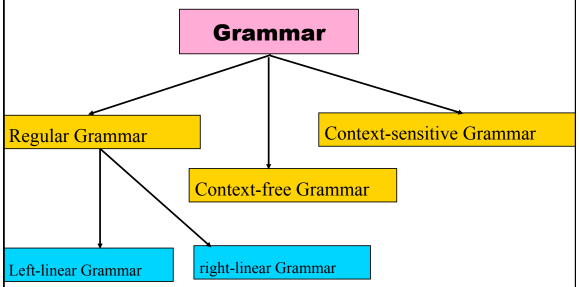
Automata and Languages

Prof. Mohamed Hamada

Software Engineering Lab.
The University of Aizu
Japan

Grammar



Today's Topics

- Grammars
- Right-Linear Grammars
- Left-Linear Grammars
- Regular Grammars
- Context-free Grammars
- Derivation: Leftmost & Rightmost
- Derivation Tree

Grammars

Definition

A grammar G is defined as G = (V, T, P, S) where:

- V : Finite set of variables/non-terminals
  (We use capital letters A,B,C,... for variables)
- T : Alphabet/Finite set of terminals
  (We use small letters a,b,c,... for terminals)
- P : Finite set of rules/productions
- S : Start symbol

$$S \in V$$
$$V \cap T = \phi$$
$$\text{Rule}: \quad \alpha \rightarrow \beta$$
$$\alpha \in (V \cup T)^{+}, \quad \beta \in (V \cup T)^{*}$$

Each grammar G defines a language L(G), which is the set of strings in T* (=Σ*) that G can generate from S.
It is all about the production rules.

Grammars

Definition

Given a grammar G = (V, T, P, S)

For a string $w=uxv$ we can apply the production rule $x \rightarrow y$ to w so we get a string $z=uyv$.
In this case we write $w \longrightarrow z$, which reads w drives z.

If $w_1 \longrightarrow w_2 \longrightarrow \dots \longrightarrow w_n$,
we say that w_1 drives w_n and we write $w_1 \longrightarrow^* w_n$

5

Grammars

Example

Given a grammar G = (V, T ,P,S)

V={A, B, C}
T={a, b, x}
S= A
And P is:

AaBx→aBAaBb
CaBx→aBAaCb
ABC→λ

6

Grammars

Definition

Let G=(V, T, P, S) be a grammar.

- $w \in (V \cup T)^*$ is a *sentential form*, if $S \Rightarrow_G^* w$
- $w \in T^*$ is a *sentence*, if $S \Rightarrow_G^* w$
- The *language of G*,
  L(G) = $\{w \in T^* \mid S \Rightarrow_G^* w\}$

7

Grammars

Some Remarks

The language L(G) = { w∈T* : S ⇒* w } contains only strings of terminals, not variables.

Notation: We summarize several rules for one variable:
A → B
A → 01 by A → B | 01 | AA
A → AA

8

Grammars

Example

Given the grammar:

$$G = \{\{S\}, \{a,b\}, \{S \to \lambda, S \to aSb\}, S)$$

The language generated by this grammar is:

$$L(G) = \{a^n b^n \mid n \geq 0\}$$

9

Right-Linear Grammars

Definition

A Grammar G= **(V, T ,P,S)** is called ***right-linear grammar*** if every production is of the form $A \to xB$, or $A \to x$ where

A,B ∈ V, x ∈ **T***

Example: The grammar
x → 0*x* | 1*y*
y → 0*x* | 1*z*
z → 0*x* | 1*z* | λ
Is a right-linear grammar.

10

Left-Linear Grammars

Definition

A Grammar G= **(V, T ,P,S)** is called ***left-linear grammar*** if every production is of the form $A \to Bx$, or $A \to x$ where

A,B ∈ V, x ∈ **T***

Example: The grammar
x → *x0* | *y1*
y → *x0* | *z1*
z → *x0* | *z1* | λ
Is a left-linear grammar.

11

Regular Grammars

Definition

A Grammar G= **(V, T ,P,S)** is called ***regular grammar*** if its is left- or right-linear

Example: The grammar
x → *x0* | *y1*
y → *x0* | *z1*
z → *x0* | *z1* | λ
Is a left-linear grammar,
hence is Regular Grammar

Example: The grammar
x → 0*x* | 1*y*
y → 0*x* | 1*z*
z → 0*x* | 1*z* | λ
Is a right-linear grammar,
Hence is Regular Grammar

12

Grammars

Example 1

Write a grammar that generate the language:

$L = \{w \in \{a,b\}^* \mid \text{length}(w) \text{ is EVEN}\}$

$$E \rightarrow \lambda \mid aaE \mid abE \mid baE \mid bbE$$

$$E \rightarrow \lambda \mid aO \mid bO$$
$$O \rightarrow aE \mid bE$$

Example 2

Write a grammar that generate the language:

$L = \{w \in \{a,b\}^* \mid w \text{ has EVEN number of } b\text{'s}\}$

$$E \rightarrow \lambda \mid aE \mid bO$$
$$O \rightarrow aO \mid bE$$

13

Grammars

Example 3

Write a grammar that generate the language:

$L = \{w \in \{a,b,c\}^* \mid w \text{ does not contain } abc\}$

$$\langle reset\rangle \rightarrow b\langle reset\rangle \mid c\langle reset\rangle \mid a\langle seenA\rangle \mid \lambda$$
$$\langle seenA\rangle \rightarrow a\langle seenA\rangle \mid c\langle reset\rangle \mid b\langle seenAB\rangle \mid \lambda$$
$$\langle seenAB\rangle \rightarrow a\langle seenA\rangle \mid b\langle reset\rangle \mid \lambda$$

14

Context-free Grammars

15

Context-free Grammars

Definition

CFG = (V, T, P, S)

- V : Finite set of variables/non-terminals
- T : Alphabet/Finite set of terminals
- P : Finite set of rules/productions
- S : Start symbol

$S \in V$

$V \cap T = \phi$

$Rule: \quad A \rightarrow \omega$

$A \in V \quad \omega \in (V \cup T)^*$

16

Context-free Grammars

Definition

- *Context-freeness:* An *A*-rule can be applied whenever *A* occurs in a string, irrespective of the context (that is, non-terminals and terminals around *A*).

17

Context-free Grammars

Derivation

- One-step Derivation

$$uAv \underset{A \to \omega}{\Rightarrow} u\omega v$$

- *w* is derivable from *v* in CFG, if there is a finite sequence of rule applications such that:

$$v \Rightarrow w_1 \Rightarrow \ldots \Rightarrow w_n = w$$

In this case we can write this derivation as $v \xrightarrow{*} w$

18

Context-free Grammars

Derivation

The derivation as $v \xrightarrow{*} w$ **is called:**

Leftmost derivation: if in every step the leftmost variable is selected for reduction

Rightmost derivation: if in every step the rightmost variable is selected for reduction

19

Context-free Grammars

Example 1

Let G = ({S}, {a,b},S,P) with for P:
•S→aSa, and S→bSb, and S→λ.
•Some *derivations* from this grammar:
• S ⇒ aSa ⇒ aaSaa ⇒ aabSbaa ⇒ aabbaa
• S ⇒ bSb ⇒ baSab ⇒ baab, and so on.

•In general S ⇒....⇒ ww^R for w∈{a,b}*.

L(G)={ww^R : w∈{a,b}*}

20

Context-free Grammars

Example 2

$$G = (\{S, A, B\}, \{a, b\}, \{S \to AB, A \to aA \mid \lambda, B \to Bb \mid \lambda\}, S)$$

$$L(G) = L(a^*b^*)$$

Leftmost Derivation:

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow aB \Rightarrow aBb \Rightarrow ab$$

Rightmost Derivation:

$$S \Rightarrow AB \Rightarrow ABb \Rightarrow Ab \Rightarrow aAb \Rightarrow ab$$

21

Context-free Grammars

Example 3

Take the CFG S → 0 | 1 | ¬(S) | (S)∨(S) | (S)∧(S), which generates all proper Boolean formulas that use "0", "1", "¬", "∨", "∧", "(" and ")".

Then "(0)∨((0)∧(1))" can be derived in the following ways…

[leftmost] S ⇒ (S)∨(S) ⇒ (0)∨(S) ⇒ (0)∨((S)∧(S)) ⇒ (0)∨((0)∧(S)) ⇒ (0)∨((0)∧(1))

[rightmost] S ⇒ (S)∨(S) ⇒ (S)∨((S)∧(S)) ⇒ (S)∨((S)∧(1)) ⇒ (S)∨((0)∧(1)) ⇒ (0)∨((0)∧(1))

[something else] S ⇒ (S)∨(S) ⇒ (0)∨(S) ⇒ (0)∨((S)∧(S)) ⇒ (0)∨((S)∧(1)) ⇒ (0)∨((0)∧(0))

22

Context-free Grammars

Example 4

Consider the CFG:

$$G = \{\{S\}, \{a, b\}, \{S \to \lambda, S \to aSb\}, S)$$

- Derivation of *aabb* is

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aabb$$

23

Context-free Grammars

Example 5

Consider the CFG G:

$$S \to aSa \mid aBa$$

$$B \to bB \mid b$$

$$L(B) = \{b^m \mid m > 0\}$$

$$L(S) = \{a^n b^m a^n \mid n > 0 \wedge m > 0\}$$

L(G)= L(S)

24

Context-free Grammars

Example 6

Consider the CFG G_1:

$$S \to aSa \mid B$$
$$B \to bB \mid \lambda$$

The language generated by G_1 is:

$$L(G_1) = \{a^n b^m a^n \mid n \geq 0 \wedge m \geq 0\}$$

Consider the CFG G_2:

$$S \to abSc \mid \lambda$$

The language generated by G_2 is:

$$L(G_2) = \{(ab)^n c^n \mid n \geq 0\}$$

25

Context-free Grammars

Example 7

Consider the CFGs G_1 and G_2 :

G_1 :

$$S \to AB$$
$$A \to aA \mid a$$
$$B \to bB \mid \lambda$$

G_2 :

$$S \to aS \mid aB$$
$$B \to bB \mid \lambda$$

The language generated by G_1 and G_2 is:

$L(G_1) = L(G_2) = L(S)$

$$L(S) = \{a^n b^m \mid m \geq 0 \wedge n > 0\}$$
$$L(S) = L(a^+ b^*)$$

26

Context-free Grammars

Example 8

Write a CFG to generate the language: $a^* ba^* ba^*$

$$S \to AbAbA$$
$$A \to aA \mid \lambda$$

$$S \to aS \mid B$$
$$B \to bA$$
$$A \to aA \mid bC$$
$$C \to aC \mid \lambda$$

Left to right generation of string.

27

Context-free Grammars

Exercise 1

WRITE A CFG FOR THE EMPTY SET

G = { {S}, Σ, ∅, S }

28

Context-free Grammars

Exercise 2

What is the CFG ({S},{(,)}, P, S) that produces the language of correct parentheses like (), (()), or ()(())?

S→ (S)|SS|λ

Context-free Grammars

Example

Consider the CFG G=({S,Z},{0,1}, P, S) with
P: S → 0S1 | 0Z1
Z → 0Z | λ
What is the language generated by G?

Answer: L(G) = $\{0^i1^j \mid i \geq j\}$

Specifically, S yields the $0^{j+k}1^j$ according to:
$S \Rightarrow 0S1 \Rightarrow \ldots \Rightarrow 0^jS1^j \Rightarrow$
$0^jZ1^j \Rightarrow 0^j0Z1^j \Rightarrow \ldots \Rightarrow 0^{j+k}Z1^j \Rightarrow 0^{j+k}\varepsilon 1^j = 0^{j+k}1^j$

Context-free Grammars

Exercise

- Can you make Context Free Grammars for the following?

a) $\{ 0^n1^n : n \geq 0\}$
b) $\{ 0^n1^m : n,m \geq 0\}$
c) Arithmetic a,b,c formulas like a+b×c+a (without ())

- Answers:

a) S → 0S1 | λ
b) S → 0S | R and R → 1R | λ
c) S → a | b | c | S+S | S×S

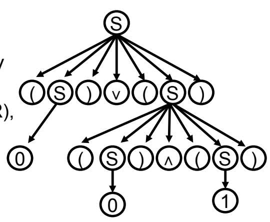
Context-free Grammars

Derivation Tree

For a CFG G=(V,T,S,P) a derivation tree has the following properties:

Example

1) The root is labeled S
2) Each leaf is from T∪{λ}
3) Each interior node is from V
4) If node has label A∈V and its children $a_1 \ldots a_n$ (from L to R), then P must have the rule $A \rightarrow a_1 \ldots a_n$ (with $a_j \in V \cup T \cup \{\lambda\}$)
5) A leaf labeled λ is a single child (has no siblings).



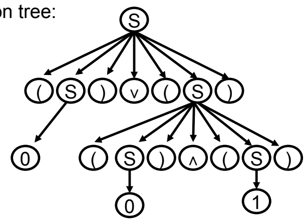
For **partial derivation trees** we have:
2a) Each leaf is from V∪T∪{λ}

Context-free Grammars

Derivation Tree: Example

Take the CFG S → 0 | 1 | ¬(S) | (S)∨(S) | (S)∧(S), which generates all proper Boolean formulas that use “0”, “1”, “¬”, “∨”, “∧”, “(“ and “)”.

The derivation S ⇒* (0)∨((0)∧(1)) can be expressed by the following derivation tree:



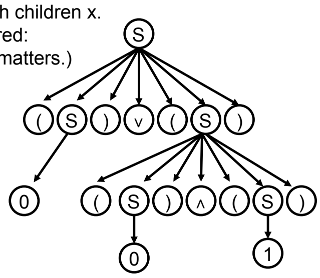
Context-free Grammars

Derivation Tree: Notes

Application of a production rule A → x is represented by node A with children x. (Note that the tree is ordered: the ordering of the nodes matters.)

The root has variable S.

The **yield** of S is expressed by the leaves of the tree.



Context-free Grammars

Derivation Tree: Notes

- Looking at a tree you see the derivation without the unnecessary information about its order.
- **Theorem**: Let G be a CFG. We have w∈L(G) if and only if there exists a derivation tree of G with yield w.
- Also, y is a sentential form of G if and only if there exists a partial derivation tree for G.
- Remember: the root always has to be S.

Context-free Grammars

Example 1

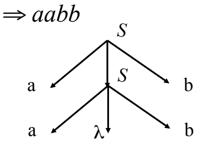
Consider the CFG G:

$$G = \{\{S\},\{a,b\},\{S \rightarrow \lambda, S \rightarrow aSb\},S)$$

- The derivation of *aabb* is:

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aabb$$

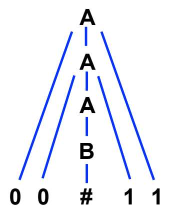
- Derivation tree is



Context-free Grammars

Example 2

A ⇒ 0A1 ⇒ 00A11 ⇒ 00B11 ⇒ 00#11



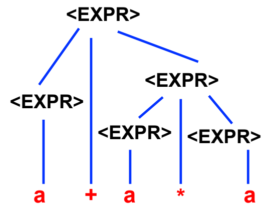
37

Context-free Grammars

Example 3

<EXPR> → <EXPR> + <EXPR>
<EXPR> → <EXPR> * <EXPR>
<EXPR> → (<EXPR>)
<EXPR> → a

Build a parse tree for a + a * a



38

Context-free Grammars

Exercise

WRITE A CFG FOR EVEN-LENGTH PALINDROMES over **Σ**={a,b,c,d} ?

S → aSa for all a ∈ Σ
S → λ

39