

CSE 105

THEORY OF COMPUTATION

"Winter" 2018

<http://cseweb.ucsd.edu/classes/wi18/cse105-ab/>

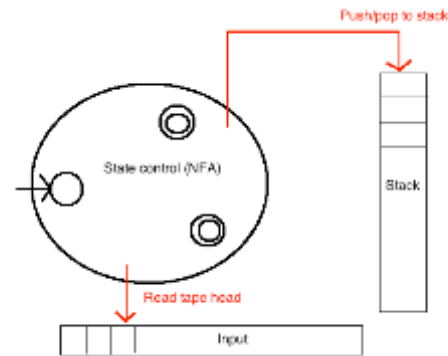
Today's learning goals

Sipser Section 2.1

- Identify the components of a formal definition of a CFG
- Derive strings in the language of a given CFG
- Determine the language of a given CFG
- Design a CFG generating a given language

PDA: NFA as **??**: RegExp

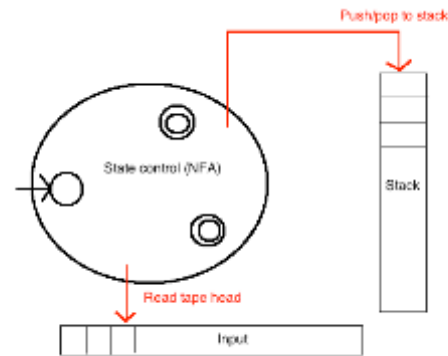
- Automata
 - String is read by the machine one character at a time, from left to right.
 - Determine if computation is successful by checking if entire string was read, and if land at an accept state.



Note: PDA based on NFA; can't always be determinized.

PDA: NFA as ?? : RegExp

- Automata
 - String is read by the machine one character at a time, from left to right.
 - Determine if computation is successful by checking if entire string was read, and if land at an accept state.
- Regular expressions and ??
 - Derive all strings in the language by following rules for required patterns.



Context-free grammar

Informally, a collection of rules used to *create* string.
CFGs *generate* languages.

Some sample rules:

$$S \rightarrow aTb$$

$$T \rightarrow aT$$

$$T \rightarrow bTS$$

$$S \rightarrow \varepsilon$$

More formally...

Context-free grammar

Sipser Def 2.2, page 102

(V, Σ, R, S)

Variables: finite set of (usually upper case) variables **V**

Terminals: finite set of alphabet symbols **Σ** $V \cap \Sigma = \emptyset$

Rules/Productions: finite set of allowed transformations **R**

$$A \rightarrow u \quad A \in V, u \in (V \cup \Sigma)^*$$

Start variable: origination of each derivation **S**

Derivation

Set of variables

Set of terminals

$$G = (\{S\}, \{0\}, R, S)$$

with the rules

$$R = \{S \xrightarrow{1} 0S, S \xrightarrow{2} 0\}$$

Sample derivation:

$$S \xrightarrow{1} 0S \xrightarrow{1} 00S \xrightarrow{2} 000$$

Start variable

One-step application of rule

String of terminals

Context-free language

Sipser p. 104

The **language generated by a CFG** (V, Σ, R, S) is

$\{ w \text{ in } \Sigma^* \mid \text{starting with the Start variable and applying sequence of rules, can derive } w \text{ on RHS} \}$

If $G = (V, \Sigma, R, S)$
the language
generated by G is
denoted $L(G)$.

Notation:
 $S \Rightarrow^* w$

Terminology: sequence of
rule applications is
derivation

Context-free language

Sipser p. 104

The language generated by CFG (V, Σ, R, S) is

$\{ w \text{ in } \Sigma^* \mid \text{starting with the Start variable and applying sequence of rules, can derive } w \text{ on RHS} \}$.

What is the language of the CFG $(\{S\}, \{0\}, R, S)$ with the rules

$R = \{ S \rightarrow 0S, S \rightarrow 0 \}$?

A. $\{0\}$

~~B. $\{0, 0S\}$~~

C. $\{0, 00, 000, \dots\}$

E. I don't know.

~~D. $\{\epsilon, 0, 00, 000, \dots\}$~~

$000 \in L(G)?$

$S \xRightarrow{p} 0S \xRightarrow{p} 00S$
 $\downarrow \textcircled{2}$

000

$L(00^*)$

Context-free language

Sipser p. 104

What is the language of the CFG ($\{S\}$, $\{0, 1\}$, R , S) with rules

$S \rightarrow 0S$

$S \rightarrow 1S$

$S \rightarrow \epsilon$

$(0^* 1^*)^*$

$w_1 w_2 \dots w_k$

LHS

$S \rightarrow 0S \mid 1S \mid \epsilon$

~~A.~~ $L(0^* 1^*)$

~~B.~~ $L(0^* \cup 1^*)$

C. $L((0 \cup 1)^*)$

D. $L((0^* 1^*)^*)$

E. I don't know.

$S \xrightarrow{0} 0S \xrightarrow{1} 01S \Rightarrow 011S$

$L(G) = \{0, 1\}^*$

011

Designing a CFG

$L = \{ abba \}$

Which CFG generates L ?

- A. $(\{S, T, V, W\}, \{a, b\}, \{S \rightarrow aT, T \rightarrow bV, V \rightarrow bW, W \rightarrow a\}, S)$
- B. $(\{Q\}, \{a, b\}, \{Q \rightarrow abba\}, Q)$
- C. $(\{X, Y\}, \{a, b\}, \{X \rightarrow aYa, Y \rightarrow bb\}, X)$
- D. All of the above
- E. None of the above

Context-free languages

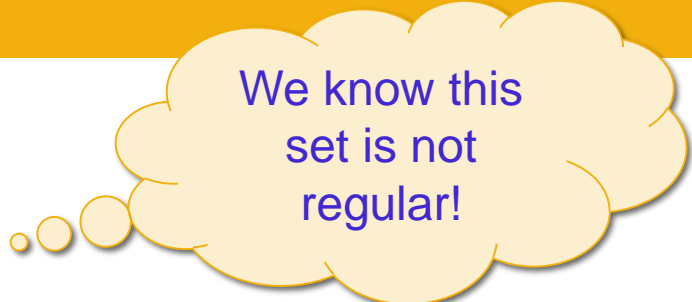
- $L(00^*)$
- $L((0U1)^*)$
- $\{abba\}$

Is any nonregular set context-free?

What about the languages that are recognized by PDAs?

Designing a CFG

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

A yellow thought bubble with a black outline and a drop shadow, containing text. It is connected to the main content by three smaller circles of decreasing size.

We know this set is not regular!

Designing a CFG

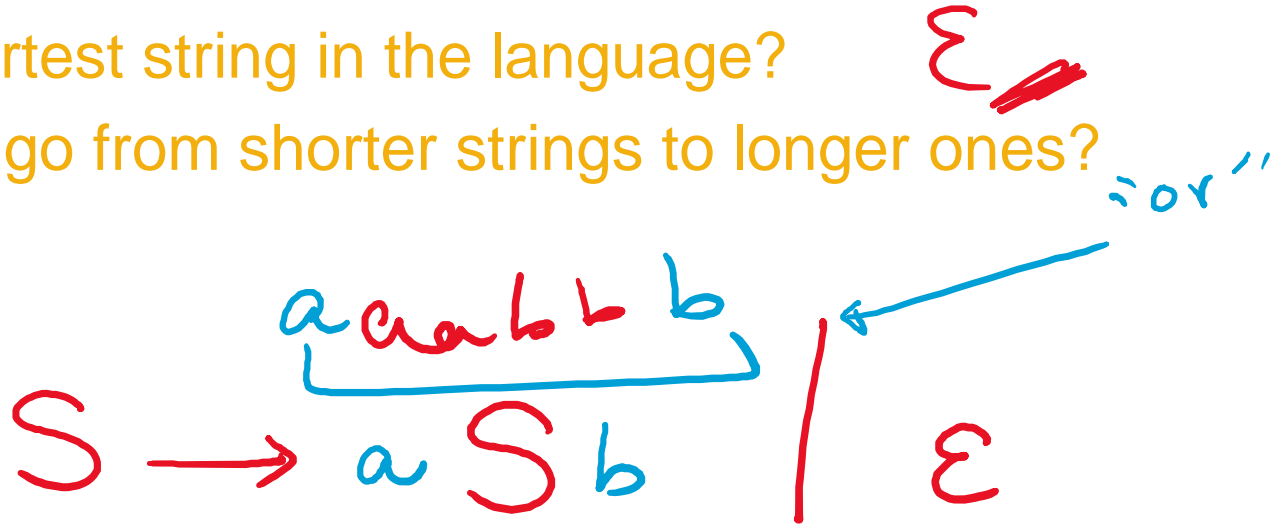
$$G = (\{S\}, \{a, b\}, R, S)$$

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

with rules
 $S \rightarrow aSb \mid \epsilon$

One approach:

- what is shortest string in the language?
- how do we go from shorter strings to longer ones?



Context-free languages

- $L(00^*)$
- $L((0U1)^*)$
- $\{abba\}$
- $\{a^n b^n \mid n \geq 0\}$

Ex: $\{0^i 1^j \mid j \geq i \geq 0\}$

recognizable by a PDA

PDAs and CFGs are equally expressive

Theorem 2.20: A language is context-free if and only if some nondeterministic PDA recognizes it.

Consequences

- Quick proof that every regular language is context free
- To prove closure of class of CFLs under a given operation, can choose two modes of proof (via CFGs or PDAs) depending on which is easier

For next time

Exam 1 next class Wednesday, February 7

- **Bring ID, pen**
- **Bring note card (half page, double sided, handwritten)**
- **Check assigned seat on Piazza**
- **Piazza will be inactive from 8AM to 3:30PM on Wednesday**