

NORMAL FORMS

Recall: We showed that CFLs = languages accepted by PDAs
Used normal forms for CFGs to show this equivalence.

Today: How to convert any grammar into these normal forms.

We used Greibach Normal form, where every rule is of the form
 $A \rightarrow cB_1 \dots B_k$, where

$$k \geq 0, c \in T, A, B_1, \dots, B_k \in NT.$$

There is another, Chomsky Normal form, where every rule is
either of the form $A \rightarrow c$ or $A \rightarrow BC$, where $c \in T, A, B, C \in NT$.

Thm: For every CFG G , there is a CFG G_1 in Chomsky Normal form,
and a CFG G_2 in Greibach Normal form s.t.

$$\mathcal{L}(G_1) = \mathcal{L}(G_2) = \mathcal{L}(G) \setminus \{\epsilon\}.$$

How do we show this?

We begin by getting rid of rules of the form

$A \rightarrow \epsilon$ ϵ -production and

$A \rightarrow B$ unit-production.

We might need to replace these with other rules!

If I remove $S \rightarrow \epsilon$ from $S \rightarrow \epsilon \mid aSb \mid bSa \mid SS$

without adding any other rules in its place,

what language does my resultant grammar generate?

How can I fix this?

Claim: For any CFG $G = (NT, T, R, S)$, there is a CFG G' with no ϵ - or unit-productions such that $L(G') = L(G) \setminus \{\epsilon\}$

Proof: Consider a set of rules \hat{R} s.t. $R \subseteq \hat{R}$, and

- ① if $A \rightarrow \alpha B \beta \in \hat{R}$ and $B \rightarrow \epsilon \in \hat{R}$, then $A \rightarrow \alpha \beta \in \hat{R}$ and
- ② if $A \rightarrow B \in \hat{R}$ and $B \rightarrow \gamma \in \hat{R}$, then $A \rightarrow \gamma \in \hat{R}$.

where $A, B \in NT$, $\alpha, \beta, \gamma \in (NT \cup T)^*$

* How do we know that \hat{R} does not keep growing infinitely large?
it is not the case that both α and β are empty, and $\gamma \neq \epsilon$.

\hat{R} is finite, and can be constructed inductively from R .

Now consider $\hat{G} = (NT, T, \hat{R}, S)$. Easy to show $L(\hat{G}) = L(G) \setminus \{\epsilon\}$
Exercise!

We now have to show that if $x \in \Sigma^+$ is derived using some sequence of application of rules in \hat{R} , it does not use any ϵ - or unit-productions. If we show this, we can throw out those rules and get the desired G' .

Consider $x \neq \epsilon$, and consider a shortest derivation for x in \hat{G} .

(i) Suppose $B \rightarrow \epsilon$ is used at some point in this derivation.

$$S \xrightarrow{*} \alpha B \gamma \longrightarrow \alpha \gamma \xrightarrow{*} x$$

At least one of α and γ must be $\neq \epsilon$.

So, there must be some $\delta, \eta, \beta, \tau$ s.t.

$$S \xrightarrow{i} \delta \eta B \beta \tau \xrightarrow{j} \alpha B \gamma \xrightarrow{l} \alpha \gamma \xrightarrow{k} x$$

But how do we get $\delta\eta B\delta\sigma$? Via some rule of the form $A \rightarrow \eta B\delta$.

So, $S \xrightarrow{i-1} \delta A\sigma \xrightarrow{1} \delta\eta B\delta\sigma \xrightarrow{j} \alpha B\gamma \xrightarrow{1} \alpha\delta \xrightarrow{k} x$. $i+j+1+k$

$S \xrightarrow{o} S \xrightarrow{1} aSb \xrightarrow{o} aSb \xrightarrow{1} ab \xrightarrow{o} ab$

But since we have $A \rightarrow \eta B\delta$ and $B \rightarrow \epsilon$ in \hat{R} , by ①,
 $A \rightarrow \eta\delta \in \hat{R}$.

So we can construct a shorter derivation for x , as follows:

$$S \xrightarrow{i-1} \delta A\sigma \xrightarrow{1} \delta\eta\delta\sigma \xrightarrow{j} \alpha\delta \xrightarrow{k} x. \quad i+j+k$$

This cuts out at least one rule application from a minimal one for x .
Contradiction!

(ii) Suppose $A \rightarrow B$ is used at some point in a minimal derivation of α .

$$S \xrightarrow{*} \alpha A \beta \rightarrow \alpha B \beta \xrightarrow{*} \alpha$$

Since $\alpha \in T^*$ and $B \in NT$, B must have been disposed of via some rule, of the form $B \rightarrow \gamma$. So, there must be δ, δ, η s.t.

$$S \xrightarrow{i} \alpha A \beta \xrightarrow{j} \alpha \delta A \eta \xrightarrow{1} \alpha \delta B \eta \xrightarrow{1} \alpha \delta \gamma \eta \xrightarrow{k} \alpha \quad i+1+j+1+k$$

But since $A \rightarrow B \in \hat{R}$ and $B \rightarrow \gamma \in \hat{R}$, $A \rightarrow \gamma \in \hat{R}$.

So, we get a shorter derivation for α as follows :

$$S \xrightarrow{i} \alpha A \beta \xrightarrow{j} \alpha \delta A \eta \xrightarrow{1} \alpha \delta \gamma \eta \xrightarrow{k} \alpha. \quad i+1+j+k$$

which contradicts the minimality of the above derivation of α .