

Example:

Let \mathcal{L} be the language formed from the terminal vocabulary $\{a,b\}$ such that for any sentence S that contains at least as many a 's as b 's, $S \in \mathcal{L}$.

Can we find a grammar G for \mathcal{L} ?

Nonterminals:

$\langle S \rangle$ --- stands for whole language

$\langle A \rangle$ --- stands for sentences with at least 1 extra a .

$\langle B \rangle$ --- stands for sentences with at most 1 extra b .

$\langle S \rangle \Rightarrow a\langle S \rangle \mid a\langle B \rangle \mid b\langle A \rangle \mid a$

$\langle A \rangle \Rightarrow a\langle S \rangle \mid b\langle A \rangle\langle A \rangle \mid a$

$\langle B \rangle \Rightarrow b\langle S \rangle \mid a\langle B \rangle\langle B \rangle \mid b$

Exercise: Prove that $L(G) = \mathcal{L}$.

Notation: Suppose we have a vocabulary $V = \{S_1, S_2, \dots, S_n\}$. We denote by $V^* = \{S_1, S_2, \dots, S_n\}^*$ the set of all sequences of symbols taken from V (including the empty string).

We use V^+ or $\{S_1, S_2, \dots, S_n\}^+$ to denote the set of all *non-empty* sequences of symbols taken from V .

Example: Consider the following grammar G

$$\begin{aligned} \langle S \rangle &\rightarrow a\langle B \rangle \mid b\langle A \rangle \\ \langle A \rangle &\rightarrow a\langle S \rangle \mid b\langle A \rangle\langle A \rangle \mid a \\ \langle B \rangle &\rightarrow b\langle S \rangle \mid a\langle B \rangle\langle B \rangle \mid b \end{aligned}$$

$L(G)$ is a subset of $\{a,b\}^+$. In particular, it is that subset whose members have exactly the same number of a's as b's.

$$\text{I.e., } L(G) = \{w \in \{a,b\}^+ \mid \#a's = \#b's \text{ in } w\}$$

To show: $L(G) \subseteq \{w \in \{a,b\}^+ \mid \#a's = \#b's \text{ in } w\}$

i.e., $\langle S \rangle \xRightarrow{*} w$ implies that w has as many a's as b's.

and $L(G) \supseteq \{w \in \{a,b\}^+ \mid \#a's = \#b's \text{ in } w\}$

i.e., if w has as many a's as b's, then $\langle S \rangle \xRightarrow{*} w$.

$\langle S \rangle \rightarrow a\langle B \rangle \mid b\langle A \rangle$

$\langle A \rangle \rightarrow a\langle S \rangle \mid b\langle A \rangle\langle A \rangle \mid a$

$\langle B \rangle \rightarrow b\langle S \rangle \mid a\langle B \rangle\langle B \rangle \mid b$

First show $L(G) \subseteq \{w \in \{a,b\}^+ \mid \#a's = \#b's \text{ in } w\}$

Proof: By induction on $n = \#steps$ in a derivation.

Note that until the final step of a derivation,

$w \notin \{a,b\}^+$. (w will contain non-terminals)

But we can prove a stronger invariant on w
that implies $\#a's = \#b's$.

At any step of a derivation, $w \in \{a,b,\langle S \rangle,\langle A \rangle,\langle B \rangle\}^+$.

Furthermore, in w, $\#a's + \#\langle A \rangle's = \#b's + \#\langle B \rangle's$

If this holds at every step, it holds after the **last** step.

But after the last step, $\#\langle A \rangle's = \#\langle B \rangle's = 0$.

Base case: $n = 1$ Only two derivations

$\langle S \rangle \Rightarrow a\langle B \rangle$ and $\langle S \rangle \Rightarrow b\langle A \rangle$

$\#a's + \#\langle A \rangle's = 1 + 0 =$

$\#a's + \#\langle A \rangle's = 0 + 1 =$

$\#b's + \#\langle B \rangle's = 0 + 1$

$\#b's + \#\langle B \rangle's = 1 + 0$

$\langle S \rangle \rightarrow a\langle B \rangle \mid b\langle A \rangle$

$\langle A \rangle \rightarrow a\langle S \rangle \mid b\langle A \rangle\langle A \rangle \mid a$

$\langle B \rangle \rightarrow b\langle S \rangle \mid a\langle B \rangle\langle B \rangle \mid b$

So suppose that for all derivations $\langle S \rangle \xRightarrow{*} w$,

of k or fewer steps, $\#a's + \#\langle A \rangle's = \#b's + \#\langle B \rangle's$.

Consider a derivation of length $k+1$:

$\langle S \rangle \Rightarrow w_1 \Rightarrow \cdots \Rightarrow w_k \Rightarrow w_{k+1}$

By the ind. hyp., in w_k , $\#a's + \#\langle A \rangle's = \#b's + \#\langle B \rangle's$.

One of the rules was used to derive w_{k+1} from w_k .

Check the effect of each rule:

$\langle A \rangle \rightarrow a\langle S \rangle \mid a$ no change

$\langle B \rangle \rightarrow b\langle S \rangle \mid b$ no change

$\langle S \rangle \rightarrow a\langle B \rangle \mid b\langle A \rangle$ each sum increases by 1

$\langle A \rangle \rightarrow b\langle A \rangle\langle A \rangle$ each sum increases by 1

$\langle B \rangle \rightarrow a\langle B \rangle\langle B \rangle$ each sum increases by 1

Thus $L(G) \subseteq \{w \in \{a,b\}^+ \mid \#a's = \#b's \text{ in } w\}$

$\langle S \rangle \rightarrow a\langle B \rangle \mid b\langle A \rangle$

$\langle A \rangle \rightarrow a\langle S \rangle \mid b\langle A \rangle\langle A \rangle \mid a$

$\langle B \rangle \rightarrow b\langle S \rangle \mid a\langle B \rangle\langle B \rangle \mid b$

Now to show $L(G) \supseteq \{w \in \{a,b\}^+ \mid \#a's = \#b's \text{ in } w\}$
it suffices to show that

1) if $w \in \{w \mid \#a's = \#b's\}$, then $\langle S \rangle \xRightarrow{*} w$

The proof is by induction on the length of w .

Once again, we aim for something stronger

(this provides us with a stronger induction hypothesis).

So in addition to 1) above, we also show

2) if $w \in \{w \mid \#a's = \#b's + 1\}$, then $\langle A \rangle \xRightarrow{*} w$

3) if $w \in \{w \mid \#a's + 1 = \#b's\}$, then $\langle B \rangle \xRightarrow{*} w$

Base cases: length of w is either 1 or 2.

If the length is 1, then either 2) or 3) applies.

2) $w = a$ the derivation is: $\langle A \rangle \Rightarrow a$

3) $w = b$ the derivation is: $\langle B \rangle \Rightarrow b$

If the length is 2, then 1) applies and $w = ab$ or $w = ba$.

$w = ab$ the derivation is $\langle S \rangle \Rightarrow a\langle B \rangle \Rightarrow ab$

$w = ba$ the derivation is $\langle S \rangle \Rightarrow b\langle A \rangle \Rightarrow ba$

$$\langle S \rangle \rightarrow a\langle B \rangle \mid b\langle A \rangle$$

$$\langle A \rangle \rightarrow a\langle S \rangle \mid b\langle A \rangle\langle A \rangle \mid a$$

$$\langle B \rangle \rightarrow b\langle S \rangle \mid a\langle B \rangle\langle B \rangle \mid b$$

Induction step: Assume 1), 2), and 3) whenever the length of w is k or less, and let $|w| = k+1$.

Then $w = aw'$ or $w = bw''$, where $|w'| = |w''| = k$.

Suppose $w = aw'$ and $w \in \{w \mid \#a's = \#b's\}$.

Then $w' \in \{w \mid \#a's + 1 = \#b's\}$

By the induction hypothesis, $\langle B \rangle \xRightarrow{*} w'$.

But then $\langle S \rangle \Rightarrow a\langle B \rangle \xRightarrow{*} aw' = w$.

Now suppose $w \in \{w \mid \#a's = \#b's + 1\}$.

Then $w' \in \{w \mid \#a's = \#b's\}$

By the induction hypothesis, $\langle S \rangle \xRightarrow{*} w'$.

But then $\langle A \rangle \Rightarrow a\langle S \rangle \xRightarrow{*} aw' = w$.

Finally suppose $w \in \{w \mid \#a's + 1 = \#b's\}$.

Then $w' \in \{w \mid \#a's + 2 = \#b's\}$

Note that $w' = xy$; both x and y have one more b than a .

By the induction hypothesis, $\langle B \rangle \xRightarrow{*} x$, and $\langle B \rangle \xRightarrow{*} y$.

But then $\langle B \rangle \Rightarrow a\langle B \rangle\langle B \rangle \xRightarrow{*} ax\langle B \rangle \Rightarrow axy = w$.

Finally, the case for $w = bw''$ is similar. □