

CHOMSKY NORMAL FORM

15/1/2010
E10.1

EXERCISE 1

Consider the following grammar:

$$S \rightarrow OS | BD | EO$$

$$B \rightarrow C | \epsilon | 1$$

$$D \rightarrow O | A | S | \epsilon$$

$$E \rightarrow BE | SE$$

$$C \rightarrow O$$

$$F \rightarrow C$$

a) Simplify the grammar.

b) Transform the simplified grammar in Chomsky normal form.

EXERCISE 2

Transform the following grammars in Chomsky normal form.

$$S \rightarrow A | B$$

$$A \rightarrow O | A | B | O | O$$

$$B \rightarrow 1 | B | O | 1 | O$$

$$S \rightarrow A | B$$

$$A \rightarrow O | S | O | B | O | O$$

$$B \rightarrow 1 | B | O | 1 | O$$

SOLUTIONS (16/01/2009)

E10.2

1) a) Elimination of ϵ -productions

Nullable symbols: $\{B, D\}$

$\{B, D, S\}$

We get: $S \rightarrow OS|O|BD|B|D|EO$

$B \rightarrow C|1$

$D \rightarrow O|AO|S$

$E \rightarrow BE|SE|E$

$C \rightarrow O$

$F \rightarrow C$

Elimination of unit productions

From reachability: $S \Rightarrow^* B, S \Rightarrow^* D, S \Rightarrow^* C$

$B \Rightarrow^* C$

$D \Rightarrow^* S, D \Rightarrow^* B, D \Rightarrow^* C$

$F \Rightarrow^* C$

We get: $S \rightarrow OS|O|BD|1|AO|EO$

$B \rightarrow O|1$

$D \rightarrow O|AO|OS|BD|1|EO$

$E \rightarrow BE|SE$

$C \rightarrow O$

$F \rightarrow O$

Elimination of useless symbols

Generating symbols: $\{0, 1\}$

$\{0, 1, S, B, D, C, F\}$

1)a) can't

E10.3

We get: $S \rightarrow 0S|0|BD|1$

$B \rightarrow 0|1$

$D \rightarrow 0|0S|BD|1$

$C \rightarrow 0$

$F \rightarrow 0$

Reachable symbols: $\{S\}$

$\{S, 0, B, D, 1\}$

We get: $S \rightarrow 0S|0|BD|1$

$B \rightarrow 0|1$

$D \rightarrow 0|0S|BD|1$

1)b) Since we do not have long productions we only need to remove "mixed bodies" (such as $0S$) by introducing some new productions ($G \rightarrow 0$).

We get: $S \rightarrow GS|0|BD|1$

$B \rightarrow 0|1$

$D \rightarrow 0|GS|BD|1$

$G \rightarrow 0$

2) Let us start with the left-hand grammar.

Elimination of unit productions

From reachability: $S \Rightarrow^* A, S \Rightarrow^* B, A \Rightarrow^* B$

We get: $S \rightarrow 0A0|00|1B0|10$

$A \rightarrow 0A0|1B0|10|00$

$B \rightarrow 1B0|10$

2) con't

E10.4

Remove "mixed bodies" ($C \rightarrow 0, D \rightarrow 1$)

We get: $S \rightarrow CAC | CC | DBC | DC$

$A \rightarrow CAC | DBC | DC | CC$

$B \rightarrow DBC | DC$

$C \rightarrow 0$

$D \rightarrow 1$

"Factor" long productions ($E \rightarrow AC, F \rightarrow BC$)

We get: $S \rightarrow CE | CC | DF | DC$

$A \rightarrow CE | DF | DC | CC$

$B \rightarrow DF | DC$

$C \rightarrow 0$

$D \rightarrow 1$

$E \rightarrow AC$

$F \rightarrow BC$

With respect to the right-hand grammar we only provide the solution. Note that we have eliminated the useless symbol A .

CNF: $S \rightarrow CE | CC | DF | DC$

$B \rightarrow DF | DC$

$C \rightarrow 0$

$D \rightarrow 1$

$E \rightarrow SC$

$F \rightarrow BC$