Review for CMSC 452 Final: Grammars

Context Free Languages

Context Free Grammar for $\{a^m b^n : m > n\}$

Context Free Grammar for $\{a^m b^n : m > n\}$

$$S \rightarrow AT$$
$$T \rightarrow aTb$$
$$T \rightarrow e$$
$$A \rightarrow Aa$$
$$A \rightarrow a$$

Context Free Grammars

Def A **Context Free Grammar** is a tuple $G = (N, \Sigma, R, S)$

- N is a finite set of **nonterminals**.
- Σ is a finite **alphabet**. Note $\Sigma \cap N = \emptyset$.
- $R \subseteq N \times (N \cup \Sigma)^*$ and are called **Rules**.
- $S \in N$, the **start symbol**.

L(G)

If A is non-terminal then the CFG gives us gives us rules like:

- $A \to AB$
- $A \to a$

L(G)

If A is non-terminal then the CFG gives us gives us rules like:

- $A \rightarrow AB$
- $A \rightarrow a$

For any string of **terminals and non-terminals** α, $A \Rightarrow \alpha$ means that, starting from A, some combination of the rules produces α.

L(G)

If A is non-terminal then the CFG gives us gives us rules like:

- $A \rightarrow AB$
- $A \rightarrow a$

For any string of **terminals and non-terminals** α, $A \Rightarrow \alpha$ means that, starting from A, some combination of the rules produces α.

Examples:

- $A \Rightarrow a$
- $A \Rightarrow aB$

L(G)

If A is non-terminal then the CFG gives us gives us rules like:

- $A \rightarrow AB$
- $A \rightarrow a$

For any string of **terminals and non-terminals** α, $A \Rightarrow \alpha$ means that, starting from A, some combination of the rules produces α.

Examples:

- $A \Rightarrow a$
- $A \Rightarrow aB$

Then we define $L(G)$ by:

$$L(G) = \{w \in \Sigma^* \mid S \Rightarrow w\}$$

Example of a Lang that is NOT a CFL

1) $\{a^n b^n c^n : n \in \mathbb{N}\}$ is NOT a CFL.

Example of a Lang that is NOT a CFL

1) $\{a^n b^n c^n : n \in \mathbb{N}\}$ is NOT a CFL.

2) $\{a^{n^2} : n \in \mathbb{N}\}$ is NOT a CFL.

Example of a Lang that is NOT a CFL

1) $\{a^n b^n c^n : n \in \mathbb{N}\}$ is NOT a CFL.

2) $\{a^{n^2} : n \in \mathbb{N}\}$ is NOT a CFL.

3) If $L \subseteq a^*$ and L is not regular than L is not a CFL.

Example of a Lang that is NOT a CFL

1) $\{a^n b^n c^n : n \in \mathbb{N}\}$ is NOT a CFL.

2) $\{a^{n^2} : n \in \mathbb{N}\}$ is NOT a CFL.

3) If $L \subseteq a^*$ and L is not regular than L is not a CFL.

One proves theorems NON CFL using the PL for CFL's (we omit)

Closure Properties and REG$\subset$ CFL

Closure Properties.

Closure Properties.

1. CFL's closed under UNION-easy.

Closure Properties.

1. CFL's closed under UNION-easy.
2. CFL's closed under CONCAT-easy.

Closure Properties.

1. CFL's closed under UNION-easy.
2. CFL's closed under CONCAT-easy.
3. CFL's closed under *-easy.

Closure Properties.

1. CFL's closed under UNION-easy.
2. CFL's closed under CONCAT-easy.
3. CFL's closed under *-easy.
4. CFL's closed under INTER-FALSE:
   $a^n b^n c^* \cap a^* b^n c^n = a^n b^n c^n$.

Closure Properties.

1. CFL's closed under UNION-easy.
2. CFL's closed under CONCAT-easy.
3. CFL's closed under *-easy.
4. CFL's closed under INTER-FALSE:
   $a^n b^n c^* \cap a^* b^n c^n = a^n b^n c^n$.
5. CFL's closed under COMPLEMENTATION: FALSE: $\overline{a^n b^n c^n}$ is CFL.

REG contained in CFL

For every **regex** α, $L(\alpha)$ is a CFL.

REG contained in CFL

For every **regex** α, $L(\alpha)$ is a CFL.

Prove by ind on the length of α.

REG contained in CFL

For every **regex** α, $L(\alpha)$ is a CFL.

Prove by ind on the length of α.

We omit from this review.

Examples of CFL's and Size of CFG's

Chomsky Normal Form

Def CFG G is in **Chomsky Normal Form** if the rules are all of the following form:

Chomsky Normal Form

Def CFG G is in **Chomsky Normal Form** if the rules are all of the following form:
1) $A \to BC$ where $A, B, C \in N$ (nonterminals).

Chomsky Normal Form

Def CFG G is in **Chomsky Normal Form** if the rules are all of the following form:
1) $A \rightarrow BC$ where $A, B, C \in N$ (nonterminals).
2) $A \rightarrow \sigma$ (where $A \in N$ and $\sigma \in \Sigma$).

Chomsky Normal Form

Def CFG G is in **Chomsky Normal Form** if the rules are all of the following form:
1) $A \rightarrow BC$ where $A, B, C \in N$ (nonterminals).
2) $A \rightarrow \sigma$ (where $A \in N$ and $\sigma \in \Sigma$).
3) $S \rightarrow e$ (where S is the start state).

Languges for Singleton Sets

All CFL's are in Chomsky Normal Form. $|G|$ is the number of rules.

Langues for Singleton Sets

All CFL's are in Chomsky Normal Form. $|G|$ is the number of rules.

1. $(\forall w \in \{a, b\}^n)(\exists G)[L(G) = \{w\} \wedge |G| = n]$.

Langues for Singleton Sets

All CFL's are in Chomsky Normal Form. $|G|$ is the number of rules.

1. $(\forall w \in \{a,b\}^n)(\exists G)[L(G) = \{w\} \wedge |G| = n]$.
2. $(\forall n \in \mathbb{N})(\exists G)[L(G) = \{a^n\} \wedge |G| = O(\log n)]$.

Langeuges for Singleton Sets

All CFL's are in Chomsky Normal Form. $|G|$ is the number of rules.

1. $(\forall w \in \{a, b\}^n)(\exists G)[L(G) = \{w\} \wedge |G| = n]$.
2. $(\forall n \in \mathbb{N})(\exists G)[L(G) = \{a^n\} \wedge |G| = O(\log n)]$.
3. Does there exist a string w such that for all G such that $L(G) = \{w\}$, $|G| = \Omega(n)$

Languges for Singleton Sets

All CFL's are in Chomsky Normal Form. $|G|$ is the number of rules.

1. $(\forall w \in \{a, b\}^n)(\exists G)[L(G) = \{w\} \wedge |G| = n]$.
2. $(\forall n \in \mathbb{N})(\exists G)[L(G) = \{a^n\} \wedge |G| = O(\log n)]$.
3. Does there exist a string w such that for all G such that $L(G) = \{w\}$, $|G| = \Omega(n)$
   Yes- can show such strings exist by counting number-of-grammars and number-of-strings.

Poly Time Algorithm for CFG Membership

Let L be a CFL. Let G be the Chomsky Normal Form CFG for L.

Poly Time Algorithm for CFG Membership

Let L be a CFL. Let G be the Chomsky Normal Form CFG for L.

$w = \sigma_1 \cdots \sigma_n$.

Poly Time Algorithm for CFG Membership

Let L be a CFL. Let G be the Chomsky Normal Form CFG for L.

$w = \sigma_1 \cdots \sigma_n$.
We want to know if $w \in L$. We assume $w \neq e$.

Poly Time Algorithm for CFG Membership

Let L be a CFL. Let G be the Chomsky Normal Form CFG for L.

$w = \sigma_1 \cdots \sigma_n$.
We want to know if $w \in L$. We assume $w \neq e$.

For $i \leq j$ let

$$\mathrm{GEN}[i,j] = \{A : A \Rightarrow \sigma_i \cdots \sigma_j\}$$

Poly Time Algorithm for CFG Membership

Let L be a CFL. Let G be the Chomsky Normal Form CFG for L.

$w = \sigma_1 \cdots \sigma_n$.
We want to know if $w \in L$. We assume $w \neq e$.

For $i \leq j$ let

$$\mathrm{GEN}[i,j] = \{A : A \Rightarrow \sigma_i \cdots \sigma_j\}$$

Find $\mathrm{GEN}[i,j]$ with Dynamic Programming.