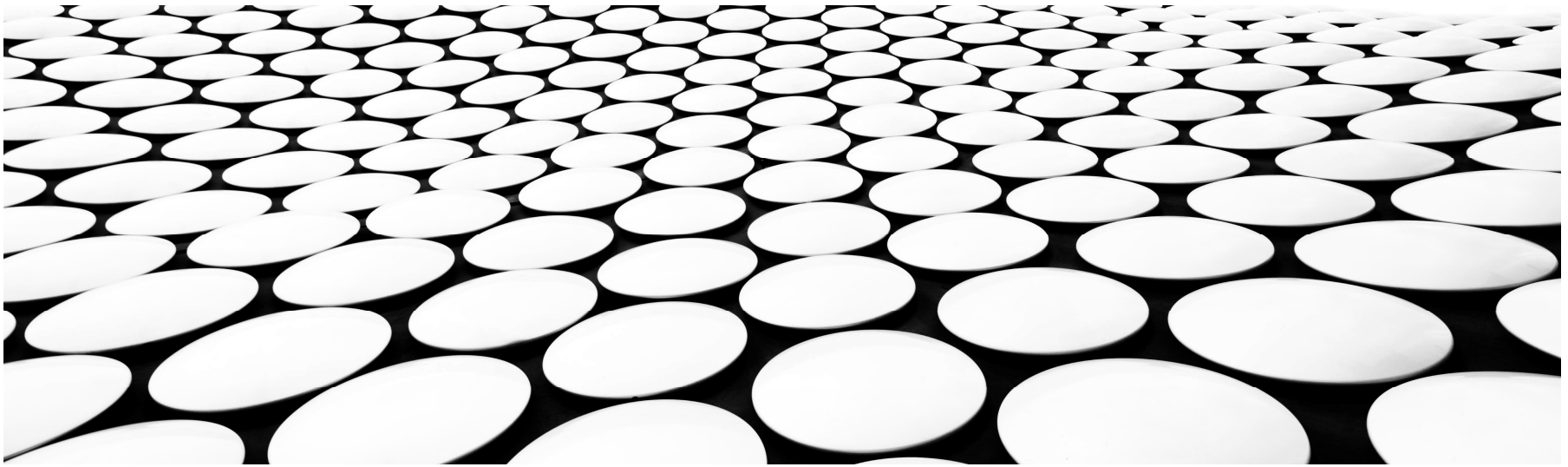
REGULAR EXPRESSIONS AND CONVERSIONS

IF 2124 TEORI BAHASA FORMAL OTOMATA



Regular expressions

A FA (NFA or DFA) is a "blueprint" for constructing a machine recognizing a regular language.

A *regular expression* is a "user-friendly," declarative way of describing a regular language.

$10^* = \{1, 10, 100, 1000, \cdots\}$

$+$

$01^* = \{0, 01, 011, 0111, \cdots\}$

Example: $\mathbf{01^* + 10^*}$ $= \{0, 1, 01, 10, 100, 011, \cdots$ dst $\}$.

Regular expressions are used in e.g.

1. UNIX `grep` command

2. UNIX Lex (Lexical analyzer generator) and Flex (Fast Lex) tools.

Operations on languages

Union:

$L \cup M = \{w : w \in L \text{ or } w \in M\}$

Concatenation:

$L.M = \{w : w = xy, x \in L, y \in M\}$

Powers:

$L^0 = \{\epsilon\},\ L^1 = L,\ L^{k+1} = L.L^k$

Kleene Closure:

$L^* = \bigcup_{i=0}^{\infty} L^i$ $\{\varepsilon, L, L^2, L^3, \ldots\}$

Question: What are $\emptyset^0$, $\emptyset^i$, and $\emptyset^*$ $\emptyset^0 = \{\varepsilon\}$ $\emptyset^i = \{\}^{i \neq 0}$ $\emptyset^* = \{\varepsilon\}$

Building regex's

Inductive definition of regex's:

Basis: $\boldsymbol{\epsilon}$ is a regex and $\boldsymbol{\emptyset}$ is a regex.
$L(\boldsymbol{\epsilon}) = \{\epsilon\}$, and $L(\boldsymbol{\emptyset}) = \emptyset$.

If $a \in \Sigma$, then $\boldsymbol{a}$ is a regex.
$L(\boldsymbol{a}) = \{a\}$.

Induction:

If E is a regex's, then (E) is a regex.
$L((E)) = L(E)$.

If E and F are regex's, then $E + F$ is a regex.
$L(E + F) = L(E) \cup L(F)$.

If E and F are regex's, then $E.F$ is a regex.
$L(E.F) = L(E).L(F)$.

If E is a regex's, then $E^\star$ is a regex.
$L(E^\star) = (L(E))^*$.

$w \subset \{0,1\}^* = \{\varepsilon, 0, 1, 00 \cdots$

Example: Regex for

$L = \{w \in \{0,1\}^* : \text{ 0 and 1 alternate in } w\}$

$(01)^* = \{\varepsilon, 01, 0101, 010101, \cdots$
$(10)^* = \{\varepsilon, 10, 1010, 101010, \cdots$
$0(10)^* = \{0, 010, 01010, \cdots$
$1(01)^* = \{1, 101, 10101, \cdots$

$$(01)^* + (10)^* + 0(10)^* + 1(01)^*$$

or, equivalently,

$$(\epsilon + 1)(01)^*(\epsilon + 0) \quad \varepsilon \Rightarrow \{\varepsilon, 0, 1, 01, 10, 1010, 010, \cdots$$

$\{\}$

Order of precedence for operators:

1. Star
2. Dot 01^*
3. Plus $+$, union.

$\emptyset = \{\} \neq \{\varepsilon\}$
$\emptyset^0 = \{\varepsilon\}$

Example: $01^* + 1$ is grouped $(0(1)^*) + 1$