Chapter 5: Context-Free Languages

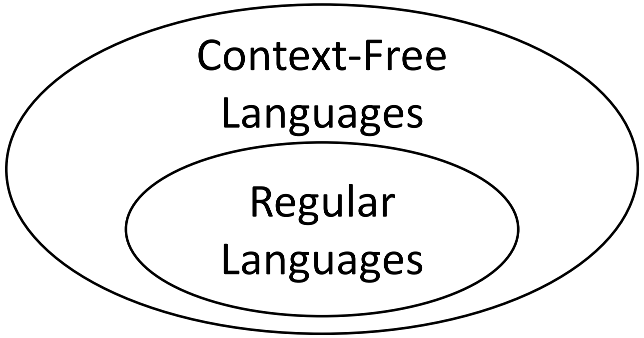
Vocabulary

- grammar
- context-free
- regular language
- derivation
- parsing
- Syntax
- Semantics
- linear grammar
- left linear grammar
- right linear grammar
- sentential form
- yield of the tree
- positive closure

Context-Free Languages (CFLs)

- If a language is regular, then we know there's a CFG for that language which means that every regular language is also a context-free.
- CFGs exists for languages known to not be regular.

$\text{RL} \subset CFL$



- CFLs describe features that are important in the design of programming languages (***syntax***) but not **semantics**.

Grammars

- Grammars are used to describe languages
- $G = (V, T, S, P)$

  V – finite set of variables

  T – Finite set of terminal symbols

  S - the start variable

  P - finite set of productions

Grammars

- Productions are given in the form
$$x \to y$$
- In general, the left side of the production indicates the symbols and/or terminals that can be replaced by the right handed side of the production.
- All the strings that can be derived from the productions of a grammar defines the language of the grammar.
- For grammar G, the language produced from the grammar is denoted $L(G)$.

Context-Free Grammars (CFGs)

A grammar is context-free given the definition
$G = (V, T, S, P)$

V – Finite set of variables

T – Finite set of terminal symbols

S – Start variable

P – Finite set of productions

All proudctions of the form: $A \rightarrow x$ where

$A \in V$ and

$x \in (V \cup T)^*$.

A language is context-free iff there exist a CFG such that $L = L(G)$.

CFGs

So what does “context” mean?

abSbbbaSabb

bSb - > baabbabab

aSa -> abbbababa

Example 1: Is $L = \{a^n b^n : n \geq 0\}$ a CFL?

$G = (\{S\}, \{a, b\}, S, P)$

$$S \to aSb$$
$$S \to \lambda$$

Does the grammar produce all strings in the language above and only strings apart of that language?

Yes.

Is the grammar context-free?

Yes.

THE LANGUAGE $L(G)$ IS A CONTEXT-FREE LANGUAGE.

Example 2: Is $L = \{ww^r : w \in \{a,b\}^*\}$ a CFL?

$G = (\{S\}, \{a,b\}, S, P)$

$$S \to aSa$$
$$S \to bSb$$
$$S \to \lambda$$

Does the grammar produce all strings in the language above and only strings apart of that language?

Yes.

Is the grammar context-free?

Yes.

THE LANGUAGE $L(G)$ IS A CONTEXT-FREE LANGUAGE.

Example 3: Is $L = \{a^n b^m : n \neq m\}$ a CFL?

Two conditions

$\underline{n < m}$

$S \rightarrow Bb$

$B \rightarrow Bb$

$B \rightarrow aBb$

$B \rightarrow \lambda$

$\underline{n > m}$

$S \rightarrow aA$

$A \rightarrow aA$

$A \rightarrow aAb$

$A \rightarrow \lambda$

Example 3: Is $L = \{a^n b^m : n \neq m\}$ a CFL?

<u>Two conditions</u>

$\underline{n < m}$

$S \rightarrow Bb$

$B \rightarrow Bb$

$B \rightarrow aBb$

$B \rightarrow \lambda$

$\underline{n > m}$

$S \rightarrow aA$

$A \rightarrow aA$

$A \rightarrow aAb$

$A \rightarrow \lambda$

Example 3: Is $L = \{a^n b^m : n \neq m\}$ a CFL?

Two conditions

$\underline{n < m}$

$S \to Bb$

$B \to Bb$

$B \to aBb$

$B \to \lambda$

$\underline{n > m}$

$S \to aA$

$A \to aA$

$A \to aAb$

$A \to \lambda$

Combined Condition

$\underline{n \neq m}$

$S \to aA$

$S \to Bb$

$A \to aA$

$A \to aAb$

$A \to \lambda$

$B \to Bb$

$B \to aBb$

$B \to \lambda$

Example 3: Is $L = \{a^n b^m : n \neq m\}$ a CFL?

Two conditions

$\underline{n < m}$

$S \rightarrow Bb$

$B \rightarrow Bb$

$B \rightarrow aBb$

$B \rightarrow \lambda$

$\underline{n > m}$

$S \rightarrow aA$

$A \rightarrow aA$

$A \rightarrow aAb$

$A \rightarrow \lambda$

Combined Condition

$\underline{n \neq m}$

$S \rightarrow aA$

$S \rightarrow Bb$

$A \rightarrow aA$

$A \rightarrow aAb$

$A \rightarrow \lambda$

$B \rightarrow Bb$

$B \rightarrow aBb$

$B \rightarrow \lambda$

Example 3: Is $L = \{a^n b^m : n \neq m\}$ a CFL?

$G = (\{S, A, B\}, \{a, b\}, S, P)$

$$S \rightarrow aA \mid Bb$$
$$A \rightarrow aA \mid aAb \mid \lambda$$
$$B \rightarrow Bb \mid aBb \mid \lambda$$

Does the grammar produce all strings in the language above and only strings apart of that language?

Yes.

Is the grammar context-free?

Yes.

THE LANGUAGE $L(G)$ IS A CONTEXT-FREE LANGUAGE.

Leftmost and Rightmost Derivation

When grammars are nonlinear and contain multiple variables in the sentential forms, their can be a difference in the order in which the variables are replaced.

Left and Right Derivations

$S \to aABb$

Example 4

$$G = (\{A, B, S\}, \{a, b\}, S, P)$$

$$S \to AB$$
$$A \to aaA$$
$$A \to \lambda$$
$$B \to bB$$
$$B \to \lambda$$

Example 4

$$G = (\{A, B, S\}, \{a, b\}, S, P)$$

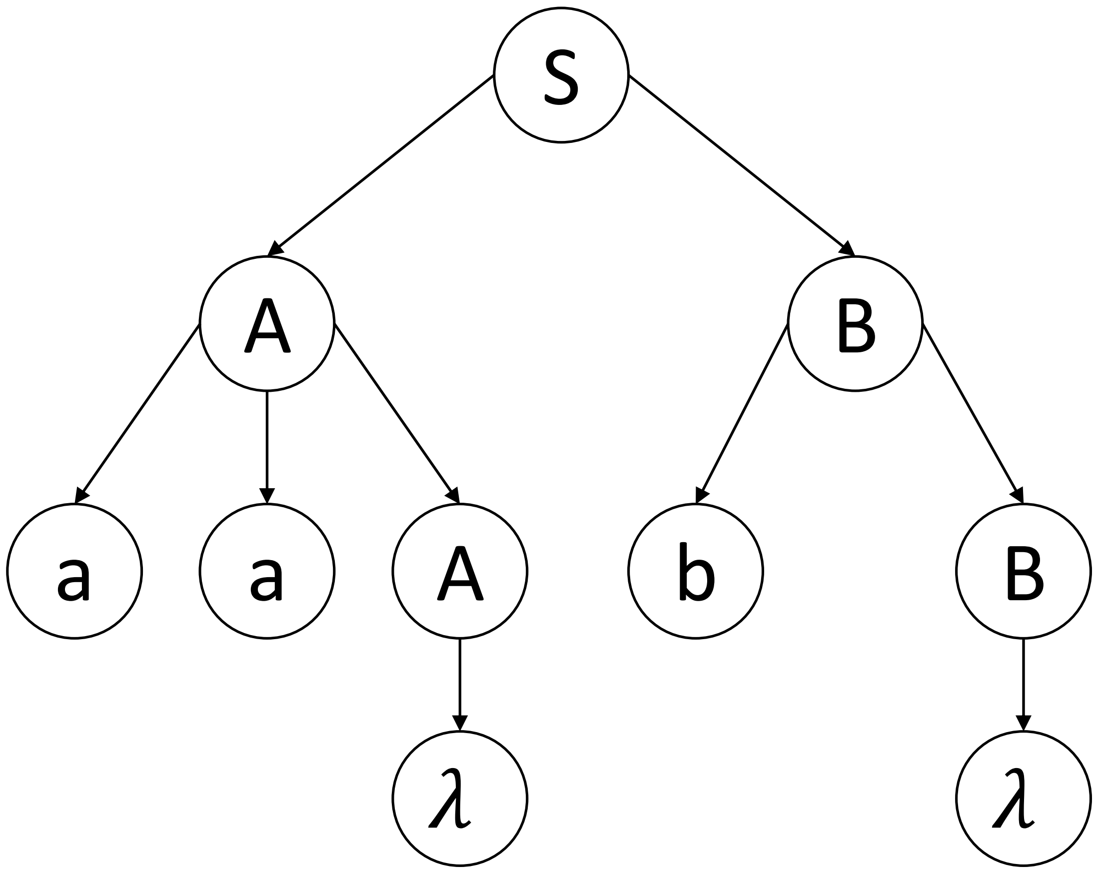
1. $S \rightarrow AB$
2. $A \rightarrow aaA$
3. $A \rightarrow \lambda$
4. $B \rightarrow bB$
5. $B \rightarrow \lambda$

$$L(G) = \{a^{2n}b^m : n \geq 0, m \geq 0\}$$

$$S \overset{1}{\Rightarrow} AB \overset{2}{\Rightarrow} aaAB \overset{3}{\Rightarrow} aaB \overset{4}{\Rightarrow} aabB \overset{5}{\Rightarrow} aab$$

<u>Leftmost Derivation</u>

String: aab



Example 4

$G = (\{A, B, S\}, \{a, b\}, S, P)$

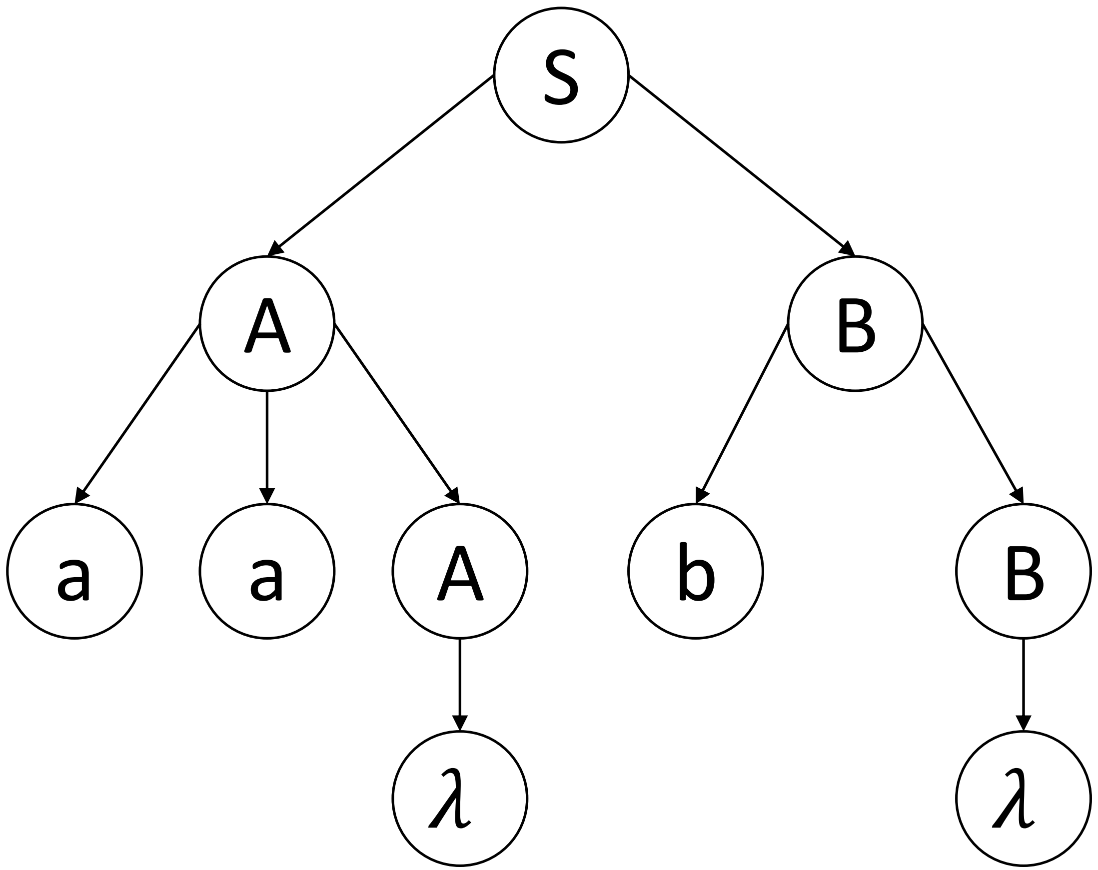
1. $S \rightarrow AB$
2. $A \rightarrow aaA$
3. $A \rightarrow \lambda$
4. $B \rightarrow bB$
5. $B \rightarrow \lambda$

$L(G) = \{a^{2n}b^m : n \geq 0, m \geq 0\}$

$$S \overset{1}{\Rightarrow} AB \overset{4}{\Rightarrow} AbB \overset{5}{\Rightarrow} Ab \overset{2}{\Rightarrow} aaAb \overset{3}{\Rightarrow} aab$$

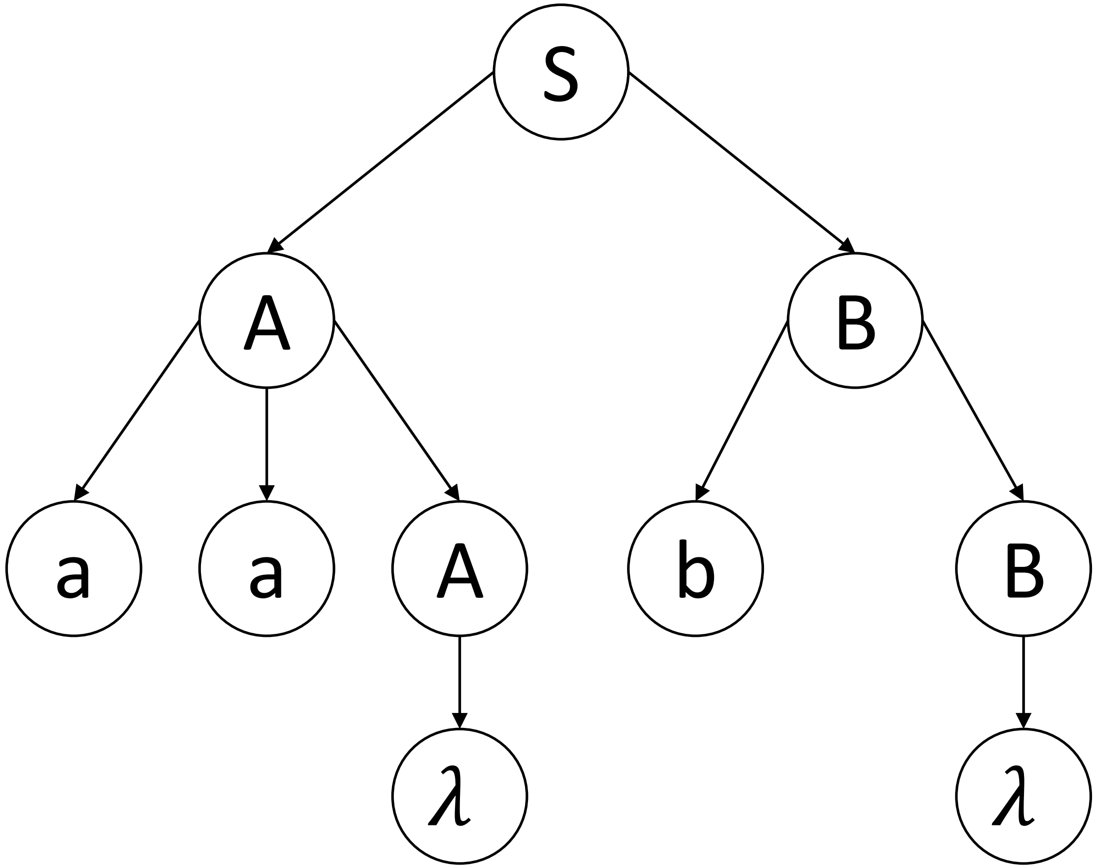
Rightmost Derivation

String: aab



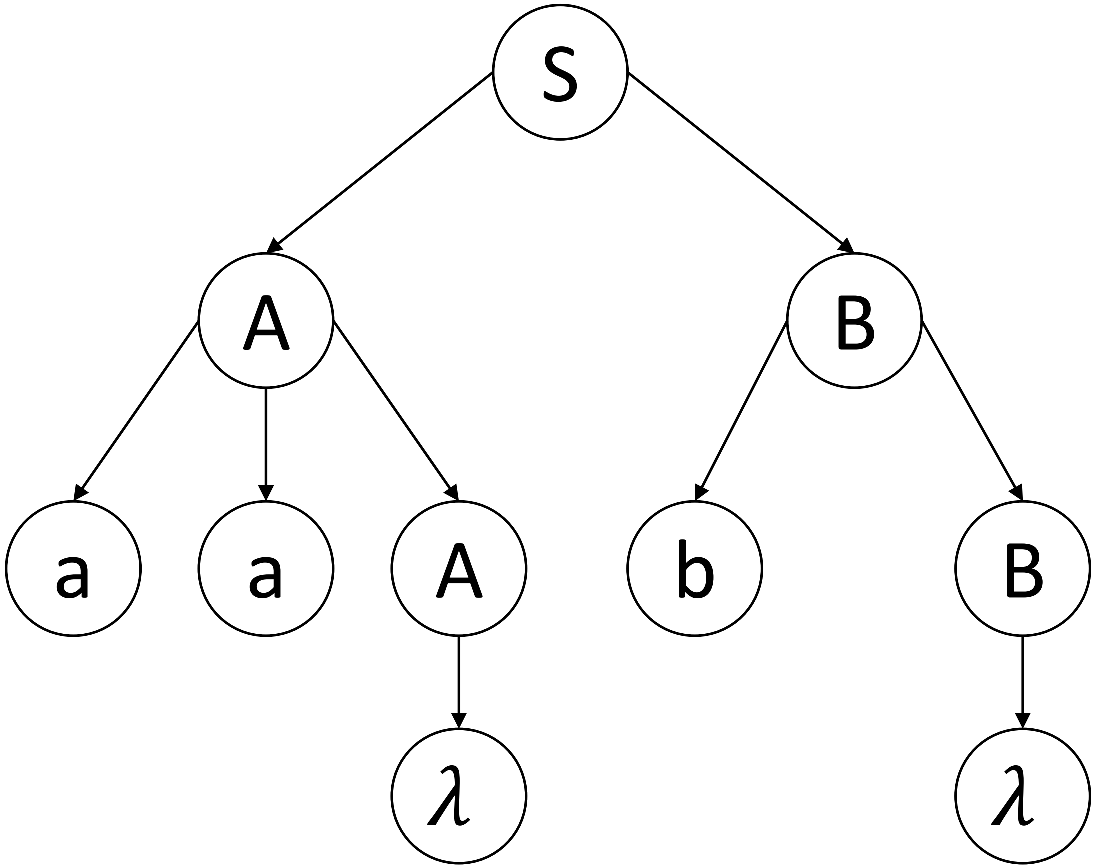
Leftmost Derivation

String: aab



Rightmost Derivation

String: aab



Parsing and Ambiguity

Parsing and Membership

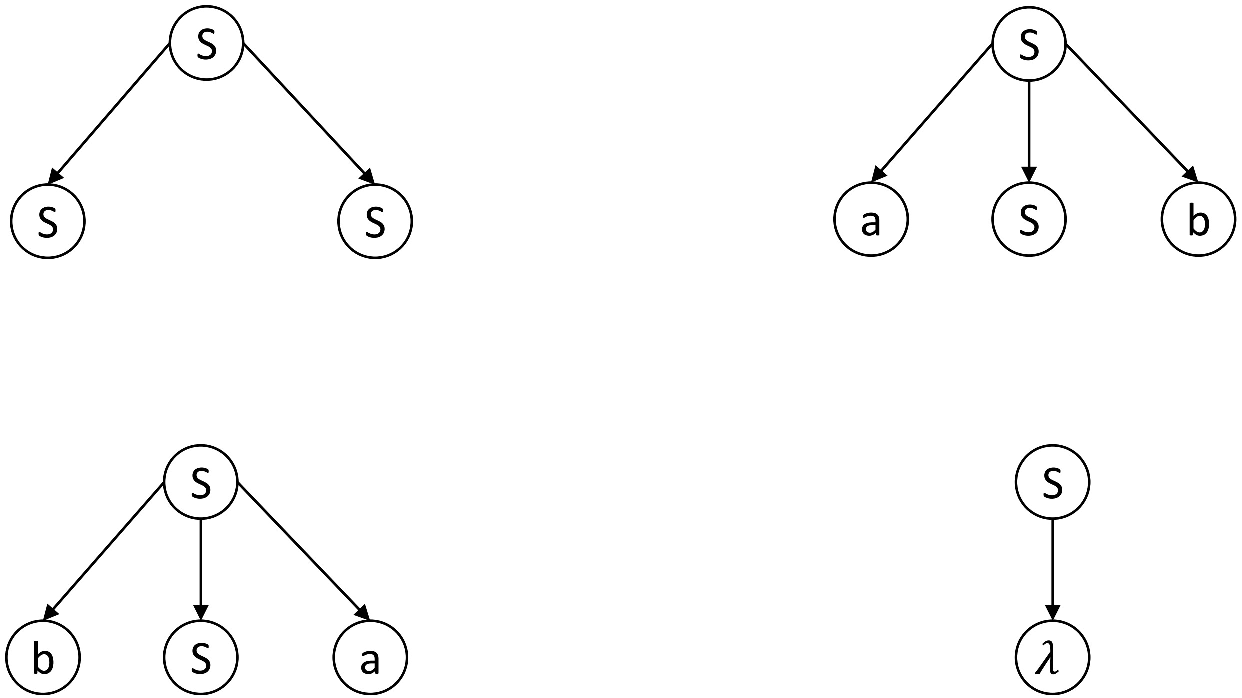
- Can we find a derivation of grammar G that generates a given w
  - if so, $w \in L(G)$.
  - if not, $w \notin L(G)$.
- Exhaustive search parsing is inefficient and may never terminate if $w \notin L(G)$.

Example 5

Consider the grammar $G = (\{S\}, \{a, b\}, S, P)$ with productions

$$S \to SS$$
$$S \to aSb$$
$$S \to bSa$$
$$S \to \lambda$$

<u>Round 1</u>



Deriving the string

aabb

Example 5

Consider the grammar $G = (\{S\}, \{a, b\}, S, P)$ with productions

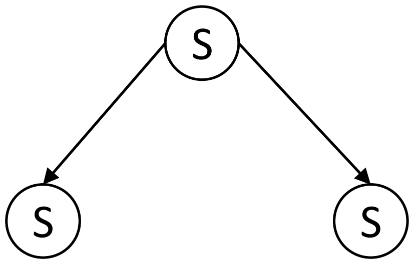
$S \to SS$

$S \to aSb$

$S \to bSa$

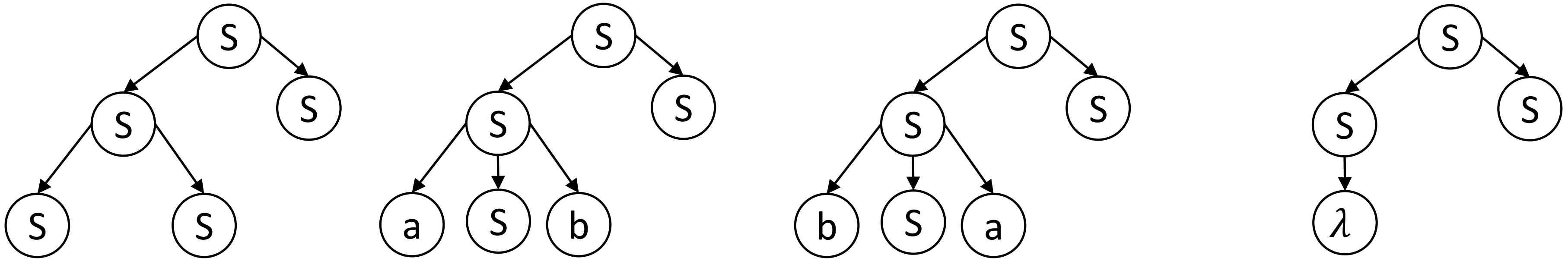
$S \to \lambda$

Round 2



Deriving the string

aabb

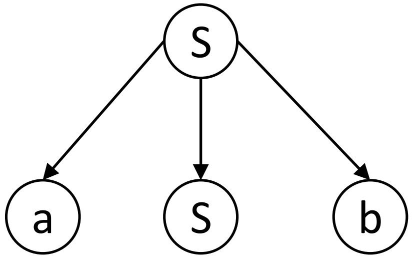


Example 5

Consider the grammar $G = (\{S\}, \{a, b\}, S, P)$ with productions

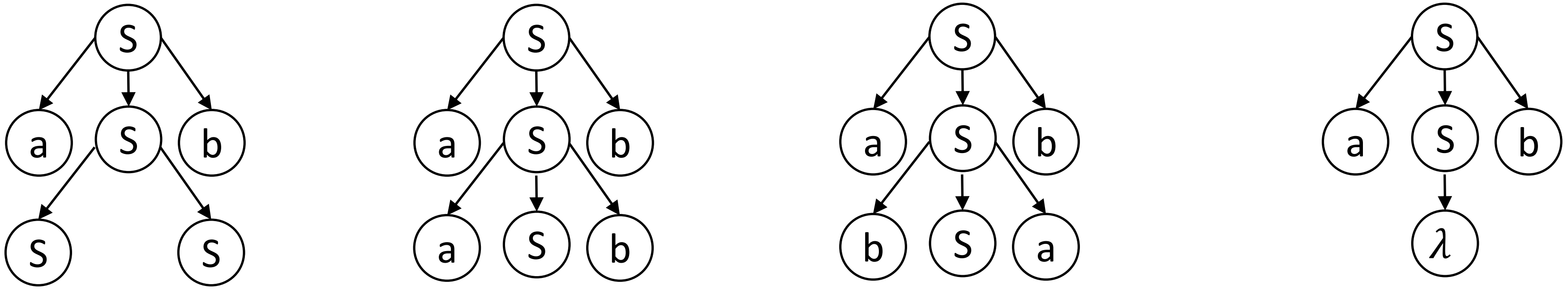
$S \rightarrow SS$
$S \rightarrow aSb$
$S \rightarrow bSa$
$S \rightarrow \lambda$

<u>Round 2</u>



Deriving the string

aabb



Parsing and Ambiguity

- Productions with a lambda (decreases the size of the sentential form)

$$A \to \lambda : A \in V$$

- Variable simply derives another variable

$$A \to B : A, B \in V$$

Simple Grammar (s-grammar)

$G = (V, T, S, P)$

Restrict Productions to:

- $A \rightarrow ax : A \in V, a \in T, x \in V^*$
- (A, a)

$S \rightarrow aS \mid bSS \mid c$

$S \rightarrow aS \mid bSS \mid aSS \mid c$

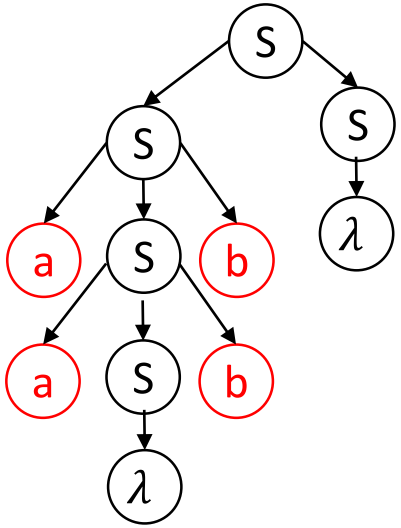
Ambiguity in Grammars and Languages

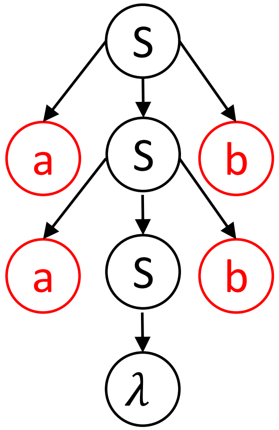
A context-free grammar G is said to be ambiguous if $w \in L(G)$ and there is at least 2 distinct derivation trees for w.

Example 5 Revisited

Consider the grammar $G = (\{S\}, \{a, b\}, S, P)$ with productions

$$S \rightarrow SS$$
$$S \rightarrow aSb$$
$$S \rightarrow bSa$$
$$S \rightarrow \lambda$$





Inherently Ambiguous Language

- If every grammar that generates L is ambiguous, then L is an inherently ambiguous language.