Context-free grammars

CENG 280

Context-free grammars

- Chomsky hierarchy
- Language generators
- Context-free grammars

Chomsky hierarchy

- Regular (Finite automaton)

Chomsky hierarchy

- Regular (Finite automaton)
- **Context-free (Non-deterministic pushdown automaton)**

Chomsky hierarchy

- Regular (Finite automaton)
- **Context-free (Non-deterministic pushdown automaton)**
- Recursively enumerable (Turing machine)

Language generators

Grammars are language generators.

- Start symbol

Language generators

Grammars are language generators.

- Start symbol
- The operation is not completely determined but limited to some rules.

Language generators

Grammars are language generators.

- Start symbol
- The operation is not completely determined but limited to some rules.
- When the process stops, it outputs a complete string $w \in \Sigma^\star$

Language generators

Grammars are language generators.

- Start symbol
- The operation is not completely determined but limited to some rules.
- When the process stops, it outputs a complete string $w \in \Sigma^{\star}$
- The language is the set of all strings that a grammar can generate.

Language generators

Grammars are language generators.

- Start symbol
- The operation is not completely determined but limited to some rules.
- When the process stops, it outputs a complete string $w \in \Sigma^\star$
- The language is the set of all strings that a grammar can generate.

Example

Consider $ba(ab \cup a)^\star b$

Language generators

Grammars are language generators.

- Start symbol
- The operation is not completely determined but limited to some rules.
- When the process stops, it outputs a complete string $w \in \Sigma^\star$
- The language is the set of all strings that a grammar can generate.

Example

Consider $ba(ab \cup a)^\star b$

Example

Formalize the "generator" concept with a grammar.

Context-free grammar

Definition (Context-free grammar)

Context-free grammar is a tuple $G = (V, \Sigma, R, S)$ where V is an alphabet, $\Sigma \subset V$ is the set of terminals, $R \subset (V \setminus \Sigma) \times V^\star$ is the finite set of rules, $S \in V \setminus \Sigma$ is the start symbol.

Context-free grammar

Definition (Context-free grammar)

Context-free grammar is a tuple $G = (V, \Sigma, R, S)$ where V is an alphabet, $\Sigma \subset V$ is the set of terminals, $R \subset (V \setminus \Sigma) \times V^{\star}$ is the finite set of rules, $S \in V \setminus \Sigma$ is the start symbol.

- $V \setminus \Sigma$ is the set of nonterminals

Context-free grammar

Definition (Context-free grammar)

Context-free grammar is a tuple $G = (V, \Sigma, R, S)$ where V is an alphabet, $\Sigma \subset V$ is the set of terminals, $R \subset (V \setminus \Sigma) \times V^\star$ is the finite set of rules, $S \in V \setminus \Sigma$ is the start symbol.

- $V \setminus \Sigma$ is the set of nonterminals
- We write $A \rightarrow_G u$ if $(A, u) \in R$.

Context-free grammar

Definition (Context-free grammar)

Context-free grammar is a tuple $G = (V, \Sigma, R, S)$ where V is an alphabet, $\Sigma \subset V$ is the set of terminals, $R \subset (V \setminus \Sigma) \times V^\star$ is the finite set of rules, $S \in V \setminus \Sigma$ is the start symbol.

- $V \setminus \Sigma$ is the set of nonterminals
- We write $A \rightarrow_G u$ if $(A, u) \in R$.
- For any two strings $u, v \in V^\star$, we write $u \Rightarrow_G v$ iff there exists $x, y, v' \in V^\star$, and $A \in V \setminus \Sigma$ such that $u = xAy$, $v = xv'y$, and $A \rightarrow_G v'$.

Context-free grammar

Definition (Context-free grammar)

Context-free grammar is a tuple $G = (V, \Sigma, R, S)$ where V is an alphabet, $\Sigma \subset V$ is the set of terminals, $R \subset (V \setminus \Sigma) \times V^\star$ is the finite set of rules, $S \in V \setminus \Sigma$ is the start symbol.

- $V \setminus \Sigma$ is the set of nonterminals
- We write $A \rightarrow_G u$ if $(A, u) \in R$.
- For any two strings $u, v \in V^\star$, we write $u \Rightarrow_G v$ iff there exists $x, y, v' \in V^\star$, and $A \in V \setminus \Sigma$ such that $u = xAy$, $v = xv'y$, and $A \rightarrow_G v'$.
- $\Rightarrow_G^\star$ is the reflexive transitive closure of $\Rightarrow_G$.

Context-free grammar

Definition (Context-free grammar)

Context-free grammar is a tuple $G = (V, \Sigma, R, S)$ where V is an alphabet, $\Sigma \subset V$ is the set of terminals, $R \subset (V \setminus \Sigma) \times V^\star$ is the finite set of rules, $S \in V \setminus \Sigma$ is the start symbol.

- $V \setminus \Sigma$ is the set of nonterminals
- We write $A \to_G u$ if $(A, u) \in R$.
- For any two strings $u, v \in V^\star$, we write $u \Rightarrow_G v$ iff there exists $x, y, v' \in V^\star$, and $A \in V \setminus \Sigma$ such that $u = xAy$, $v = xv'y$, and $A \to_G v'$.
- $\Rightarrow_G^\star$ is the reflexive transitive closure of $\Rightarrow_G$.
- $L(G) = \{w \in \Sigma^\star \mid S \Rightarrow_G^\star w\}$ is the language generated by G.

Context-free grammar

Definition (Context-free grammar)

Context-free grammar is a tuple $G = (V, \Sigma, R, S)$ where V is an alphabet, $\Sigma \subset V$ is the set of terminals, $R \subset (V \setminus \Sigma) \times V^\star$ is the finite set of rules, $S \in V \setminus \Sigma$ is the start symbol.

- $V \setminus \Sigma$ is the set of nonterminals
- We write $A \to_G u$ if $(A, u) \in R$.
- For any two strings $u, v \in V^\star$, we write $u \Rightarrow_G v$ iff there exists $x, y, v' \in V^\star$, and $A \in V \setminus \Sigma$ such that $u = xAy$, $v = xv'y$, and $A \to_G v'$.
- $\Rightarrow_G^\star$ is the reflexive transitive closure of $\Rightarrow_G$.
- $L(G) = \{w \in \Sigma^\star \mid S \Rightarrow_G^\star w\}$ is the language generated by G.
- $w_0 \Rightarrow_G w_1 \Rightarrow_G \ldots \Rightarrow_G w_n$: a derivation of w_n from w_0 in G, where $w_i \in V^\star$.

Context-free languages

Definition

A language L is context-free if $L = L(G)$ for a context-free grammar G.

Definition

A language L is context-free if $L = L(G)$ for a context-free grammar G.

Example

Write a derivation for "baaabb" (using the grammar rules for $\mathcal{L}(ba(ab \cup a)^{\star} b)$).

Context-free languages

Definition

A language L is context-free if $L = L(G)$ for a context-free grammar G.

Example

Write a derivation for "baaabb" (using the grammar rules for $\mathcal{L}(ba(ab \cup a)^\star b)$).

Example

$G = (V, \Sigma, R, S)$, $V = \{S, a, b\}$, $\Sigma = \{a, b\}$,

$$R : S \rightarrow aSb, S \rightarrow e$$

Context-free languages

Definition

A language L is context-free if $L = L(G)$ for a context-free grammar G.

Example

Write a derivation for "baaabb" (using the grammar rules for $\mathcal{L}(ba(ab \cup a)^\star b)$).

Example

$G = (V, \Sigma, R, S)$, $V = \{S, a, b\}$, $\Sigma = \{a, b\}$,

$$R : S \to aSb, S \to e$$

"Some context-free languages are not regular."

Context-free languages

Example

Show that regular language $L = \mathcal{L}(a^\star b)$ is context free.

Context-free languages

Example

Show that regular language $L = \mathcal{L}(a^\star b)$ is context free.

Example (Mathematical expressions)

$G = (V, \Sigma, R, E)$
$V = \{+, x, (,), id, T, F, E\}, \Sigma = \{+, x, (,), id, \}$,
$R = \{E \rightarrow E + T, T \rightarrow TxF, F \rightarrow (E), E \rightarrow T, T \rightarrow F, F \rightarrow id\}$

Context-free languages

Example

Show that regular language $L = \mathcal{L}(a^\star b)$ is context free.

Example (Mathematical expressions)

$G = (V, \Sigma, R, E)$
$V = \{+, x, (,), id, T, F, E\}, \Sigma = \{+, x, (,), id, \}$,
$R = \{E \to E + T, T \to TxF, F \to (E), E \to T, T \to F, F \to id\}$

Is it regular?

Context-free languages

Example

Show that regular language $L = \mathcal{L}(a^\star b)$ is context free.

Example (Mathematical expressions)

$G = (V, \Sigma, R, E)$
$V = \{+, x, (,), id, T, F, E\}, \Sigma = \{+, x, (,), id, \}$,
$R = \{E \rightarrow E + T, T \rightarrow TxF, F \rightarrow (E), E \rightarrow T, T \rightarrow F, F \rightarrow id\}$

Is it regular?

Example

Properly balanced left and right parenthesis: $V = \{S, (,)\}$, $\Sigma = \{(,)\}$.
$R = \{S \rightarrow ..., S \rightarrow ...., S \rightarrow ...$

Context-free languages

Example

Show that regular language $L = \mathcal{L}(a^\star b)$ is context free.

Example (Mathematical expressions)

$G = (V, \Sigma, R, E)$
$V = \{+, x, (,), id, T, F, E\}, \Sigma = \{+, x, (,), id, \}$,
$R = \{E \rightarrow E + T, T \rightarrow TxF, F \rightarrow (E), E \rightarrow T, T \rightarrow F, F \rightarrow id\}$

Is it regular?

Example

Properly balanced left and right parenthesis: $V = \{S, (,)\}$, $\Sigma = \{(,)\}$.
$R = \{S \rightarrow ..., S \rightarrow ...., S \rightarrow ...$

Example

Show that $L = \{ww^R : w \in \{a, b\}^\star\}$ is CFL.

Context-free languages

Theorem

All regular languages are context-free.

Context-free languages

Theorem

All regular languages are context-free.

Proof by direct construction: Let regular language L be accepted by $M = (K, \Sigma, \delta, s, F)$. The same language is generated by the grammar $G = (V, \Sigma, R, S)$,

- $V = K \cup \Sigma$
- $S = s$
- $R = \{q \to ap \mid \delta(q, a) = p\} \cup \{q \to e \mid q \in F\}$

Context-free languages

Theorem

All regular languages are context-free.

Proof by direct construction: Let regular language L be accepted by $M = (K, \Sigma, \delta, s, F)$. The same language is generated by the grammar $G = (V, \Sigma, R, S)$,

- $V = K \cup \Sigma$
- $S = s$
- $R = \{q \rightarrow ap \mid \delta(q, a) = p\} \cup \{q \rightarrow e \mid q \in F\}$

Ex: Write a DFA, and then a grammar for $\mathcal{L}(a^\star b)$

Theorem

All regular languages are context-free.

Proof by direct construction: Let regular language L be accepted by $M = (K, \Sigma, \delta, s, F)$. The same language is generated by the grammar $G = (V, \Sigma, R, S)$,

- $V = K \cup \Sigma$
- $S = s$
- $R = \{q \rightarrow ap \mid \delta(q, a) = p\} \cup \{q \rightarrow e \mid q \in F\}$

Ex: Write a DFA, and then a grammar for $\mathcal{L}(a^\star b)$

Other proof methods:

- CFL are accepted by pushdown automata, which is a generalization of FA
- CFL are closed under union, concatenation, and Kleene star. $\emptyset$ and $\{a\}$ are context-free, hence CFL has to contain RL.