)$ with
 - 1 Q : finite set of states
 - 2 Σ : input alphabet

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

Definitions

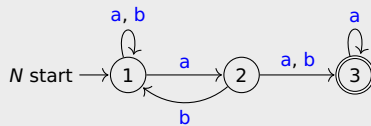
- **nondeterministic finite automaton (NFA)** is quintuple $N = (Q, \Sigma, \Delta, s, F)$ with
 - ① Q : finite set of states
 - ② Σ : input alphabet
 - ③ $\Delta : Q \times \Sigma \rightarrow 2^Q$: transition function
 - ④ $S \subseteq Q$: set of start states

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

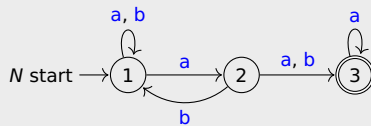
Example

$N = (Q, \Sigma, \Delta, S, F)$



Example

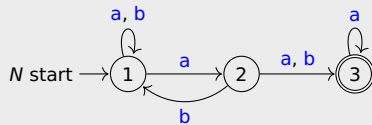
$N = (Q, \Sigma, \Delta, S, F)$



① $Q = \{1, 2, 3\}$

Example

$N = (Q, \Sigma, \Delta, S, F)$

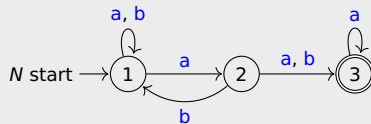


① $Q = \{1, 2, 3\}$

② $\Sigma = \{a, b\}$

Example

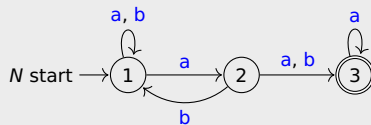
$N = (Q, \Sigma, \Delta, S, F)$



- ① $Q = \{1, 2, 3\}$
- ② $\Sigma = \{a, b\}$
- ③ $\Delta : Q \times \Sigma \rightarrow 2^Q$

Example

$N = (Q, \Sigma, \Delta, S, F)$

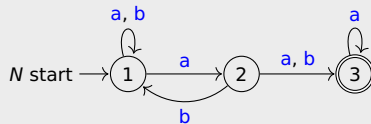


- ① $Q = \{1, 2, 3\}$
- ② $\Sigma = \{a, b\}$
- ③ $\Delta : Q \times \Sigma \rightarrow 2^Q$

Δ	a	b
1	$\{1, 2\}$	$\{1\}$
2	$\{3\}$	$\{1, 3\}$
3	$\{3\}$	\emptyset

Example

$N = (Q, \Sigma, \Delta, S, F)$

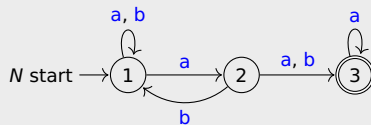


- ① $Q = \{1, 2, 3\}$
- ② $\Sigma = \{a, b\}$
- ③ $\Delta : Q \times \Sigma \rightarrow 2^Q$
- ④ $S = \{1\}$

Δ	a	b
1	{1, 2}	{1}
2	{3}	{1, 3}
3	{3}	\emptyset

Example

$N = (Q, \Sigma, \Delta, S, F)$



- ① $Q = \{1, 2, 3\}$
- ② $\Sigma = \{a, b\}$
- ③ $\Delta : Q \times \Sigma \rightarrow 2^Q$
- ④ $S = \{1\}$
- ⑤ $F = \{3\}$

Δ	a	b
1	$\{1, 2\}$	$\{1\}$
2	$\{3\}$	$\{1, 3\}$
3	$\{3\}$	\emptyset

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

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

$$\widehat{\Delta}(A, \epsilon) := A$$

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

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

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

$$\widehat{\Delta}(A, \epsilon) := A$$

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

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

Definitions

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

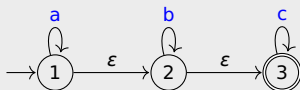
Definitions

- NFA with ε -transitions (NFA_ε) is sextuple $N = (Q, \Sigma, \varepsilon, \Delta, S, F)$ such that
 - 1 $\varepsilon \notin \Sigma$

Definitions

- NFA with ε -transitions (NFA_ε) is sextuple $N = (Q, \Sigma, \varepsilon, \Delta, S, F)$ such that
 - ① $\varepsilon \notin \Sigma$
 - ② $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$

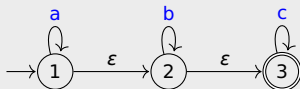
Example



Definitions

- NFA with ε -transitions (NFA_ε) is sextuple $N = (Q, \Sigma, \varepsilon, \Delta, S, F)$ such that
 - ① $\varepsilon \notin \Sigma$
 - ② $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$
- ε -closure of set $A \subseteq Q$ is defined as $C_\varepsilon(A) = \bigcup \{\widehat{\Delta}_{N_\varepsilon}(A, x) \mid x \in \{\varepsilon\}^*\}$

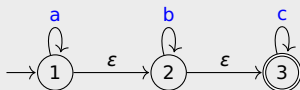
Example



Definitions

- NFA with ε -transitions (NFA_ε) is sextuple $N = (Q, \Sigma, \varepsilon, \Delta, S, F)$ such that
 - $\varepsilon \notin \Sigma$
 - $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$
- ε -closure of set $A \subseteq Q$ is defined as $C_\varepsilon(A) = \bigcup \{\widehat{\Delta}_{N_\varepsilon}(A, x) \mid x \in \{\varepsilon\}^*\}$

Example



$$C_\varepsilon(\{1\}) = \{1, 2, 3\}$$

Definitions

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

① $\varepsilon \notin \Sigma$

② $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$

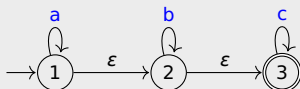
- ε -closure of set $A \subseteq Q$ is defined as $C_\varepsilon(A) = \bigcup \{\widehat{\Delta}_{N_\varepsilon}(A, x) \mid x \in \{\varepsilon\}^*\}$

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

$$\widehat{\Delta}_N(A, \varepsilon) = C_\varepsilon(A)$$

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

Example



$$C_\varepsilon(\{1\}) = \{1, 2, 3\}$$

Definitions

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

① $\varepsilon \notin \Sigma$

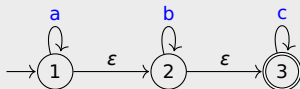
② $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$

- ε -closure of set $A \subseteq Q$ is defined as $C_\varepsilon(A) = \bigcup \{\widehat{\Delta}_{N_\varepsilon}(A, x) \mid x \in \{\varepsilon\}^*\}$
- $\widehat{\Delta}_N: 2^Q \times \Sigma^* \rightarrow 2^Q$ is inductively defined by

$$\widehat{\Delta}_N(A, \varepsilon) = C_\varepsilon(A)$$

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

Example



$$C_\varepsilon(\{1\}) = \{1, 2, 3\}$$

$$\widehat{\Delta}(\{1\}, b) =$$

Definitions

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

① $\varepsilon \notin \Sigma$

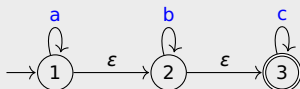
② $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$

- ε -closure of set $A \subseteq Q$ is defined as $C_\varepsilon(A) = \bigcup \{\widehat{\Delta}_{N_\varepsilon}(A, x) \mid x \in \{\varepsilon\}^*\}$
- $\widehat{\Delta}_N: 2^Q \times \Sigma^* \rightarrow 2^Q$ is inductively defined by

$$\widehat{\Delta}_N(A, \varepsilon) = C_\varepsilon(A)$$

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

Example



$$C_\varepsilon(\{1\}) = \{1, 2, 3\}$$

$$\widehat{\Delta}(\{1\}, b) = C_\varepsilon(\{1\})$$

Definitions

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

① $\varepsilon \notin \Sigma$

② $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$

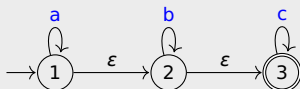
- ε -closure of set $A \subseteq Q$ is defined as $C_\varepsilon(A) = \bigcup \{\widehat{\Delta}_{N_\varepsilon}(A, x) \mid x \in \{\varepsilon\}^*\}$

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

$$\widehat{\Delta}_N(A, \varepsilon) = C_\varepsilon(A)$$

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

Example



$$C_\varepsilon(\{1\}) = \{1, 2, 3\}$$

$$\widehat{\Delta}(\{1\}, b) = C_\varepsilon(\Delta(1, b)) \cup C_\varepsilon(\Delta(2, b)) \cup C_\varepsilon(\Delta(3, b))$$

Definitions

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

① $\varepsilon \notin \Sigma$

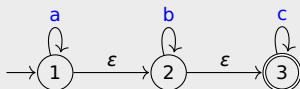
② $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$

- ε -closure of set $A \subseteq Q$ is defined as $C_\varepsilon(A) = \bigcup \{\widehat{\Delta}_{N_\varepsilon}(A, x) \mid x \in \{\varepsilon\}^*\}$
- $\widehat{\Delta}_N: 2^Q \times \Sigma^* \rightarrow 2^Q$ is inductively defined by

$$\widehat{\Delta}_N(A, \varepsilon) = C_\varepsilon(A)$$

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

Example



$$C_\varepsilon(\{1\}) = \{1, 2, 3\}$$

$$\widehat{\Delta}(\{1\}, b) = C_\varepsilon(\emptyset) \cup C_\varepsilon(\{2\}) \cup C_\varepsilon(\emptyset)$$

Definitions

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

① $\varepsilon \notin \Sigma$

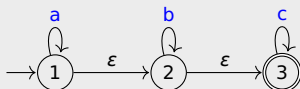
② $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$

- ε -closure of set $A \subseteq Q$ is defined as $C_\varepsilon(A) = \bigcup \{\widehat{\Delta}_{N_\varepsilon}(A, x) \mid x \in \{\varepsilon\}^*\}$
- $\widehat{\Delta}_N: 2^Q \times \Sigma^* \rightarrow 2^Q$ is inductively defined by

$$\widehat{\Delta}_N(A, \varepsilon) = C_\varepsilon(A)$$

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

Example



$$C_\varepsilon(\{1\}) = \{1, 2, 3\}$$

$$\widehat{\Delta}(\{1\}, b) = \emptyset \cup \{2, 3\} \cup \emptyset$$

Definitions

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

① $\varepsilon \notin \Sigma$

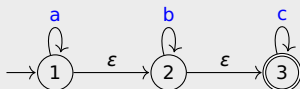
② $N_\varepsilon = (Q, \Sigma \cup \{\varepsilon\}, \Delta, S, F)$ is NFA over alphabet $\Sigma \cup \{\varepsilon\}$

- ε -closure of set $A \subseteq Q$ is defined as $C_\varepsilon(A) = \bigcup \{\widehat{\Delta}_{N_\varepsilon}(A, x) \mid x \in \{\varepsilon\}^*\}$
- $\widehat{\Delta}_N: 2^Q \times \Sigma^* \rightarrow 2^Q$ is inductively defined by

$$\widehat{\Delta}_N(A, \varepsilon) = C_\varepsilon(A)$$

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

Example



$$C_\varepsilon(\{1\}) = \{1, 2, 3\}$$

$$\widehat{\Delta}(\{1\}, b) = \{2, 3\}$$

Theorem

every set accepted by NFA is regular

Theorem

every set accepted by NFA is regular

Theorem

every set accepted by NFA_ϵ is regular

Theorem

every set accepted by NFA is regular

Theorem

every set accepted by NFA_ϵ is regular

Theorem

regular sets are effectively closed under **concatenation**

Theorem

every set accepted by NFA is regular

Theorem

every set accepted by NFA_ϵ is regular

Theorem

regular sets are effectively closed under **concatenation**

Theorem

regular sets are effectively closed under **asterate**

Outline

- 1 A Quick Recap
- 2 Pattern Matching
- 3 Regular Expressions
- 4 Homomorphisms

Pattern matching is important for

- lexical analysis of programs
- scripting languages (Perl, Ruby)
- search engines (Google Code Search)
- DNA analysis

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern foo

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Some Patterns

`^` matches beginning of line

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Some Patterns

- `^` matches beginning of line
- `$` matches end of line

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Some Patterns

- | | |
|-----------------|----------------------------------|
| <code>^</code> | matches beginning of line |
| <code>\$</code> | matches end of line |
| <code>c</code> | matches character <code>c</code> |

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Some Patterns

<code>^</code>	matches beginning of line	<code>.</code>	matches any character
<code>\$</code>	matches end of line		
<code>c</code>	matches character <code>c</code>		

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Some Patterns

<code>^</code>	matches beginning of line	<code>.</code>	matches any character
<code>\$</code>	matches end of line	<code>[abc]</code>	matches a or b or c
<code>c</code>	matches character c		

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Some Patterns

<code>^</code>	matches beginning of line	<code>.</code>	matches any character
<code>\$</code>	matches end of line	<code>[abc]</code>	matches a or b or c
<code>c</code>	matches character c	<code>[a-zA-Z]</code>	matches any letter

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Some Patterns

<code>^</code>	matches beginning of line	<code>.</code>	matches any character
<code>\$</code>	matches end of line	<code>[abc]</code>	matches a or b or c
<code>c</code>	matches character c	<code>[a-zA-Z]</code>	matches any letter

Example

`grep "0" file` returns lines containing 0

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Some Patterns

<code>^</code>	matches beginning of line	<code>.</code>	matches any character
<code>\$</code>	matches end of line	<code>[abc]</code>	matches a or b or c
<code>c</code>	matches character c	<code>[a-zA-Z]</code>	matches any letter

Example

```
grep "0" file    returns lines containing 0
grep "0$" file   returns lines ending with 0
```

Applications of Regular expressions: grep

- `grep foo file` returns lines in file containing pattern `foo`
- basis for more powerful tools like `awk`, `sed`, `perl`

Some Patterns

<code>^</code>	matches beginning of line	<code>.</code>	matches any character
<code>\$</code>	matches end of line	<code>[abc]</code>	matches a or b or c
<code>c</code>	matches character c	<code>[a-zA-Z]</code>	matches any letter

Example

```
grep "0" file      returns lines containing 0
grep "0$" file     returns lines ending with 0
grep "b.g" file    returns lines containing e.g. bag, big, bug, buggy
```

Pattern matching is important for

- lexical analysis of programs
- scripting languages (Perl, Ruby)
- search engines (Google Code Search)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

Pattern matching is important for

- lexical analysis of programs
- scripting languages (Perl, Ruby)
- search engines (Google Code Search)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

$$\frac{\text{atomic pattern } \alpha \quad L(\alpha)}{\mathbf{a} \in \Sigma \quad \{a\}}$$

-

Pattern matching is important for

- lexical analysis of programs
- scripting languages (Perl, Ruby)
- search engines (Google Code Search)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$\mathbf{a} \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$

-

Pattern matching is important for

- lexical analysis of programs
- scripting languages (Perl, Ruby)
- search engines (Google Code Search)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
• \emptyset	\emptyset

Pattern matching is important for

- lexical analysis of programs
- scripting languages (Perl, Ruby)
- search engines (Google Code Search)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$\mathbf{a} \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
• \emptyset	\emptyset
$\#$	Σ

Pattern matching is important for

- lexical analysis of programs
- search engines (Google Code Search)
- scripting languages (Perl, Ruby)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
\emptyset	\emptyset
$\#$	Σ
$@$	Σ^*

Pattern matching is important for

- lexical analysis of programs
- search engines (Google Code Search)
- scripting languages (Perl, Ruby)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
\emptyset	\emptyset
$\#$	Σ
$@$	Σ^*

compound pattern α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$

Pattern matching is important for

- lexical analysis of programs
- search engines (Google Code Search)
- scripting languages (Perl, Ruby)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
\emptyset	\emptyset
$\#$	Σ
$@$	Σ^*

compound pattern α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$

Pattern matching is important for

- lexical analysis of programs
- search engines (Google Code Search)
- scripting languages (Perl, Ruby)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
\emptyset	\emptyset
$\#$	Σ
$@$	Σ^*

compound pattern α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta \gamma$	$L(\beta)L(\gamma)$

Pattern matching is important for

- lexical analysis of programs
- search engines (Google Code Search)
- scripting languages (Perl, Ruby)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
• \emptyset	\emptyset
$\#$	Σ
$@$	Σ^*

compound pattern α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta \gamma$	$L(\beta)L(\gamma)$
β^*	$L(\beta)^*$

Pattern matching is important for

- lexical analysis of programs
- search engines (Google Code Search)
- scripting languages (Perl, Ruby)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
• \emptyset	\emptyset
$\#$	Σ
$@$	Σ^*

compound pattern α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta\gamma$	$L(\beta)L(\gamma)$
β^*	$L(\beta)^*$
β^+	$L(\beta)^+$

Pattern matching is important for

- lexical analysis of programs
- search engines (Google Code Search)
- scripting languages (Perl, Ruby)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
• \emptyset	\emptyset
$\#$	Σ
$@$	Σ^*

compound pattern α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta \gamma$	$L(\beta)L(\gamma)$
β^*	$L(\beta)^*$
β^+	$L(\beta)^+$
$\sim \beta$	$\sim L(\beta) = \Sigma^* - L(\beta)$

Pattern matching is important for

- lexical analysis of programs
- search engines (Google Code Search)
- scripting languages (Perl, Ruby)
- DNA analysis

Definitions

- pattern is string α that represents set of strings $L(\alpha) \subseteq \Sigma^*$

atomic pattern α	$L(\alpha)$
$a \in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
• \emptyset	\emptyset
$\#$	Σ
$@$	Σ^*

compound pattern α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta \gamma$	$L(\beta)L(\gamma)$
β^*	$L(\beta)^*$
β^+	$L(\beta)^+$
$\sim \beta$	$\sim L(\beta) = \Sigma^* - L(\beta)$

- string $x \in \Sigma^*$ **matches** pattern α if $x \in L(\alpha)$

Example

pattern

@a@a@a@

matched string

strings containing at least 3 occurrences of a

Example

pattern	matched string
@a@a@a@	strings containing at least 3 occurrences of <i>a</i>
@a@b@	strings containing <i>a</i> followed later by <i>b</i>

Example

pattern	matched string
@a@a@a@	strings containing at least 3 occurrences of <i>a</i>
@a@b@	strings containing <i>a</i> followed later by <i>b</i>
#n ~ a	single letters except <i>a</i>

Example

pattern	matched string
@a@a@a@	strings containing at least 3 occurrences of <i>a</i>
@a@b@	strings containing <i>a</i> followed later by <i>b</i>
#n ~ <i>a</i>	single letters except <i>a</i>
(#n ~ <i>a</i>)*	strings without <i>a</i>

Questions

- how difficult is pattern matching?

Questions

- how difficult is pattern matching?
- is pattern equivalence ($L(\alpha) = L(\beta)$) decidable?

Questions

- how difficult is pattern matching?
- is pattern equivalence ($L(\alpha) = L(\beta)$) decidable?
- which operators are **redundant**?

Questions

- how difficult is pattern matching?
- is pattern equivalence ($L(\alpha) = L(\beta)$) decidable?
- which operators are **redundant**?

$$\varepsilon \quad \equiv \quad \sim (\# @) \equiv \emptyset^*$$

Questions

- how difficult is pattern matching?
- is pattern equivalence ($L(\alpha) = L(\beta)$) decidable?
- which operators are **redundant**?

$$\begin{array}{lcl} \epsilon & \equiv & \sim (\# @) \equiv \emptyset^* \\ @ & \equiv & \#^* \end{array}$$

Questions

- how difficult is pattern matching?
- is pattern equivalence ($L(\alpha) = L(\beta)$) decidable?
- which operators are **redundant**?

$$\begin{array}{lll}
 \varepsilon & \equiv & \sim (\# @) \equiv \emptyset^* \\
 @ & \equiv & \#^* \\
 \alpha^+ & \equiv & \alpha \alpha^*
 \end{array}$$

Questions

- how difficult is pattern matching?
- is pattern equivalence ($L(\alpha) = L(\beta)$) decidable?
- which operators are **redundant**?

$$\begin{array}{lll}
 \varepsilon & \equiv & \sim (\# @) \equiv \emptyset^* \\
 @ & \equiv & \#^* \\
 \alpha^+ & \equiv & \alpha \alpha^* \\
 \# & \equiv & a_1 \dots a_n \quad \text{if } \Sigma = \{a_1 \dots a_n\}
 \end{array}$$

Questions

- how difficult is pattern matching?
- is pattern equivalence ($L(\alpha) = L(\beta)$) decidable?
- which operators are **redundant**?

$$\begin{array}{lll} \varepsilon & \equiv & \sim (\# @) \equiv \emptyset^* \\ @ & \equiv & \#^* \\ \alpha^+ & \equiv & \alpha \alpha^* \\ \# & \equiv & a_1 \dots a_n \quad \text{if } \Sigma = \{a_1 \dots a_n\} \\ \alpha \cap \beta & \equiv & \sim (\sim \alpha + \sim \beta) \end{array}$$

Questions

- how difficult is pattern matching?
- is pattern equivalence ($L(\alpha) = L(\beta)$) decidable?
- which operators are **redundant**?

ϵ	\equiv	$\sim (\# @) \equiv \emptyset^*$	
@	\equiv	$\#^*$	
α^+	\equiv	$\alpha\alpha^*$	
#	\equiv	$a_1 \dots a_n$	if $\Sigma = \{a_1 \dots a_n\}$
$\alpha \cap \beta$	\equiv	$\sim (\sim \alpha + \sim \beta)$	
$\sim \alpha$	\equiv	?	

Questions

- how difficult is pattern matching?
- is pattern equivalence ($L(\alpha) = L(\beta)$) decidable?
- which operators are **redundant**?

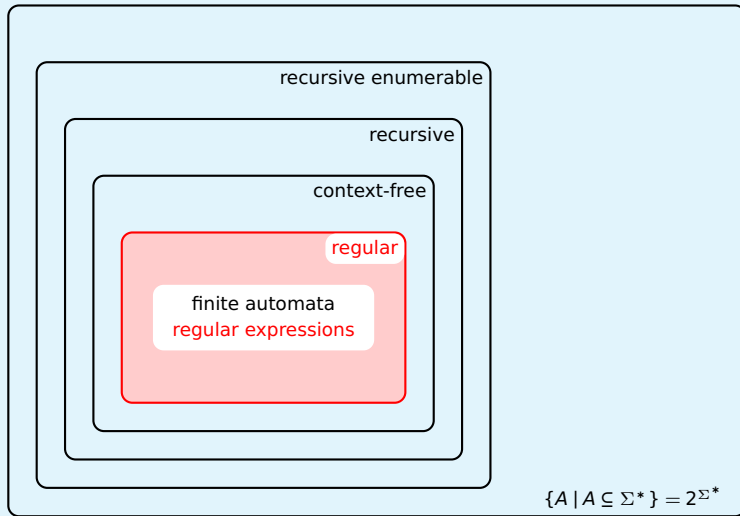
ε	\equiv	$\sim (\# @) \equiv \emptyset^*$	
@	\equiv	$\#^*$	
α^+	\equiv	$\alpha\alpha^*$	
#	\equiv	$a_1 \dots a_n$	if $\Sigma = \{a_1 \dots a_n\}$
$\alpha \cap \beta$	\equiv	$\sim (\sim \alpha + \sim \beta)$	
$\sim \alpha$	\equiv	?	

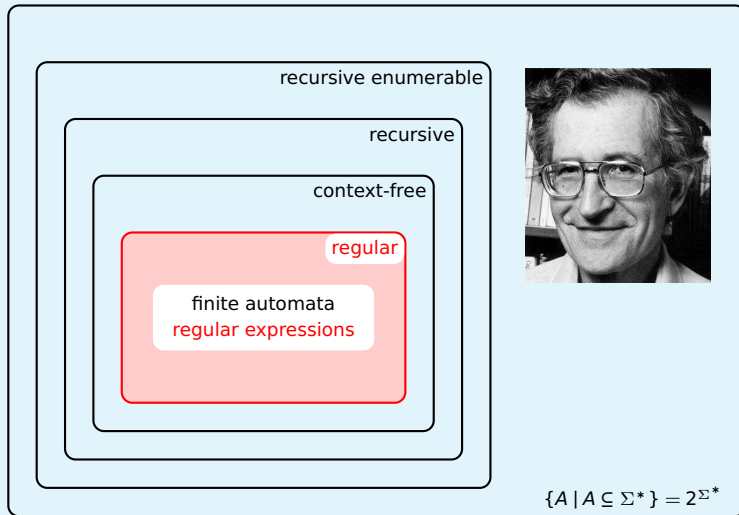
Notation

$\alpha \equiv \beta$ if $L(\alpha) = L(\beta)$

Outline

- 1 A Quick Recap
- 2 Pattern Matching
- 3 Regular Expressions**
- 4 Homomorphisms





Definition

regular expressions are restricted patterns which use only

$$a \in \Sigma \quad \varepsilon \quad \emptyset \quad \alpha + \beta \quad \alpha^* \quad \alpha\beta$$

Definition

regular expressions are restricted patterns which use only

$$a \in \Sigma \quad \varepsilon \quad \emptyset \quad \alpha + \beta \quad \alpha^* \quad \alpha\beta$$

Theorem

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

Definition

regular expressions are restricted patterns which use only

$$a \in \Sigma \quad \varepsilon \quad \emptyset \quad \alpha + \beta \quad \alpha^* \quad \alpha\beta$$

Theorem

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

- for all $A \subseteq \Sigma^*$
- ① A is regular
 - \iff ② $A = L(\alpha)$ for some pattern α
 - \iff ③ $A = L(\alpha)$ for some regular expression α

Definition

regular expressions are restricted patterns which use only

$$a \in \Sigma \quad \varepsilon \quad \emptyset \quad \alpha + \beta \quad \alpha^* \quad \alpha\beta$$

Theorem

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

- for all $A \subseteq \Sigma^*$
- ① A is regular
 - \iff ② $A = L(\alpha)$ for some pattern α
 - \iff ③ $A = L(\alpha)$ for some regular expression α

Proof.

- ③ \implies ② trivial (every regular expression is a pattern)

Definition

regular expressions are restricted patterns which use only

$$a \in \Sigma \quad \varepsilon \quad \emptyset \quad \alpha + \beta \quad \alpha^* \quad \alpha\beta$$

Theorem

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

- for all $A \subseteq \Sigma^*$
- ① A is regular
 - \iff ② $A = L(\alpha)$ for some pattern α
 - \iff ③ $A = L(\alpha)$ for some regular expression α

Proof.

- ③ \implies ② trivial (every regular expression is a pattern)
- ② \implies ① induction on α (see slides #16 – 17)

Definition

regular expressions are restricted patterns which use only

$$a \in \Sigma \quad \varepsilon \quad \emptyset \quad \alpha + \beta \quad \alpha^* \quad \alpha\beta$$

Theorem

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

- for all $A \subseteq \Sigma^*$
- ① A is regular
 - \iff ② $A = L(\alpha)$ for some pattern α
 - \iff ③ $A = L(\alpha)$ for some regular expression α

Proof.

- ③ \implies ② trivial (every regular expression is a pattern)
- ② \implies ① induction on α (see slides #16 – 17)
- ① \implies ③ (see slide #18)

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

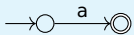
① atomic pattern $\Sigma = \{a_1, \dots, a_n\}$

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

① atomic pattern $\Sigma = \{a_1, \dots, a_n\}$

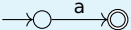
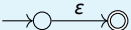
α	$L(\alpha)$	finite automaton
$a \in \Sigma$	$\{a\}$	

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

① atomic pattern $\Sigma = \{a_1, \dots, a_n\}$

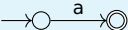
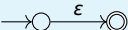

α	$L(\alpha)$	finite automaton
$a \in \Sigma$	$\{a\}$	
ϵ	$\{\epsilon\}$	

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

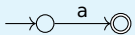
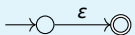

① atomic pattern $\Sigma = \{a_1, \dots, a_n\}$

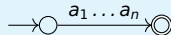
α	$L(\alpha)$	finite automaton
$a \in \Sigma$	$\{a\}$	
ϵ	$\{\epsilon\}$	
\emptyset	\emptyset	

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular
induction on pattern α

① atomic pattern $\Sigma = \{a_1, \dots, a_n\}$

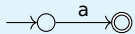
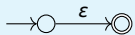

α	$L(\alpha)$	finite automaton
$a \in \Sigma$	$\{a\}$	
ϵ	$\{\epsilon\}$	
\emptyset	\emptyset	

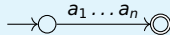

α	$L(\alpha)$	finite automaton
$\#$	Σ	

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular
induction on pattern α

① atomic pattern $\Sigma = \{a_1, \dots, a_n\}$

α	$L(\alpha)$	finite automaton
$a \in \Sigma$	$\{a\}$	
ϵ	$\{\epsilon\}$	
\emptyset	\emptyset	

α	$L(\alpha)$	finite automaton
$\#$	Σ	
$@$	Σ^*	

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

2 compound patterns

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

2 compound patterns

$$\frac{\alpha \quad L(\alpha)}{\beta + \gamma \quad L(\beta) \cup L(\gamma)}$$

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

2 compound patterns

α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

② compound patterns

α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta\gamma$	$L(\beta)L(\gamma)$

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

② compound patterns

α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta\gamma$	$L(\beta)L(\gamma)$

α	$L(\alpha)$
β^*	$L(\beta)^*$

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

② compound patterns

α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta\gamma$	$L(\beta)L(\gamma)$

α	$L(\alpha)$
β^*	$L(\beta)^*$
β^+	$L(\beta)^+$

Proof. (2 \implies 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

② compound patterns

α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta\gamma$	$L(\beta)L(\gamma)$

α	$L(\alpha)$
β^*	$L(\beta)^*$
β^+	$L(\beta)^+$
$\sim \beta$	$\sim L(\beta)$

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

2 compound patterns

α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta\gamma$	$L(\beta)L(\gamma)$

α	$L(\alpha)$
β^*	$L(\beta)^*$
β^+	$L(\beta)^+$
$\sim \beta$	$\sim L(\beta)$

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

Proof. (2 \Rightarrow 1)

statement: for any pattern α , $L(\alpha)$ is regular

induction on pattern α

2 compound patterns

α	$L(\alpha)$
$\beta + \gamma$	$L(\beta) \cup L(\gamma)$
$\beta \cap \gamma$	$L(\beta) \cap L(\gamma)$
$\beta\gamma$	$L(\beta)L(\gamma)$

α	$L(\alpha)$
β^*	$L(\beta)^*$
β^+	$L(\beta)^+$
$\sim \beta$	$\sim L(\beta)$

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

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

Proof. (1 \implies 3 – An idea)

given $\text{NFA}_\varepsilon N_\varepsilon = (Q, \Sigma, \varepsilon, \Delta, S, F)$

$\forall Y \subseteq Q \quad \forall u, v \in Q$ construct regular expression α_{uv}^Y such that

$x \in L(\alpha_{uv}^Y) \iff \exists$ a path from u to v labeled x ($v \in \widehat{\Delta}(\{u\}, x)$)
such that all intermediate states belong to Y

Proof. (1 \implies 3 – An idea)

given $NFA_{\epsilon} N_{\epsilon} = (Q, \Sigma, \epsilon, \Delta, S, F)$

$\forall Y \subseteq Q \quad \forall u, v \in Q$ construct regular expression α_{uv}^Y such that

$$x \in L(\alpha_{uv}^Y) \iff \boxed{\exists \text{ a path from } u \text{ to } v \text{ labeled } x \ (v \in \widehat{\Delta}(\{u\}, x)) \text{ such that all intermediate states belong to } Y}$$

Definitions

$$\bullet \alpha_{uv}^{\emptyset} := \begin{cases} a_1 + \dots + a_k & \text{if } u \neq v \text{ and } k > 0 \\ \emptyset & \text{if } u \neq v \text{ and } k = 0 \\ a_1 + \dots + a_k + \epsilon & \text{if } u = v \text{ and } k > 0 \\ \epsilon & \text{if } u = v \text{ and } k = 0 \end{cases}$$

$$\{a_1, \dots, a_k\} := \{a \in \Sigma \cup \{\epsilon\} \mid v \in \Delta(u, a)\}$$

Proof. (1 \implies 3 – An idea)

given $NFA_\epsilon N_\epsilon = (Q, \Sigma, \epsilon, \Delta, S, F)$

$\forall Y \subseteq Q \quad \forall u, v \in Q$ construct regular expression α_{uv}^Y such that

$$x \in L(\alpha_{uv}^Y) \iff \boxed{\exists \text{ a path from } u \text{ to } v \text{ labeled } x \text{ (} v \in \widehat{\Delta}(\{u\}, x) \text{) such that all intermediate states belong to } Y}$$

Definitions

- $\alpha_{uv}^\emptyset := \begin{cases} a_1 + \dots + a_k & \text{if } u \neq v \text{ and } k > 0 \\ \emptyset & \text{if } u \neq v \text{ and } k = 0 \\ a_1 + \dots + a_k + \epsilon & \text{if } u = v \text{ and } k > 0 \\ \epsilon & \text{if } u = v \text{ and } k = 0 \end{cases} \quad \{a_1, \dots, a_k\} := \{a \in \Sigma \cup \{\epsilon\} \mid v \in \Delta(u, a)\}$
- $\alpha_{uv}^Y := \alpha_{uv}^{Y-\{q\}} + \alpha_{uq}^{Y-\{q\}} (\alpha_{qq}^{Y-\{q\}})^* \alpha_{qv}^{Y-\{q\}} \quad \text{for some fixed } q \in Y$

Proof. (1 \implies 3 – An idea)

given $\text{NFA}_\varepsilon N_\varepsilon = (Q, \Sigma, \varepsilon, \Delta, S, F)$

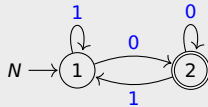
$\forall Y \subseteq Q \quad \forall u, v \in Q$ construct regular expression α_{uv}^Y such that

$x \in L(\alpha_{uv}^Y) \iff \exists \text{ a path from } u \text{ to } v \text{ labeled } x \text{ (} v \in \widehat{\Delta}(\{u\}, x) \text{)}$
such that all intermediate states belong to Y

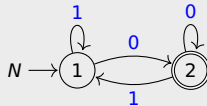
Theorem

$$L(N_\varepsilon) = L\left(\sum_{s \in S, t \in F} \alpha_{st}^Q\right)$$

Example

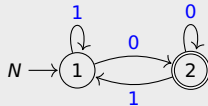


Example



$L(N) = L(\alpha)$ with

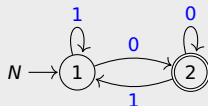
Example



$$L(N) = L(\alpha) \text{ with}$$

$$\alpha = \alpha_{12}^{\{1,2\}}$$

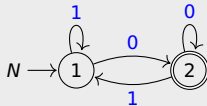
Example



$L(N) = L(\alpha)$ with

$$\alpha = \alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

Example

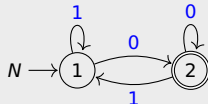


$L(N) = L(\alpha)$ with

$$\alpha = \alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

Example



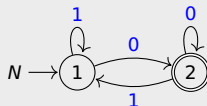
$L(N) = L(\alpha)$ with

$$\alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{22}^{\{1\}} = \alpha_{22}^{\emptyset} + \alpha_{21}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

Example



$L(N) = L(\alpha)$ with

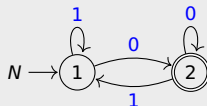
$$\alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{22}^{\{1\}} = \alpha_{22}^{\emptyset} + \alpha_{21}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{12}^{\emptyset} = 0$$

Example



$L(N) = L(\alpha)$ with

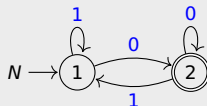
$$\alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{22}^{\{1\}} = \alpha_{22}^{\emptyset} + \alpha_{21}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{12}^{\emptyset} = 0 \quad \alpha_{11}^{\emptyset} = 1 + \epsilon$$

Example



$L(N) = L(\alpha)$ with

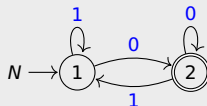
$$\alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{22}^{\{1\}} = \alpha_{22}^{\emptyset} + \alpha_{21}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{12}^{\emptyset} = 0 \quad \alpha_{11}^{\emptyset} = 1 + \epsilon \quad \alpha_{22}^{\emptyset} = 0 + \epsilon$$

Example



$L(N) = L(\alpha)$ with

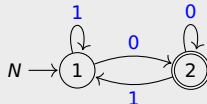
$$\alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{22}^{\{1\}} = \alpha_{22}^{\emptyset} + \alpha_{21}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{12}^{\emptyset} = 0 \quad \alpha_{11}^{\emptyset} = 1 + \epsilon \quad \alpha_{22}^{\emptyset} = 0 + \epsilon \quad \alpha_{21}^{\emptyset} = 1$$

Example



$L(N) = L(\alpha)$ with

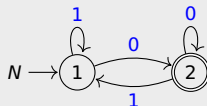
$$\alpha = \alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset} = 0 + (1 + \epsilon)(1 + \epsilon)^* 0$$

$$\alpha_{22}^{\{1\}} = \alpha_{22}^{\emptyset} + \alpha_{21}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset}$$

$$\alpha_{12}^{\emptyset} = 0 \quad \alpha_{11}^{\emptyset} = 1 + \epsilon \quad \alpha_{22}^{\emptyset} = 0 + \epsilon \quad \alpha_{21}^{\emptyset} = 1$$

Example



$L(N) = L(\alpha)$ with

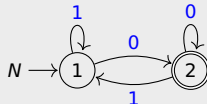
$$\alpha = \alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset} = 0 + (1 + \epsilon)(1 + \epsilon)^* 0$$

$$\alpha_{22}^{\{1\}} = \alpha_{22}^{\emptyset} + \alpha_{21}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset} = (0 + \epsilon) + 1(1 + \epsilon)^* 0$$

$$\alpha_{12}^{\emptyset} = 0 \quad \alpha_{11}^{\emptyset} = 1 + \epsilon \quad \alpha_{22}^{\emptyset} = 0 + \epsilon \quad \alpha_{21}^{\emptyset} = 1$$

Example



$L(N) = L(\alpha)$ with

$$\alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

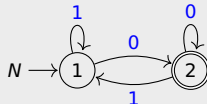
$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset} = 0 + (1 + \epsilon)(1 + \epsilon)^* 0$$

$$\alpha_{22}^{\{1\}} = \alpha_{22}^{\emptyset} + \alpha_{21}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset} = (0 + \epsilon) + 1(1 + \epsilon)^* 0$$

$$\alpha_{12}^{\emptyset} = 0 \quad \alpha_{11}^{\emptyset} = 1 + \epsilon \quad \alpha_{22}^{\emptyset} = 0 + \epsilon \quad \alpha_{21}^{\emptyset} = 1$$

$$\alpha = (0 + (1 + \epsilon)(1 + \epsilon)^* 0) + (0 + (1 + \epsilon)(1 + \epsilon)^* 0)((0 + \epsilon) + 1(1 + \epsilon)^* 0)^*((0 + \epsilon) + 1(1 + \epsilon)^* 0)$$

Example



$L(N) = L(\alpha)$ with

$$\alpha_{12}^{\{1,2\}} = \alpha_{12}^{\{1\}} + \alpha_{12}^{\{1\}} (\alpha_{22}^{\{1\}})^* \alpha_{22}^{\{1\}} \quad (q = 2)$$

$$\alpha_{12}^{\{1\}} = \alpha_{12}^{\emptyset} + \alpha_{11}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset} = 0 + (1 + \epsilon)(1 + \epsilon)^* 0$$

$$\alpha_{22}^{\{1\}} = \alpha_{22}^{\emptyset} + \alpha_{21}^{\emptyset} (\alpha_{11}^{\emptyset})^* \alpha_{12}^{\emptyset} = (0 + \epsilon) + 1(1 + \epsilon)^* 0$$

$$\alpha_{12}^{\emptyset} = 0 \quad \alpha_{11}^{\emptyset} = 1 + \epsilon \quad \alpha_{22}^{\emptyset} = 0 + \epsilon \quad \alpha_{21}^{\emptyset} = 1$$

$$\alpha = (0 + (1 + \epsilon)(1 + \epsilon)^* 0) + (0 + (1 + \epsilon)(1 + \epsilon)^* 0)((0 + \epsilon) + 1(1 + \epsilon)^* 0)^*((0 + \epsilon) + 1(1 + \epsilon)^* 0)$$

$$\equiv (0 + 1)^* 0$$

Outline

- 1 A Quick Recap
- 2 Pattern Matching
- 3 Regular Expressions
- 4 Homomorphisms**

Theorem

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

Theorem

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

Definitions

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

$$h(\varepsilon) = \varepsilon$$

$$h(xy) = h(x)h(y)$$

Theorem

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

Definitions

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

$$h(\varepsilon) = \varepsilon \qquad h(xy) = h(x)h(y)$$

so homomorphism is completely determined by its effect on Σ

Theorem

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

Definitions

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

$$h(\varepsilon) = \varepsilon \qquad h(xy) = h(x)h(y)$$

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 ”**

Theorem

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

Definitions

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

$$h(\varepsilon) = \varepsilon$$

$$h(xy) = h(x)h(y)$$

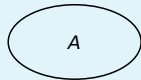
so homomorphism is completely determined by its effect on Σ

- 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 ”

2^{Σ^*}

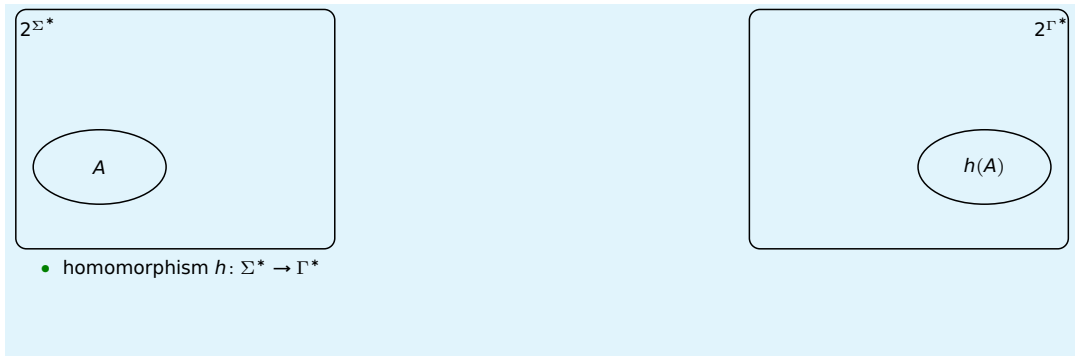
- homomorphism $h: \Sigma^* \rightarrow \Gamma^*$

 2^{Γ^*}

2^{Σ^*} 

- homomorphism $h: \Sigma^* \rightarrow \Gamma^*$

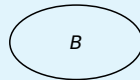
 2^{Γ^*}





2^{Σ^*}

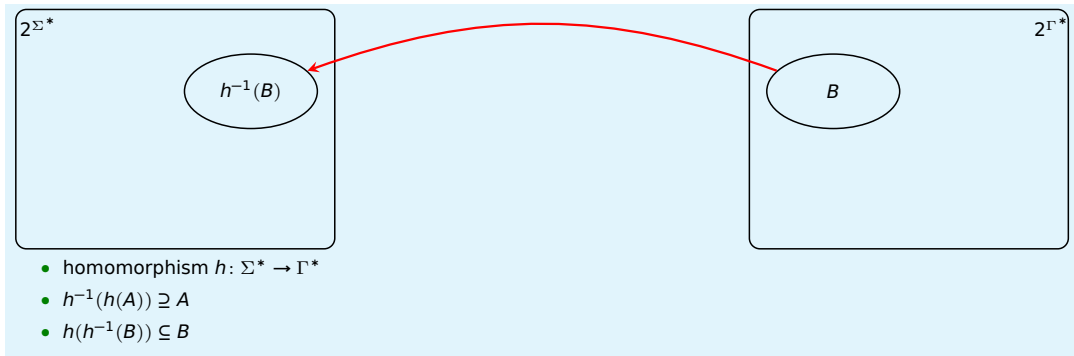
- homomorphism $h: \Sigma^* \rightarrow \Gamma^*$
- $h^{-1}(h(A)) \supseteq A$

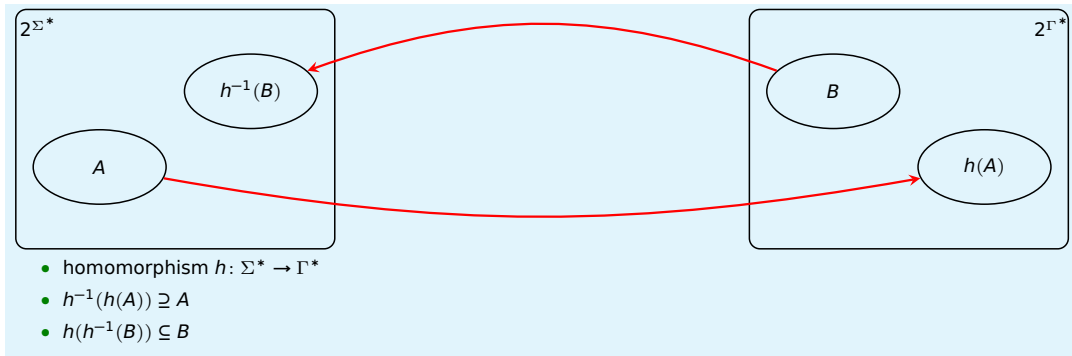
 2^{Γ^*} 

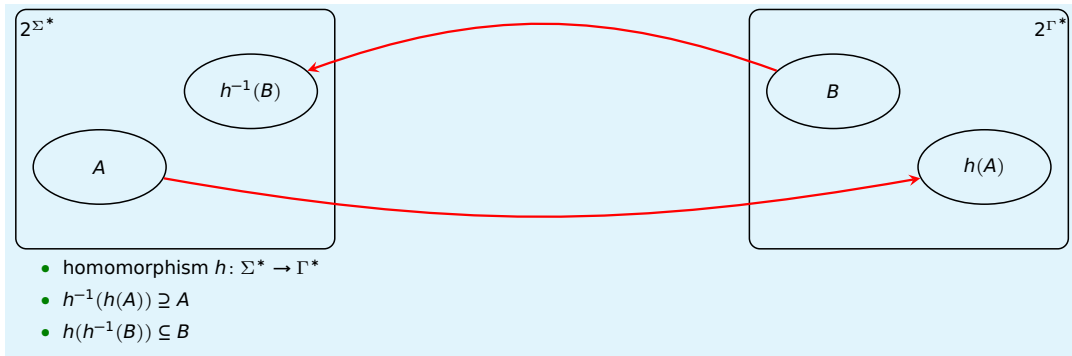
2^{Σ^*} $h^{-1}(B)$

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

 2^{Γ^*} B

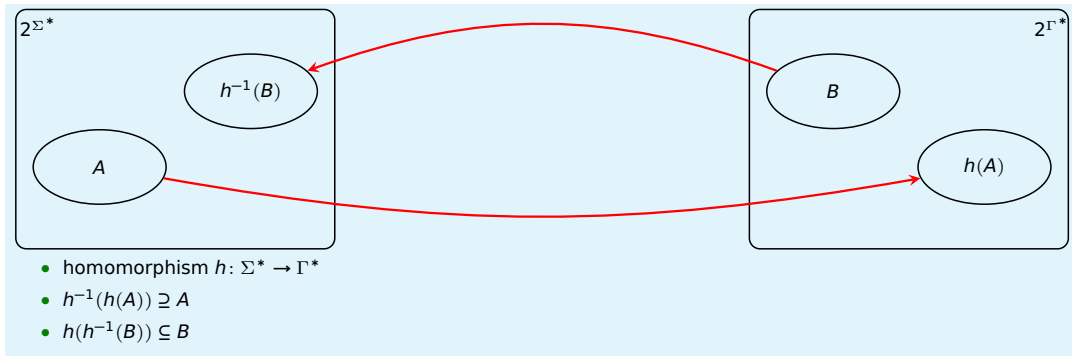






Example

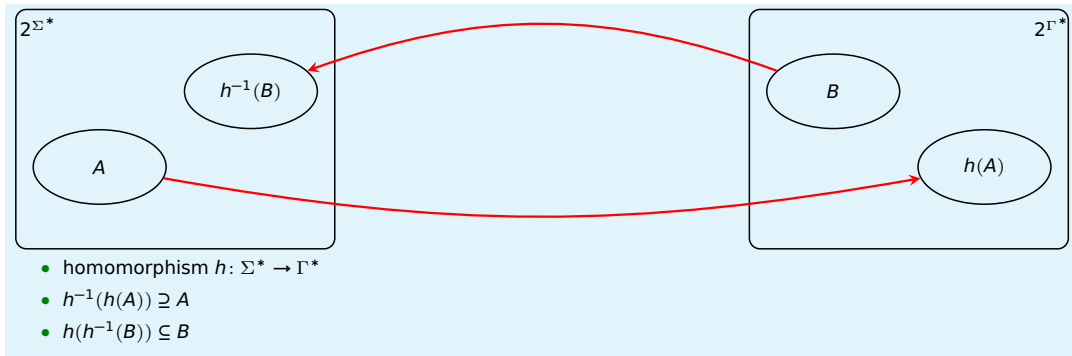
$$\Sigma = \Gamma = \{0, 1\} \quad h(0) = 11 \quad h(1) = 1$$



Example

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

- $h^{-1}(h(A)) = h^{-1}(\{11\}) = \{0, 11\} \supset A$



Example

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

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

Lemma

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

Lemma

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

Proof.

- $\Sigma = \{0, 1\}$ and $\Gamma = \{0, 1, 2\}$

Lemma

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

Proof.

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

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

Lemma

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

Proof.

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

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

- $h^{-1}(A) = \{x \mid h(x) \in A\}$

Lemma

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

Proof.

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

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

- $h^{-1}(A) = \{x \mid h(x) \in A\}$
- $h^{-1}(A) \cap L((0+1)^*2(0+1)^*) = \{x2y \mid x1y \in A\}$

Lemma

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

Proof.

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

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

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

Lemma

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

Proof.

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

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

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

Theorem

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

Theorem

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

Proof.

- DFA $M = (Q, \Gamma, \delta, s, F)$

Theorem

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

Proof.

- DFA $M = (Q, \Gamma, \delta, s, F)$
- homomorphism $h: \Sigma^* \rightarrow \Gamma^*$

Theorem

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

Proof.

- DFA $M = (Q, \Gamma, \delta, s, F)$
- homomorphism $h: \Sigma^* \rightarrow \Gamma^*$
- $h^{-1}(L(M)) = L(M')$ for DFA $M' = (Q, \Sigma, \delta', s, F)$ with $\delta'(q, a) := \widehat{\delta}(q, h(a))$

Theorem

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

Proof.

- DFA $M = (Q, \Gamma, \delta, s, F)$
- homomorphism $h: \Sigma^* \rightarrow \Gamma^*$
- $h^{-1}(L(M)) = L(M')$ for DFA $M' = (Q, \Sigma, \delta', s, F)$ with $\delta'(q, a) := \widehat{\delta}(q, h(a))$
- claim: $\widehat{\delta}'(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

Theorem

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

Proof.

- DFA $M = (Q, \Gamma, \delta, s, F)$
- homomorphism $h: \Sigma^* \rightarrow \Gamma^*$
- $h^{-1}(L(M)) = L(M')$ for DFA $M' = (Q, \Sigma, \delta', s, F)$ with $\delta'(q, a) := \widehat{\delta}(q, h(a))$
 - claim: $\widehat{\delta}'(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$
 - proof of claim: induction on $|x|$ (see next slide)

proof of the claim

claim: $\widehat{\delta'}(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

proof of the claim

claim: $\widehat{\delta'}(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

- base case: $|x| = 0$ thus $x = \varepsilon$

$$\widehat{\delta'}(q, \varepsilon) = q = \widehat{\delta}(q, h(\varepsilon))$$

proof of the claim

claim: $\widehat{\delta'}(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

- base case: $|x| = 0$ thus $x = \varepsilon$

$$\widehat{\delta'}(q, \varepsilon) = q = \widehat{\delta}(q, h(\varepsilon))$$

- step case: $|x| > 0$ thus $x = ya$ s.t. $|y| = |x| - 1$ with IH: $\widehat{\delta'}(q, y) = \widehat{\delta}(q, h(y))$

proof of the claim

claim: $\widehat{\delta'}(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

- base case: $|x| = 0$ thus $x = \varepsilon$

$$\widehat{\delta'}(q, \varepsilon) = q = \widehat{\delta}(q, h(\varepsilon))$$

- step case: $|x| > 0$ thus $x = ya$ s.t. $|y| = |x| - 1$ with IH: $\widehat{\delta'}(q, y) = \widehat{\delta}(q, h(y))$

$$\widehat{\delta'}(q, ya) = \delta'(\widehat{\delta'}(q, y), a) \quad \text{(by definition of } \widehat{\delta'})$$

proof of the claim

claim: $\widehat{\delta'}(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

- base case: $|x| = 0$ thus $x = \varepsilon$

$$\widehat{\delta'}(q, \varepsilon) = q = \widehat{\delta}(q, h(\varepsilon))$$

- step case: $|x| > 0$ thus $x = ya$ s.t. $|y| = |x| - 1$ with IH: $\widehat{\delta'}(q, y) = \widehat{\delta}(q, h(y))$

$$\begin{aligned} \widehat{\delta'}(q, ya) &= \delta'(\widehat{\delta'}(q, y), a) && \text{(by definition of } \widehat{\delta'}) \\ &= \delta'(\widehat{\delta}(q, h(y)), a) && \text{(by induction hypothesis IH)} \end{aligned}$$

proof of the claim

claim: $\widehat{\delta'}(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

- base case: $|x| = 0$ thus $x = \varepsilon$

$$\widehat{\delta'}(q, \varepsilon) = q = \widehat{\delta}(q, h(\varepsilon))$$

- step case: $|x| > 0$ thus $x = ya$ s.t. $|y| = |x| - 1$ with IH: $\widehat{\delta'}(q, y) = \widehat{\delta}(q, h(y))$

$$\begin{aligned} \widehat{\delta'}(q, ya) &= \delta'(\widehat{\delta'}(q, y), a) && \text{(by definition of } \widehat{\delta'}) \\ &= \delta'(\widehat{\delta}(q, h(y)), a) && \text{(by induction hypothesis IH)} \\ &= \widehat{\delta}(\widehat{\delta}(q, h(y)), h(a)) && \text{(by definition of } \delta') \end{aligned}$$

proof of the claim

claim: $\widehat{\delta'}(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

- base case: $|x| = 0$ thus $x = \varepsilon$

$$\widehat{\delta'}(q, \varepsilon) = q = \widehat{\delta}(q, h(\varepsilon))$$

- step case: $|x| > 0$ thus $x = ya$ s.t. $|y| = |x| - 1$ with IH: $\widehat{\delta'}(q, y) = \widehat{\delta}(q, h(y))$

$$\begin{aligned}\widehat{\delta'}(q, ya) &= \delta'(\widehat{\delta'}(q, y), a) && \text{(by definition of } \widehat{\delta'}) \\ &= \delta'(\widehat{\delta}(q, h(y)), a) && \text{(by induction hypothesis IH)} \\ &= \widehat{\delta}(\widehat{\delta}(q, h(y)), h(a)) && \text{(by definition of } \delta') \\ &= \widehat{\delta}(q, h(y)h(a)) && \text{(by distributivity of } \widehat{\delta} - \text{w3.pdf, slide 10)}\end{aligned}$$

proof of the claim

claim: $\widehat{\delta'}(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

- base case: $|x| = 0$ thus $x = \varepsilon$

$$\widehat{\delta'}(q, \varepsilon) = q = \widehat{\delta}(q, h(\varepsilon))$$

- step case: $|x| > 0$ thus $x = ya$ s.t. $|y| = |x| - 1$ with IH: $\widehat{\delta'}(q, y) = \widehat{\delta}(q, h(y))$

$$\begin{aligned}\widehat{\delta'}(q, ya) &= \delta'(\widehat{\delta'}(q, y), a) && \text{(by definition of } \widehat{\delta'}) \\ &= \delta'(\widehat{\delta}(q, h(y)), a) && \text{(by induction hypothesis IH)} \\ &= \widehat{\delta}(\widehat{\delta}(q, h(y)), h(a)) && \text{(by definition of } \delta') \\ &= \widehat{\delta}(q, h(y)h(a)) && \text{(by distributivity of } \widehat{\delta} - \text{w3.pdf, slide 10)} \\ &= \widehat{\delta}(q, h(ya)) && \text{(by definition of homomorphism)}\end{aligned}$$

proof of the claim

claim: $\widehat{\delta}'(q, x) = \widehat{\delta}(q, h(x)) \quad \forall x \in \Sigma^* \quad \forall q \in Q$

- base case: $|x| = 0$ thus $x = \varepsilon$

$$\widehat{\delta}'(q, \varepsilon) = q = \widehat{\delta}(q, h(\varepsilon))$$

- step case: $|x| > 0$ thus $x = ya$ s.t. $|y| = |x| - 1$ with IH: $\widehat{\delta}'(q, y) = \widehat{\delta}(q, h(y))$

$$\begin{aligned}\widehat{\delta}'(q, ya) &= \delta'(\widehat{\delta}'(q, y), a) && \text{(by definition of } \widehat{\delta}') \\ &= \delta'(\widehat{\delta}(q, h(y)), a) && \text{(by induction hypothesis IH)} \\ &= \widehat{\delta}(\widehat{\delta}(q, h(y)), h(a)) && \text{(by definition of } \delta') \\ &= \widehat{\delta}(q, h(y)h(a)) && \text{(by distributivity of } \widehat{\delta} - \text{w3.pdf, slide 10)} \\ &= \widehat{\delta}(q, h(ya)) && \text{(by definition of homomorphism)} \\ &= \widehat{\delta}(q, h(x))\end{aligned}$$



Proof. (closedness under complement homomorphic preimage)

statement: $L(M') = h^{-1}(L(M))$

Proof. (closedness under complement homomorphic preimage)statement: $L(M') = h^{-1}(L(M))$

$$\forall x \in \Sigma^*, x \in L(M') \iff \widehat{\delta'}(s, x) \in F \quad \text{(by definition of acceptance)}$$

Proof. (closedness under complement homomorphic preimage)

statement: $L(M') = h^{-1}(L(M))$

$$\begin{array}{lll} \forall x \in \Sigma^*, x \in L(M') & \iff & \widehat{\delta'}(s, x) \in F & \text{(by definition of acceptance)} \\ & \iff & \widehat{\delta}(s, h(x)) \in F & \text{(by claim proven in slide 25)} \end{array}$$

Proof. (closedness under complement homomorphic preimage)

statement: $L(M') = h^{-1}(L(M))$

$$\begin{array}{lll}
 \forall x \in \Sigma^*, x \in L(M') & \iff & \widehat{\delta'}(s, x) \in F & \text{(by definition of acceptance)} \\
 & \iff & \widehat{\delta}(s, h(x)) \in F & \text{(by claim proven in slide 25)} \\
 & \iff & h(x) \in L(M) & \text{(by definition of acceptance)}
 \end{array}$$

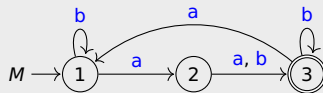
Proof. (closedness under complement homomorphic preimage)statement: $L(M') = h^{-1}(L(M))$

$$\begin{aligned}\forall x \in \Sigma^*, x \in L(M') &\iff \widehat{\delta'}(s, x) \in F && \text{(by definition of acceptance)} \\ &\iff \widehat{\delta}(s, h(x)) \in F && \text{(by claim proven in slide 25)} \\ &\iff h(x) \in L(M) && \text{(by definition of acceptance)} \\ &\iff x \in h^{-1}(L(M)) && \text{(by definition of homomorphic preimage)}\end{aligned}$$



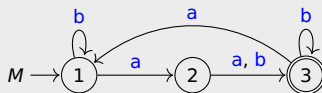
Example

- DFA M



Example

- DFA M

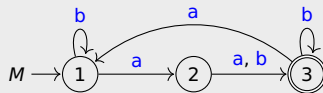


- homomorphism $h: \{a, b, c\}^* \rightarrow \{a, b\}^*$

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

Example

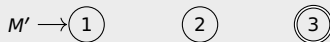
- DFA M



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

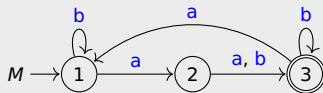
$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

- DFA M'



Example

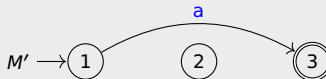
- DFA M



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

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

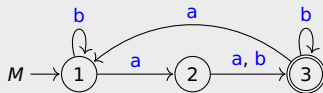
- DFA M'



$$\delta'(1, a) = \hat{\delta}(1, aa) = 3$$

Example

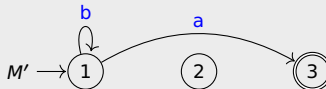
- DFA M



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

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

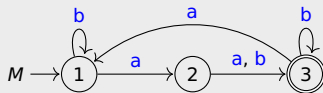
- DFA M'



$$\delta'(1, b) = \widehat{\delta}(1, \varepsilon) = 1$$

Example

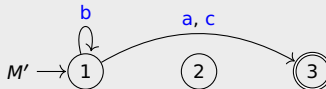
- DFA M



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

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

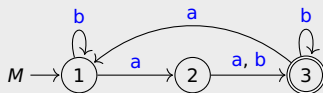
- DFA M'



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

Example

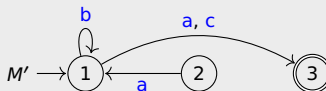
- DFA M



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

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

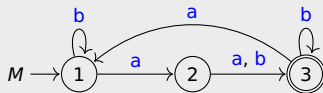
- DFA M'



$$\delta'(2, a) = \hat{\delta}(2, aa) = 1$$

Example

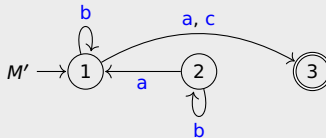
- DFA M



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

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

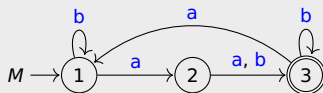
- DFA M'



$$\delta'(2, b) = \hat{\delta}(2, \varepsilon) = 2$$

Example

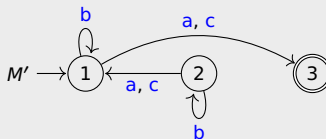
- DFA M



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

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

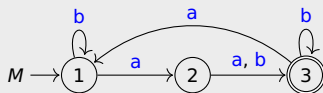
- DFA M'



$$\delta'(2, c) = \widehat{\delta}(2, bab) = 1$$

Example

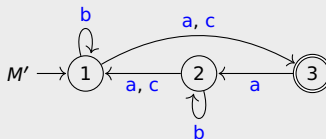
- DFA M



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

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

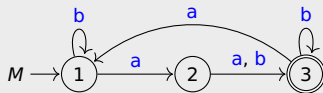
- DFA M'



$$\delta'(3, a) = \hat{\delta}(3, aa) = 2$$

Example

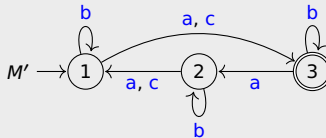
- DFA M



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

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

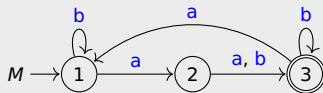
- DFA M'



$$\delta'(3, b) = \hat{\delta}(3, \varepsilon) = 3$$

Example

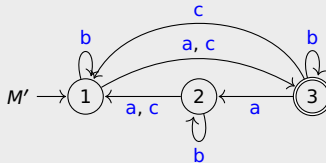
- DFA M



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

$$h(a) = aa \quad h(b) = \varepsilon \quad h(c) = bab$$

- DFA M'



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

Theorem

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

Theorem

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

Proof.

- regular expression α over Σ

Theorem

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

Proof.

- regular expression α over Σ
- homomorphism $h: \Sigma^* \rightarrow \Gamma^*$

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:

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:

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

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:

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

$$\boldsymbol{\epsilon'} = \boldsymbol{\epsilon}$$

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:

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

$$\epsilon' = \epsilon$$

$$\emptyset' = \emptyset$$

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:

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

$$\boldsymbol{\epsilon'} = \boldsymbol{\epsilon}$$

$$\boldsymbol{\emptyset'} = \boldsymbol{\emptyset}$$

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

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:

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

$$\epsilon' = \epsilon$$

$$\emptyset' = \emptyset$$

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

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

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:

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

$$\epsilon' = \epsilon$$

$$\emptyset' = \emptyset$$

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

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

$$(\beta^*)' = (\beta')^*$$

Definitions

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

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

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

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.

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

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.

$$\begin{aligned}
 D_k &= \{x \in (\{0, 1\} \times \{0, 1\})^* \mid x \text{ contains at most } k \text{ pairs } (0, 1) \text{ or } (1, 0)\} \quad \text{is regular} \\
 &= \{x \in (\{0, 1\} \times \{0, 1\})^* \mid H(\text{fst}(x), \text{snd}(x)) \leq k\} \\
 N_k(A) &= \text{fst}(\text{snd}^{-1}(A) \cap D_k)
 \end{aligned}$$

Example

- $A = \{0011\}$ $k = 2$

•

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

-

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

0	0	1	1
---	---	---	---

-

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

0	0	1	1
---	---	---	---

1	0	1	1
---	---	---	---

-

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

0	0	1	1
---	---	---	---

1	0	1	1
---	---	---	---

0	1	1	1
---	---	---	---

•

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

0	0	1	1
---	---	---	---

1	0	1	1
---	---	---	---

0	1	1	1
---	---	---	---

0	0	0	1
---	---	---	---

-

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

0	0	1	1
---	---	---	---

1	0	1	1
---	---	---	---

0	1	1	1
---	---	---	---

0	0	0	1
---	---	---	---

0	0	1	0
---	---	---	---

•

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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
---	---	---	---

-

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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																				

-

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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																

-

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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												

•

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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								

-

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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

•

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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	

- $snd^{-1}(A)$ 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		

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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	

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

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

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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	

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

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

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ 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		

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

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

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

0 0 1 1	1 0 1 1	0 1 1 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		

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

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

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

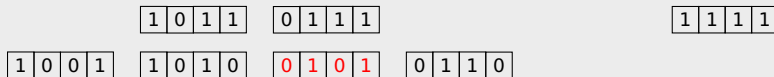
0 0 1 1	1 0 1 1	0 1 1 1		1 1 1 1
1 0 0 1	1 0 1 0	0 1 0 1	0 1 1 0	

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

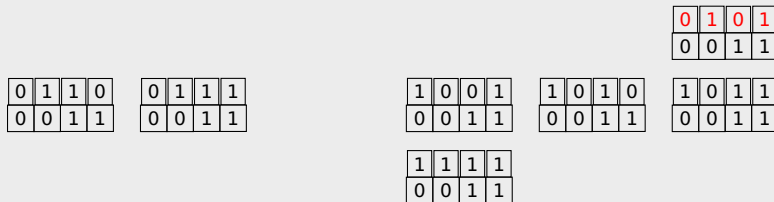
		0 0 1 1		0 1 0 1
		0 0 1 1		0 0 1 1
0 1 1 0	0 1 1 1		1 0 0 1	1 0 1 0
0 0 1 1	0 0 1 1		0 0 1 1	0 0 1 1
			1 1 1 1	
			0 0 1 1	

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

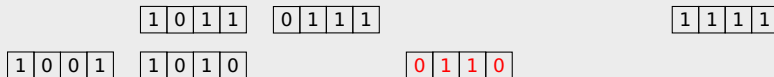


- $\text{fst}(\text{snd}^{-1}(A) \cap D_k) = N_k(A)$

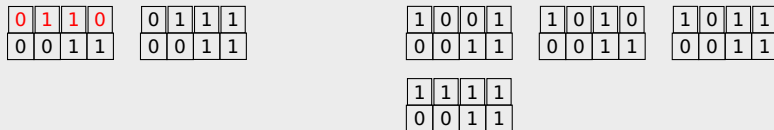


Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

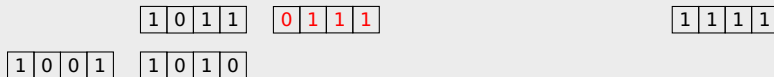


- $\text{fst}(\text{snd}^{-1}(A) \cap D_k) = N_k(A)$

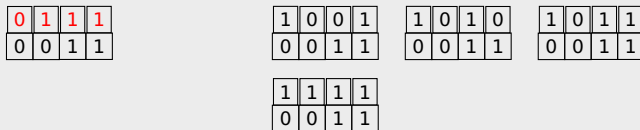


Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of



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



Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

1	0	1	1
---	---	---	---

1	0	0	1
---	---	---	---

1	0	1	0
---	---	---	---

1	1	1	1
---	---	---	---

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

1	0	0	1
0	0	1	1

1	0	1	0
0	0	1	1

1	0	1	1
0	0	1	1

1	1	1	1
0	0	1	1

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

1	0	1	1
---	---	---	---

1	1	1	1
---	---	---	---

1	0	1	0
---	---	---	---

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

1	0	1	0
0	0	1	1

1	0	1	1
0	0	1	1

1	1	1	1
0	0	1	1

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

1	0	1	1
---	---	---	---

1	1	1	1
---	---	---	---

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

1	0	1	1
0	0	1	1

1	1	1	1
0	0	1	1

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of

1	1	1	1
---	---	---	---

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

1	1	1	1
0	0	1	1

Example

- $A = \{0011\}$ $k = 2$
- $N_k(A)$ consists of
- $\text{fst}(\text{snd}^{-1}(A) \cap D_k) = N_k(A)$

Thanks! & Questions?