THEORY OF AUTOMATA & FORMAL LANGUAGES

Kleene Star and Kleene Plus

Generate strings by repeating symbols or whole words

Powers → Star → Plus → Membership

Alphabet powers

First recall how strings of a fixed length are generated.

Construct powers of Σ

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

$\Sigma^0 = \{\varepsilon\}$

$\Sigma^1 = \{a, b\}$

$\Sigma^2 = \{aa, ab, ba, bb\}$

$\Sigma^3 = \{aaa, aab, aba, abb, baa, bab, bba, bbb\}$

Understand the notation

Σ^n contains all strings of length n.

Choose one symbol for each position.

For n = 2: $2 \times 2 = 4$ strings.

For n = 3: $2 \times 2 \times 2 = 8$ strings.

$|\Sigma^n| = 2^n$ for this alphabet.

Each power describes exactly one length.

Kleene star of an alphabet

Kleene star gathers strings of every possible finite length.

Build Σ*

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

Length 0: ε

Length 1: a, b

Length 2: aa, ab, ba, bb

Continue for every finite length.

Formal definition

$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \ldots$

$\Sigma^* = \{\varepsilon, a, b, aa, ab, ba, bb, \ldots\}$

$\varepsilon \in \Sigma^*$ because length 0 is allowed.

$abba \in \Sigma^*$ because all symbols are valid.

$ac \notin \Sigma^*$ because c is not in Σ.

Star means zero or more repetitions; it never includes an infinite string.

Kleene plus of an alphabet

Kleene plus gathers all nonempty strings over the alphabet.

Build Σ^+

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

Length 1: a, b

Length 2: aa, ab, ba, bb

Length 3: aaa, aab, ...

Continue for all positive lengths.

Compare with star

$\Sigma^+ = \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \ldots$

$\Sigma^+ = \Sigma^* - \{\varepsilon\}$

$\Sigma^* = \Sigma^+ \cup \{\varepsilon\}$

$ab \in \Sigma^+$ and $ab \in \Sigma^*$

$\varepsilon \notin \Sigma^+$ but $\varepsilon \in \Sigma^*$

For an alphabet, the only difference between star and plus is ε.

Powers of a language

Now repeat entire strings rather than individual alphabet symbols.

Example: L = {a, bb}

$L^0 = \{\varepsilon\}$

$L^1 = \{a, bb\}$

$a \cdot a = aa$

$a \cdot bb = abb$; $bb \cdot a = bba$

$bb \cdot bb = bbbb$

Build and interpret L^2

$L^2 = \{aa, abb, bba, bbbb\}$

Each result joins two members of L.

Length may vary: $|aa| = 2$.

$|bbbb| = 4$, although it uses two blocks.

L^3 includes $a \cdot bb \cdot bb = abbbb$.

The exponent counts blocks, not necessarily individual symbols.

Kleene star of a language

Use zero or more blocks from a language, in any order.

Example: L = {ab}

$L^0 = \{\varepsilon\}$

$L^1 = \{ab\}$

$L^2 = \{abab\}$

$L^3 = \{ababab\}$

$L^* = \{\varepsilon, ab, abab, ababab, \ldots\}$

Formal definition and testing

$L^* = L^0 \cup L^1 \cup L^2 \cup \ldots$

$L^* = \{(ab)^n : n \geq 0\}$

$abab \in L^*$ (two blocks)

$aba \notin L^*$ (incomplete block)

$ba \notin L^*$ (wrong block order)

Repeat the entire block ab; do not repeat a and b independently.

Kleene plus of a language

Use one or more blocks; a block itself may be the empty string.

Example: L = {ab}

$L^+ = L^1 \cup L^2 \cup L^3 \cup \ldots$

$L^+ = \{ab, abab, ababab, \ldots\}$

$\varepsilon \notin L^+$ in this example.

$L^* = L^+ \cup \{\varepsilon\}$

$L^+ = LL^* = L^*L$

Important exception

Let $L = \{\varepsilon, a\}$.

One block ε produces ε.

Two blocks $a \cdot a$ produce aa.

$L^+ = \{\varepsilon, a, aa, aaa, \ldots\}$

Here $L^+ = L^*$ because $\varepsilon \in L$.

Unlike alphabet plus, language plus can contain ε.

Membership by block decomposition

Decide whether a string can be formed by joining allowed blocks.

Example: L = {a, bb}

aabb = a · a · bb → yes

bba = bb · a → yes

bbbb = bb · bb → yes

ab = a · b → no: b is not a block.

ε uses zero blocks → in L*.

Boundary cases

If L = ∅, no positive block is possible.

$\emptyset^* = \{\varepsilon\}$; $\emptyset^+ = \emptyset$

If L = {ε}, every repetition is ε.

$\{\varepsilon\}^* = \{\varepsilon\}$; $\{\varepsilon\}^+ = \{\varepsilon\}$

For every L, $\varepsilon \in L^*$.

Check whole blocks and their order, not just the allowed symbols.

Practice and explained answers

Apply the definitions before moving to language operations.

Try these questions

1. List Σ^2 for $\Sigma = \{0, 1\}$.

2. Is ε in Σ^+?

3. For $L = \{ab\}$, list L^2.

4. Is aabb in $\{a, bb\}^+$?

5. Is ε in $\{\varepsilon, a\}^+$?

Answers and reasoning

1. $\{00, 01, 10, 11\}$

2. No: positive length is required.

3. $\{abab\}$, two copies of ab.

4. Yes: $a \cdot a \cdot bb$.

5. Yes: choose the single block ε.

Next: union, intersection, complement, concatenation and other language operations.