

Oct 19

COMPSCI 3331

Fall 2022

# What's next?

- ▶ Midterm marks: after reading week.
- ▶ Assignment 1: Marks available now.
- ▶ Assignment 2: due today.
- ▶ Quiz 3 marks available now.
- ▶ Quiz 4 after reading week - **Lectures 8 and 9.**
- ▶ Assignment 3: available by Nov 8 at the **latest.**
- ▶ Office Hours Thursday - 1-3 PM.

# CNF - $\epsilon$ productions

$S \rightarrow ABDE$   
 $A \rightarrow BDa \mid \underline{D}$   
 $B \rightarrow aaE$   
 $D \rightarrow \epsilon \mid \text{ } \mid G$   
 $E \rightarrow DD \mid aab$   
 $F \rightarrow DE \mid \epsilon$   
 $G \rightarrow aaa$

( $D, E$  are nullable)  
 $\downarrow$  anyone that is nullable can be removed.  
 $AB$   $B$  etc....

Which variables are nullable?

$D, F, A, E$

# CNF - unit pairs

remove all  $\epsilon$  production

$S \rightarrow ABDE \mid ABD \mid ABE \mid BDE \mid AB \mid BD \mid BE \mid B$

$A \rightarrow BDa \mid D \mid Ba \mid aaa$

$B \rightarrow aa\bar{E} \mid aa$

$D \rightarrow G$

$E \rightarrow DDD \mid aab \mid aaa$

$F \rightarrow DE \mid DIE \mid aaa$

$G \rightarrow aaa$

What are the unit pairs?

pairs of variables such as  $A \rightarrow B$  (one to one).

e.g.  $SB$   $FD$   $AG$

$AD$   $\bar{F}\bar{E}$

$DG$   $\bar{E}G$  ( $A \rightarrow B \rightarrow C$ )

$\bar{E}D$   $FG$

# CNF - useless symbols : get rid of the rules that could be simplified.

$S \rightarrow$

$A \rightarrow$

$B \rightarrow$

$D \rightarrow$

$E \rightarrow$

In fact,  $F$  is unreachable.   
 So it is deleted

~~$F \rightarrow$~~   $DE|~~E~~|aaa|DD|aab$  (unit pairs are eliminated)  
 $G \rightarrow$

What are reachable / generating symbols?

Assume  $H \rightarrow HH$ , it is useless too.

# CNF - letter production

$S \rightarrow$

$A \rightarrow$

$B \rightarrow aa\bar{E} \Rightarrow B \rightarrow x_a x_a \bar{E}$

$D \rightarrow$

$E \rightarrow$

$X_a \rightarrow a$

$X_b \rightarrow b$

$X_c \rightarrow c$

$\bar{C} \rightarrow \bar{x}_a \bar{x}_b \bar{x}_c$

# CNF

$$S \rightarrow \text{ABDE} \rightarrow \gamma_{ABD} \bar{E}$$

$$A \rightarrow$$

$$B \rightarrow$$

$$D \rightarrow$$

$$E \rightarrow$$

$$X_a \rightarrow a$$

$$X_b \rightarrow b$$

$$X_c \rightarrow c$$

$$\gamma_{ABD} \rightarrow \text{ABD} \rightarrow \gamma_{AB} D$$

$$\gamma_{AB} \rightarrow AB$$

breaking big rules into shorter rules with length 2.

get rid of  $\epsilon$  productions

↓

get rid of unit pairs.

↓

get rid of useless steps.

↓

convert long-unterminal into short rules.

↓

rules having single letter/two unterminals only.