B1 1914078 TOC

# Tutorial

Q1) i) $S \rightarrow A \mid B$
$A \rightarrow 0A1 \mid \varepsilon$
$B \rightarrow 1B0 \mid \varepsilon$

By steps :-

(a) Make a new start symbol if it's in RHS of any production, but this step doesn't satisfy.

(b) Remove null productions

On removing $A \rightarrow \varepsilon$ :
$S \rightarrow A \mid B \mid \varepsilon$
$A \rightarrow 0A1 \mid 01$
$B \rightarrow 1B0 \mid \varepsilon$

On removing $B \rightarrow \varepsilon$ :
$S \rightarrow A \mid B \mid \varepsilon$
$A \rightarrow 0A1 \mid 01$
$B \rightarrow 1B0 \mid 10$

On removing $S \rightarrow \varepsilon$ :
$S \rightarrow A \mid B$
$A \rightarrow 0A1 \mid 01$
$B \rightarrow 1B0 \mid 10$

(c) Removing unit productions

Removing $S \rightarrow A$ :
$S \rightarrow 0A1 \mid 01 \mid B$

Removing $S \rightarrow B$ :
$S \rightarrow 0A1 \mid 01 \mid 1B0 \mid 10$

(d) Making production to CNF by making new production. We have

$S \rightarrow OA1 \mid 01 \mid 1BO \mid 10$
$A \rightarrow OA1 \mid 01$
$B \rightarrow 1BO \mid 10$

Making productions

$D \rightarrow 0$
$E \rightarrow 1$

$S \rightarrow DAE \mid DE \mid EBD \mid ED$
$A \rightarrow DAE \mid DE$
$B \rightarrow EBD \mid ED$
$D \rightarrow 0$
$E \rightarrow 1$

Now,

$X \rightarrow DA$
$Y \rightarrow EB$

$S \rightarrow XE \mid DE \mid YD \mid ED$
$A \rightarrow XE \mid DE$
$B \rightarrow YD \mid ED$
$D \rightarrow 0$
$E \rightarrow 1$
$X \rightarrow DA$
$Y \rightarrow EB$

So in CNF, we have

$G = (\{S, A, B, D, E, X, Y\}, \{0, 1\}, S, T)$

where $T \Rightarrow S \rightarrow XE \mid DE \mid YD \mid ED$

Q1) ii) $G = (\{E, T, F\}, \{a, +, *, (, )\}, R, E)$

$E \rightarrow E + T \mid T$
$T \rightarrow T * F \mid F$
$F \rightarrow (E) \mid a$

**Step 1:** E appears in RHS so we introduce a new state/variable E'

$E' \rightarrow E$
$E \rightarrow E + T \mid T$
$T \rightarrow T * F \mid F$
$F \rightarrow (E) \mid a$

**Step 2:** There are no null productions

**Step 3:** Removal of unit production ($E' \rightarrow E$, $E \rightarrow T$, $T \rightarrow F$)

**Removing $T \rightarrow F$**

$E' \rightarrow E$
$E \rightarrow E + T \mid T$
$T \rightarrow T * F \mid (E) \mid a$
$F \rightarrow (E) \mid a$

**Removing $E \rightarrow T$**

$E' \rightarrow E$
$E \rightarrow E + T \mid T * F \mid (E) \mid a$
$F \rightarrow (E) \mid a$
$T \rightarrow T * F \mid (E) \mid a$

**Removing $E' \rightarrow E$**

$E' \rightarrow E + T \mid T * F \mid (E) \mid a$

$E \rightarrow E+T \mid T * F \mid (E) \mid a$

$T \rightarrow T * F \mid (E) \mid a$

$F \rightarrow (E) \mid a$

**Step 4 :-**

$E' \rightarrow E+T \mid T*F \mid (E) \mid a$

$E \rightarrow E+T \mid T*F \mid (E) \mid a$

$T \rightarrow T*F \mid (E) \mid a$

$F \rightarrow (E) \mid a$

$E \rightarrow EA \mid TB \mid Cc \mid a$

$E \rightarrow EA \mid TB \mid Cc \mid a$

$T \rightarrow TB \mid Cc \mid a$

$F \rightarrow Cc \mid a$

$A \rightarrow +T$

$B \rightarrow *F$

$C \rightarrow E)$

**Step 5:** Now changing the production

$X \rightarrow ($

$Y \rightarrow +$

$Z \rightarrow *$

$U \rightarrow )$

$E' \rightarrow EA \mid TB \mid XC \mid a$

$E \rightarrow EA \mid TB \mid XC \mid a$

$T \rightarrow TB \mid XC \mid a$

$F \rightarrow XC \mid a$

$A \rightarrow YT$

$B \rightarrow ZF$

$C \rightarrow EU$

02) Chomsky Normal form (CNF) puts some constraints on the grammar rules while preserving the same language. The benifit is that if a grammar is in CNF, then we can avoid the ambiguity problem during parsing.

- Another benifit of CNF is that it provide an upper bound for parsing complexity. For any string w with length n, the derivation must be $2n-1$ steps: $n-1$ steps to obtain length n, and another n steps to replace the variables with terminals.
- The worst case is just to enumerate all possible derivations with $2n-1$ steps which is exponential but at least is soluable.