Closure of Regular Languages Under Kleene Star

Consider the following language

\[ A = \{ w \mid w \text{ is a string of 0s and 1s with any number of 0s and ending in a 1}\}. \]
NFA with a Single Accept State

**Theorem.** Prove that any NFA can be converted to an equivalent one that has a single accept state.

**Proof.**
Back to the Start

- Languages have been the center of our attention
  - What are languages again?
- So far we have described languages using English mostly
- Example
  - $A = \{ w \mid w$ is a string of 0s and 1s containing an even number of 1s $\}$
  - $B = \{ w \mid w$ is a string of as and bs containing the substring aba $\}$?
- What if there was a better way?
  - Regular expressions = Language for specifying languages
Regular Expressions
Regular Expressions in the Wild

- Some Applications
  - Text processing
  - Data validation.
    - Email or password format
  - Simple parsing (Ding ding - compilers!)
  - Syntax highlighting
  - `grep`
Definition

We recursively define regular expression over an alphabet $\Sigma$ as

1. $\varepsilon$
2. $\emptyset$
3. $a$ for some $a$ in the alphabet $\Sigma$,
4. $(R \cup S)$, if $R$ and $S$ are regular expressions,
5. $(R \circ S)$, if $R$ and $S$ are regular expressions,
6. $(R)^*$, if $R$ is a regular expression.
$L(R^*) = L(R)^*$

**Semantics**

What are the languages described by these expressions?

1. $L(\epsilon) = \{\epsilon\}$
2. $L(\emptyset) = \emptyset$
3. $L(a) = \{a\}$
4. $L(R \cup S) = L(R) \cup L(S)$
5. $L(R \cdot S) = L(R) \cdot L(S)$
6. $L(R^*) = L(R)^*$
Exercise

- Define regular expressions for the following languages. Consider $\Sigma = \{a, b\}$.
  - All strings containing no $b$’s
    $$a^*$$
  - All strings containing exactly one $b$
    $$a^*ba^*$$
  - All strings containing at least one $b$
    $$(a \cup b)^*b(a \cup b)^*$$
  - All strings containing the substring $aba$
    $$(a \cup b)^*aba(a \cup b)^*$$
  - All strings with length 3
    $$(a \cup b)(a \cup b)(a \cup b)$$
  - All strings that have even length
    $$((a \cup b)(a \cup b))^*$$
  - All strings that start and end with the same symbol
    $$a(a \cup b)^*a \cup a \cup b(a \cup b)^*b \cup b$$
NFAs and Regular expressions
NFAs and Regular expressions

- What is the language recognized by this NFA?
- Can you describe it as a regular expression? \((ba \cup bba)^*\)
Regular Expressions vs. Regular Languages

- Is there any relationship?

**Theorem.** A language is regular if and only if a regular expression describes it.
Given a regular expression \( R \), we want to show how to construct an NFA \( N \) such that \( L(R) = L(N) \).
Regular Expressions $\Rightarrow$ Regular Languages

**Theorem.** A language is regular if and only if a regular expression describes it.

**Proof.** By induction on the definition of regular expressions.

**Base case:**
1. If $a \in \Sigma$, then $a$ is regular.
2. $\varepsilon$ is regular.
3. $\emptyset$ is regular.

**Inductive step:**
4. If $R$ and $S$ are regular, then $(R \cup S)$ is regular.
5. If $R$ and $S$ are regular, then $(R \circ S)$ is regular.
6. If $R$ is regular, then $(R)^*$ is regular.
Proof in Action

Build an NFA that recognizes the regular expression \((ab \cup a)^*\)
Proof in Action

Build an NFA that recognizes the regular expression $a(a \cup b)^*a$