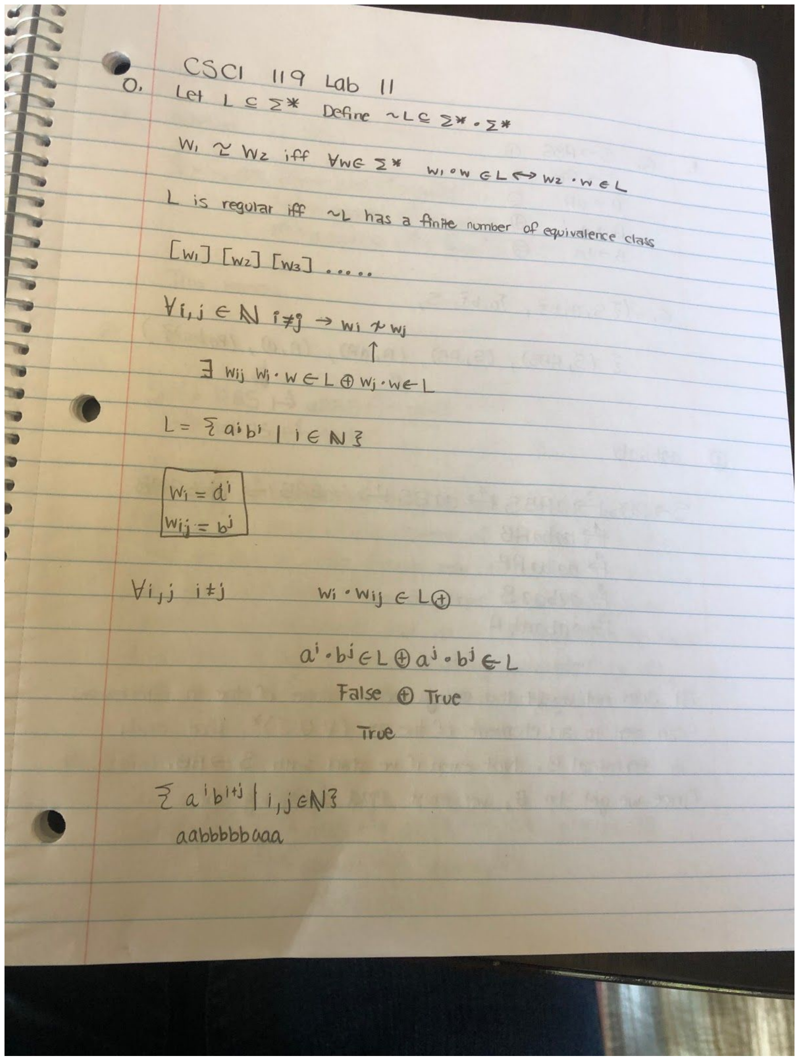
Madhuri Pyreddy
20 November 2019

CSCI 119 Lab 11

1. Question 0. Use the Myhill-Nerode Theorem to prove that the language {aibi+jaj∣i,j∈N} is not regular.



CSCI 119 Lab 11

0. Let $L \subseteq \Sigma^*$ Define $\sim_L \subseteq \Sigma^* \cdot \Sigma^*$

$w_1 \simeq w_2$ iff $\forall w \in \Sigma^*$ $w_1 \circ w \in L \leftrightarrow w_2 \cdot w \in L$

$L$ is regular iff $\sim_L$ has a finite number of equivalence class

$[w_1]\ [w_2]\ [w_3]$ .....

$\forall i,j \in \mathbb{N}\ i \neq j \rightarrow w_i \not\sim w_j$

$\exists\ w_{ij}\ w_i \cdot w \in L \oplus w_j \cdot w \in L$

$L = \{a^i b^i \mid i \in \mathbb{N}\}$

$\boxed{\begin{array}{l} w_i = a^i \\ w_{ij} = b^j \end{array}}$

$\forall i,j\ i \neq j$ $\quad w_i \cdot w_{ij} \in L \oplus$

$a^i \cdot b^j \in L \oplus a^j \cdot b^j \in L$

False $\oplus$ True

True

$\{a^i b^{i+j} \mid i,j \in \mathbb{N}\}$

aabbbbbaaa

2. Question 1. Consider the following context-free grammar G1, with start symbol S:

   S⟶ABS∣AB
   A⟶aA∣a
   B⟶bA.

   Which of the following strings are in L(G1) and which are not? Provide derivations or parse trees for those that are in L(G1) and reasons for those that are not:

aabaab
aaaaba
aabbaa
abaaba

1. $G_1$ $S \to ABS$ ①
$S \to AB$ ②
$A \to aA$ ③
$A \to a$ ④
$B \to ba$ ⑤

$G_1 = (\{S, A, B\}, \{a, b\}, S,$

$\{(S, ABS), (S, AB), (A, aA), (A, a), (B, ba)\})$

$R$

① aabaab

$S \to ABS \overset{3}{\mapsto} aABS \overset{4}{\mapsto} aaBS \overset{1}{\mapsto} aaBAB \overset{5}{\mapsto} aabAAB$
$\overset{4}{\mapsto} aabaAB$
$\overset{4}{\mapsto} aabaAB$
$\overset{4}{\mapsto} aabaaB$
$\overset{5}{\mapsto} aabaabA$

G1 does not accept this string because none of the non-terminals can step to an element of the set $(V \cup \Sigma)^*$ that ends in terminal B. Not even if we start with $S \to AB$. Once we get to B, we must read something else.

② aaaaba

$$S \xmapsto{2} AB \xmapsto{3} aAB \xmapsto{3} aaAB \xmapsto{3} aaaAB \xmapsto{4} aaaaB \xmapsto{5} aaaabA \xmapsto{4} \boxed{aaaaba}$$

This works.

③ aabbaa

$$S \xmapsto{1} ABS \xmapsto{3} aABS \xmapsto{4} aaBS \xmapsto{5} aabAS \xmapsto{4} aabaS$$

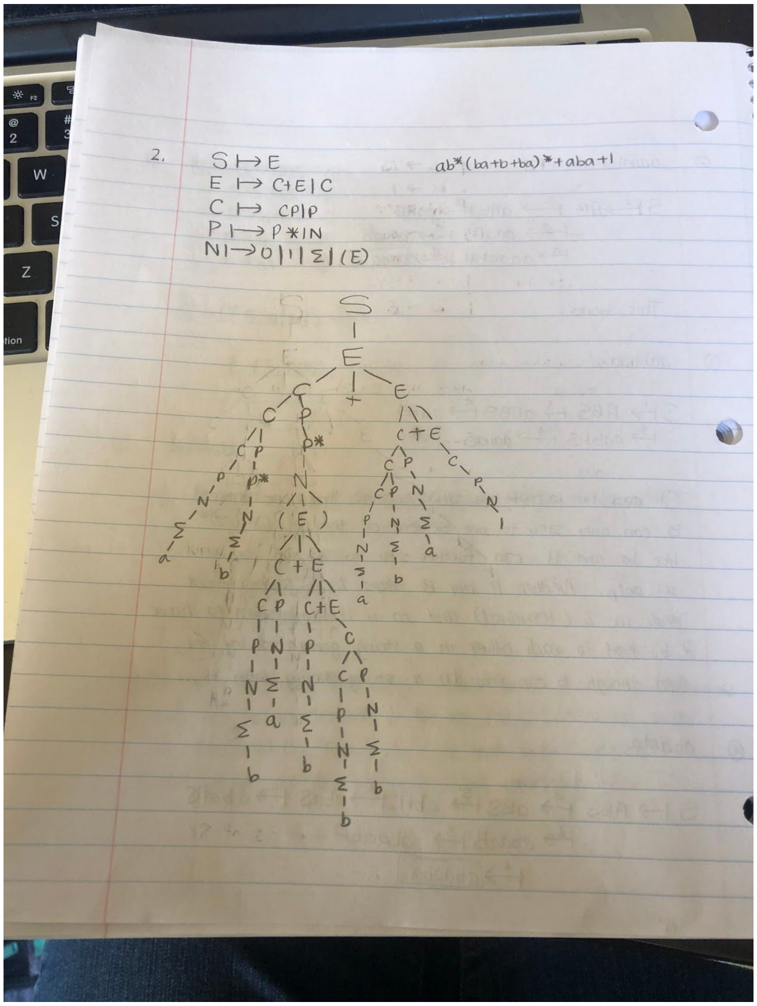
G1 does not accept this string because the non-terminal B can only step to one element of the set $(V \cup \Sigma)^*$ like ba and A can further step to an aA or an a only. Neither A nor B steps to a string that ends in b (terminal) and so it is impossible to have 2 b's next to each other in a string generated by G1, even though B can stop to a string starting with b.

④ abaaba

$$S \xmapsto{1} ABS \xmapsto{4} aBS \xmapsto{5} abAS \xmapsto{4} abaS \xmapsto{2} abaAS \xmapsto{4} abaaB \xmapsto{5} abaabA \xmapsto{4} \boxed{abaaba}$$

3. Question 2. Let Σ be a finite alphabet, and define Σ′=Σ∪{0,1,+,∗,(,)}. Find a context-free grammar G2 with Σ′ as the set of nonterminals which generates exactly the regular expressions over the alphabet Σ. Your grammar should be such that regular expressions are parsed unambiguously, where

   ∗ is postfix and has the highest precedence
   concatenation is left-associative and has the next highest precedence
   + is right-associative and has the lowest precedence.
   Show the derivation or parse tree of the string ab∗(ba+b+bb)∗+aba+1 in G2.




4. Question 3. Find a context-free grammar G3 that generates exactly the same strings as the regular expression at the end of Question 2.

3. $S \to X \mid Y \mid Z$

$X \to AB$

$A \to aU$

$U \to Ub \mid \varepsilon$

$B \to BV \mid \varepsilon$

$V \to L \mid M \mid N$

$L \to TS$

$M \to T$

$N \to TT$

$S \to a$

$T \to b$

$Y \to STS$

$Z \to \varepsilon$

4. $\{a^i b^j a^i \mid i \geq 0, j \geq 1\}$

$A \to B_1 B_2$ or $A \to a$, where $B_1, B_2 \in V$ and $a \in \Sigma$

Rules for Chomsky Form

if we have a rule $S \to A$

we can't just say $S \to AB$

and have $B \to \varepsilon$

Since we have to go to a terminal symbol and $\varepsilon$ is not a terminal.

We do global change of replacing all S's with A's and then S is the start symbol but ~~it~~ disappears so we have to make A the start symbol.

Suppose $S \to A$

and $S \to \varepsilon \mid A\ B$

but $A \to C\ S\ D$ then to get rid of S's there are 2 possibilities: $A \to CD$ and $A \to CABD$

5. Question 4. A CFG G=(V,Σ,S,R) is in Chomsky normal form if every rule in R is of the form A→B1B2 or A→a, where B1,B2∈V and a∈Σ. Give a grammar in Chomsky normal form that generates the language {aibjai∣i≥0,j≥1}.
   Question 5. Let Σ={a,b}. Prove that the CFG G5 with rules

   S→aSb∣bSa∣SS∣ε

   generates the set of all strings in Σ∗ with an equal number of as and bs. Do this by defining two recursive functions

$a,b:\Sigma* \to N$

such that a(w) is the number of as in w and b(w) is the number of bs in w and finding and proving the appropriate condition on w and its prefixes, as we did with the balanced-parentheses grammar.

3.

| | |
|---|---|
| $S \to X \mid Y \mid Z$ | $L \to TS$ |
| $X \to AB$ | $M \to T$ |
| $A \to aU$ | $N \to TT$ |
| $U \to Ub \mid \varepsilon$ | $S \to a$ |
| $B \to BV \mid \varepsilon$ | $T \to b$ |
| $V \to L \mid M \mid N$ | $Y \to STS$ |
| | $Z \to \varepsilon$ |

4. $\{a^i b^j a^i \mid i \geq 0, j \geq 1\}$

$A \to B_1 B_2$ or $A \to a$, where $B_1, B_2 \in V$ and $a \in \Sigma$

Rules for Chomsky Form

if we have a rule $S \to A$

we can't just say $S \to AB$

and have $B \to \varepsilon$

Since we have to go to a terminal symbol and $\varepsilon$ is not a terminal.

We do global change of replacing all S's with A's and then S is the start symbol but ~~it~~ disappears so we have to make A the start symbol.

Suppose $S \to A$

and $S \to \varepsilon \mid A\ B$

but $A \to C\ S\ D$ then to get rid of S's there are 2 possibilities: $A \to CD$ and $A \to CABD$

4. Strings accepted by given language :-

b, aba, aabaa, aaa baaa, .....
bb, abba, aabbaa, aaa bbaaa .....
bbb , abbba, aabbbaa, aaabbbaaa

Not a regular language can't write a FSM that accepts this language because we can't remember the a's we read before b's with finite number of states, but it's Context Free Language.

Required CNF is

Step 1: removing null rules :-

$S \rightarrow aSa \mid B$ | $B' \rightarrow bB' \mid b$
$B \rightarrow bB' \mid b$

Step 2: removing unit rules
remove $S \rightarrow B$

$S \rightarrow aSa \mid bB' \mid b$
$B \rightarrow bB' \mid b$
$B' \rightarrow bB' \mid b$

Step 3: removing more than 1 terminal on RHS | 1 terminal with non-terminals

$S \rightarrow ASA \mid ZB' \mid b$
$B \rightarrow ZB' \mid b$
$B' \rightarrow ZB' \mid b$
$A \rightarrow a$
$Z \rightarrow b$

4\. Step 4 ε removing more than 2 non-terminal on the RHS

$S \to AY \mid ZB' \mid b$
$B \to ZB' \mid b$
$B' \to ZB' \mid b$
$A \to a$
$Z \to b$
$Y \to SA$

5\. $S \to aSb \mid bSa \mid SS \mid \varepsilon$

$V = \{S\}$ $\Sigma = \{a, b\}$

the claim is that strings with equal number of a's & b's

Define $a, b : (V \cup \Sigma)^* \to \mathbb{N}$ by

| | |
|---|---|
| $a(\varepsilon) = 0$ | $b(\varepsilon) = 0$ |
| $a(aw) = 1 + a(w)$ | $b(aw) = b(w)$ |
| $a(bw) = a(w)$ | $b(bw) = 1 + b(w)$ |
| $a(Sw) = a(w)$ | $b(Sw) = b(w)$ |

Theorem: for all $w \in \Sigma^*$

$S \overset{*}{\Rightarrow} w$, if there are equal number of a's and b's in $w$.

6. Question 5. Let Σ={a,b}. Prove that the CFG G5 with rules

   S⟶aSb∣bSa∣SS∣ε

   generates the set of all strings in Σ∗ with an equal number of as and bs. Do this by defining two recursive functions

   a,b:Σ∗⟶N

   such that a(w) is the number of as in w and b(w) is the number of bs in w and finding and proving the appropriate condition on w and its prefixes, as we did with the balanced-parentheses grammar.

4. Step 4 & removing more than 2 non-terminal on the RHS:

$S \to AY \mid ZB' \mid b$

$B \to ZB' \mid b$

$B' \to ZB' \mid b$

$A \to a$

$Z \to b$

$Y \to SA$

5. $S \to aSb \mid bSa \mid SS \mid \varepsilon$

$V = \{S\}$ $\Sigma = \{a, b\}$

the claim is that strings with equal number of a's & b's

Define $a, b : (V \cup \Sigma)^* \to \mathbb{N}$ by

| | |
|---|---|
| $a(\varepsilon) = 0$ | $b(\varepsilon) = 0$ |
| $a(a\omega) = 1 + a(\omega)$ | $b(a\omega) = b(\omega)$ |
| $a(b\omega) = a(\omega)$ | $b(b\omega) = 1 + b(\omega)$ |
| $a(S\omega) = a(\omega)$ | $b(S\omega) = b(\omega)$ |

Theorem: for all $\omega \in \Sigma^*$

$S \overset{*}{\Rightarrow} \omega$. if there are equal number of a's and b's in $\omega$.

5. Lemma 1: For all $w \in (V \cup \Sigma)^*$, if $S \overset{*}{\mapsto} w$, then $w$ is a string with equal number of $a$s and $b$s.

Proof: By induction on the derivation on $S \overset{*}{\mapsto} w$. 2 rules are:

$$\frac{}{\alpha \overset{*}{\mapsto} \alpha} \text{(S1)} \qquad \frac{\alpha \overset{*}{\mapsto} \beta \quad \beta \mapsto \gamma}{\alpha \overset{*}{\mapsto} \gamma} \text{(S2)}$$

Case S1: $w = S$

Hence $a(w) = a(S) = 0 = b(S) = b(w)$, so $w$ satisfies requirement.

Case S2: $S \overset{*}{\mapsto} w'$, and $w' \mapsto w$

By IH on assumption that $S \overset{*}{\mapsto} w'$, we get that $w'$ satisfies consequent of left-to-right lemma we are proving.

By rule of $\mapsto$, rule $V \to \gamma$ in the grammar and 2 strings $\alpha, \beta \in (V \cup \Sigma)^*$ such that $w' = \alpha V \beta$

$w' = \alpha \gamma \beta$

But $V = S$ is only non-terminal in G5 and it follows from the IH and it follows from the IH that $a(w') = b(w') \Rightarrow a(\alpha V \beta) = b(\alpha V \beta)$

$\Rightarrow a(\alpha S \beta) = b(\alpha S \beta)$

But $(A+)$ and $(B+)$ state: –

$a(\alpha \circ S\beta) = a(\alpha) + a(S\beta)$

$= a(\alpha) + a(\beta)$

5.

and $b(\alpha \circ S \beta) = b(\alpha) + b(S\beta)$

$= b(\alpha) + b(\beta)$

(by associativity of $\circ$)

$(*) \rightarrow a(\alpha) + a(\beta) = b(\alpha) + b(\beta)$

Case (R1): $w = \alpha a S b \beta$

Using (A+) (B+) & $(*)$

$a(\alpha a S b \beta) = a(\alpha) + a(aSb\beta)$

$= a(\alpha) + a(aSb) + a(\beta)$

(by associativity of $\circ$)

(by associativity of $\circ$)

$= a(\alpha) + a(a) + a(Sb) + a(\beta)$

$= a(\alpha) + a(a) + a(S) + a(b) + a(\beta)$

(by associativity of $\circ$)

(by associativity of $\circ$)

$= a(\alpha) + 1 + 0 + 0 + a(\beta)$

$= b(\alpha) + 0 + 0 + 1 + b(\beta)$

$= b(\alpha) + b(aSb) + b(\beta)$

$= b(\alpha a S b \beta)$

Case (R2): $w = \alpha b S a \beta$

Using (A+) (B+) and $(*)$

$a(\alpha b S a \beta) = a(\alpha) + 1 + a(\beta)$

$= b(\alpha) + 1 + b(\beta)$

$= b(\alpha b S a \beta)$

5. Case (R3) and (R4): $\omega = \alpha SS\beta$ and $\omega = \alpha\beta$.
there are no extra a's and b's. So consequent follows even more easily from (*), along same lines as in (R1). QED

<u>Lemma 2:</u> For $\omega \in \{a,b\}^*$, if $a(\omega) = b(\omega)$, then $S \xrightarrow{*} \omega$
By string induction on $|w|$. Assume that (L2) is true for all strings $w$ $|w| < n$

<u>Case 1:</u> $n=0$

(P1) $\Rightarrow$ if $\alpha \to \beta$ then $\alpha \xrightarrow{*} \beta$
(P2) $\Rightarrow$ $\vdash\!\xrightarrow{*}$ is transitive
(P3) $\Rightarrow$ if $\gamma \vdash\!\xrightarrow{*} \gamma_2$ then for any strings $\alpha$ and $\beta$, we have $\alpha\gamma_1\beta \vdash\!\xrightarrow{*} \alpha\gamma_2\beta$
where $\alpha, \beta, \gamma \in (V \cup \Sigma)^*$

if $n=0$, then $w=\varepsilon$ and by (R4) $S \vdash\!\xrightarrow{*} \varepsilon$ because $S \to \varepsilon$ (P1)

<u>Case 2;</u> $n \geq 1$ and $\omega$ has a deeper prefix $u$ such that $a(u) = b(u)$
Let $v$ be such that $w = u \cdot v$

<u>Case 3:</u> $a(u) = b(u)$ & $a(v) = b(v)$
It follows that $a(w) = b(w)$ by facts of (A+) and (B+) that letter is true.

5. We conclude that in both cases $S \overset{*}{\mapsto} u$, $S \overset{*}{\mapsto} v$ since lengths of $u$ and $v$ are smaller than $n$ in accordance with strong IH.

Use (R3) & (P3) to get:

$$S \mapsto SS \overset{*}{\mapsto} uS \overset{*}{\mapsto} u \cdot v$$

It follows from (P1) and 2 applications of (P2) that $S \overset{*}{\mapsto} w$

Case 3: $n \geq 1$ and $w$ doesn't have proper prefix such that $a(w) = b(u)$

But $w$ can't end there otherwise $a(w) = 1$ and $b(w) = 0$ or $a(w) = 0$ as the case maybe so $|w| \geq 2$

But if $w$ started with an a then it has to end with a b, otherwise it ended in an a.

$w = ua$,

then $u$ would be proper prefix of $w$ so

$a(w) = 1 + a(u)$ and $b(w) = b(u)$

$a(w) > b(w)$ (violates condition of balanced a's & b's of $w$)

And similarly, if $w$ started with a b then it has to end with a.

5. Finally, to show w's is balanced in a's and b's

$$a(w') = a(w) - 1$$
$$b(w) - 1 = b(w')$$

Since $|w'| < n$, our strong IH applies here, hence $S \overset{*}{\mapsto} w'$

∴ using (R1)/(R2) and (P3), we get

$$S \mapsto aSb \overset{*}{\mapsto} aw'b$$
$$\text{or } S \mapsto bSa \overset{*}{\mapsto} bw'a$$

by (P1) and (P2) $S \overset{*}{\mapsto} w$ $\boxed{\text{QED}}$

6. $G_6: S \longrightarrow bS \mid Sa \mid aSb \mid \varepsilon$

$L(G_6) = \{\varepsilon, a, b, aa, bb, ab, ba, aaa, aab, aba, baa, abb, bab, bba, bbb\}$

$L(G_6)$ = all strings matching R.E. $(a+b)^*$

Lemma 1: $\forall w \in (V \cup \Sigma)^*$, if $S \overset{*}{\mapsto} w$, then $w$ is any string over $\Sigma$ including empty string.

Proof: By induction on derivation of $S \overset{*}{\mapsto} w$, 2 rules of which are

$$\frac{}{\alpha \overset{*}{\mapsto} \alpha}\ (S1) \qquad \frac{\alpha \overset{*}{\mapsto} \beta \quad \beta \mapsto \gamma}{\alpha \overset{*}{\mapsto} \gamma}\ (S2)$$

7. Question 6. What set is generated by the following grammar?

$S \rightarrow bS \mid Sa \mid aSb \mid \varepsilon$

Give a proof (in the same style as in Question 5).

5. Finally, to show w's is balanced in a's and b's

$$a(w') = a(w) - 1$$
$$b(w) - 1 = b(w')$$

Since $|w'| < n$, our strong IH applies here, hence $S \overset{*}{\mapsto} w'$

$\therefore$ using (R1) / (R2) and (P3), we get

$$S \mapsto aSb \overset{*}{\mapsto} aw'b$$
or $$S \mapsto bSa \overset{*}{\mapsto} bw'a$$

by (P1) and (P2) $S \overset{*}{\mapsto} w$ $\boxed{QED}$

6. $G_6 : S \longrightarrow bS \mid Sa \mid aSb \mid \varepsilon$

$L(G_6) = \{ \varepsilon, a, b, aa, bb, ab, ba, aaa, aab, aba, baa, abb, bab, bba, bbb \}$

$L(G_6)$ = all strings matching R.E.

$$(a+b)^*$$

<u>Lemma 1</u>: $\forall w \in (V \cup \Sigma)^*$, if $S \overset{*}{\mapsto} w$, then $w$ is any string over $\Sigma$ including empty string.

<u>Proof</u>: By induction on derivation of $S \overset{*}{\mapsto} w$, 2 rules of which are

$$\frac{}{\alpha \overset{*}{\mapsto} \alpha} \; (S1) \qquad \frac{\alpha \overset{*}{\mapsto} \beta \quad \beta \mapsto \gamma}{\alpha \overset{*}{\mapsto} \gamma} \; (S2)$$

6. <u>Case S1:</u> $w = S$

By rules of $G_6$, $S$ can be any string over $\Sigma$ or in $\Sigma^*$ including the empty string, so $w$ satisfies the consequent.

<u>Case S2:</u> $S \overset{*}{\longmapsto} w'$ and $w' \longmapsto w$

Hence being a smaller string, we can apply the IH on $w'$ to get that $w'$ satisfies consequent.

<u>new goal:</u> $w$ satisfies consequent as well. It's an invariant of step solution

Define $\longmapsto : (V \cup \Sigma)^* \to (V \cup \Sigma)^*$

in $\alpha, \beta, \gamma \in (V \cup \Sigma)^*$

and $(V, \gamma) \in R$ then

$V \to \gamma \qquad \alpha\ \beta \in (V \cup \Sigma)^*$

$\alpha\ V \beta \longmapsto \alpha \gamma \beta$

By this rule, if $V \to \gamma$ and $w' = \alpha V \beta$ then

$w = \alpha \gamma \beta$

But $V = S$ is only non-determinant in $G_6$ and it follows from the IH that $w = \alpha S \beta$ is any string in $\Sigma^*$

Now, we have 1 case to consider for each rule on $G_6$ to check whether $w$ satisfies consequent.

6. Case (R1): $w = \alpha b S B$

But if b is included anywhere inside $w$, it still is a string drawn from $\Sigma^*$,

Hence, it satisfies consequent. ✓

Case (R2): $w = \alpha S a \beta$

But, if a is included anywhere inside $w'$, it still is a string drawn from $\Sigma^*$,

It satisfies consequent. ✓

Case (R3): $w = \alpha a S b \beta$

But, if either a or b or both can be included anywhere inside $w'$, no matter in what side, it still is a string drawn from $\Sigma^*$. Also, it satisfies consequent. ✓

Case (R4): $w = \alpha \beta$

If $w'$ is already a string of a's and b's over $\Sigma$ including the empty string, and S can step to any string in $\Sigma^*$ then both $\alpha$ & $\beta$ can step to any string in $\Sigma^*$ either individually or together. Also, it satisfies the consequent, completing the proof. $\boxed{QED}$

Lemma 2: If $w \in (V \cup \Sigma)^*$ be arbitrary, $n = |w|$ and for the lemma hold for all strings of length $< n$ (by Strong IH)

6. Proper prefix: if $w = u \cdot v$, then $u$ is a prefix of $w$ but if $0 < |u| < |w|$, then $u$ is proper prefix of $w$.

Properties of multistep relation

(P1) $\Rightarrow$ if $\alpha \to \beta$ then $\alpha \xmapsto{*} \beta$

(P2) $\Rightarrow$ $\xmapsto{*}$ is transitive

(P3) $\Rightarrow$ if $\gamma \xmapsto{*} \gamma_2$, then for any strings $\alpha$ and $\beta$ we have $\alpha\gamma, \beta \xmapsto{*} \alpha \gamma_2 \beta$ where $\alpha, \beta, \gamma \in (V \cup \Sigma)^*$

Case 1: $n=0$

then $w = \varepsilon$ and by (R4) $S \xmapsto{*} w = \varepsilon$ because $S \to \varepsilon$ (P1)

Case 2: $n \geq 1$ and $w$ has proper prefix $u$ such that $u$ satisfies the antecedent. Let $v$ be such that $w = u \cdot v$

Claim: $u$ satisfies ancedent.

This is true by assumption and the latter follows from the fact that since $w$ is drawn from $\Sigma^*$, as well it follows that $w$ has to be drawn from $\Sigma^*$ as well because if $v$ is drawn from any set other than $\Sigma^*$, then that would violate the assumption that the whole string is drawn from $\Sigma^*$.

We apply strong IH to both $u$ and $V$ because the

6. lengths of both are smaller than , and we conclude that antecedent is true in both cases:

$$S \overset{*}{\longmapsto} u, \quad S \overset{*}{\longmapsto} v$$

Using (P3), we derived $u \cdot v$ from S except that we don't have a rule that says

$$S \rightarrow SS.$$

Therefore, there's no way to split $\omega$ such that S steps to concatenation of splits. Assumption is wrong, which disapproves lemma 2.

For all $w \in \Sigma^*$, if $S \overset{*}{\longmapsto} w$ then w is a string drawn from $\Sigma^*$, but not the other way around.