Assignment I1

Jens Kristian Refsgaard Nielsen

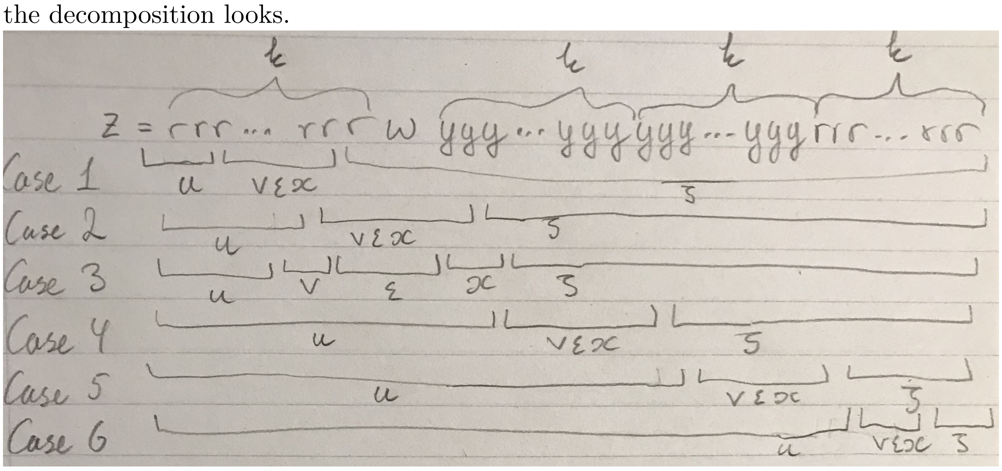
14. marts 2019

Question 1.

Claim: the language of the Spanish flag $L = \{r^n w y^{2n} r^n \mid n \geq 0\}$ is not context-free
Proof: we proceed by using the contraposition of the pumping lemma for context-free languages that states:

$$
\begin{aligned}
&X \subseteq \Sigma^* \text{ is not context free if:} \\
&\forall k \geq 1 \exists z \in X : |z| > k : \\
&\forall u, v, \epsilon, x, \zeta \in \Sigma^* : z = uv\epsilon x\zeta, |v\epsilon x| \leq k, |vx| > 0 : \\
&\exists i \geq 0 : uv^i\epsilon x^i\zeta \notin X
\end{aligned}
$$

So let $k \geq 1$ be given, we then choose $z = r^k w y^{2k} r^k$, and let a decomposition $u, v, \epsilon, x, \zeta \in \Sigma^*$: $z = uv\epsilon x\zeta, |v\epsilon x| \leq k, |vx| > 0$ be given. Now we proceed by looking at the possible cases of how the decomposition looks.



Case 1: $u = r^l, v\epsilon x = r^j, \zeta = r^{k-l-j} w y^{2k} r^k$ for $0 \leq l \leq k$ and $0 < j < k - l$
The condition $|vx| > 0$ forces atleast one of v and x to have atleast one r, so if we choose to pump with $i = 2$ we get atleast one more r in $v\epsilon x$ so $v^2\epsilon x^2 = r^p$ for $p > j$, consider then

$$uv^2\epsilon x^2\zeta = r^l r^p r^{k-l-j} w y^{2k} r^k = r^{k-j+p} w y^{2k} r^k \notin L$$

Because $p > j$ we get more r's in front of our expression than in the back, which is not allowed in L, concluding the case.

Case 2: If w is part of either v or x, simply pump $i = 2$ and we get two w's which our language doesnt support, completing the small case.

Case 3: $u = r^{k-l-i}, v = r^l, \epsilon = r^i w y^r, x = y^j, \zeta = y^{2k-j} r^k$ for $0 < l + i + 1 + r + j \leq k$ and $0 \leq i < k - l$ and $0 \leq r$
Note that either l or j is non zero so if we pump $i = 2$ we either get too many r's in the front or too many y's compared to the r's in ζ, showing that $uv^2\epsilon x^2\zeta \notin L$ completing the case. We have l and j not both zero, if

Case 4: $u = r^k w y^l, v\epsilon x = y^j, \zeta = y^{2k-j-l} r^k$ for $0 < j \leq k$ and $0 \leq l \leq k$
Once again the $|vx| > 0$ condition forces either v or x to have atleast one y, so when we pump $i = 2$ we get more than $2k$ y's in our decomposition and

$$uv^2\epsilon x^2\zeta = r^k w y^p r^k \overset{p>2k}{\notin} L$$

Case 5: $u = r^k w y^{2k-l}, v\epsilon x = y^l r^j, \zeta = r^{k-j}$ for $0 < l + j \leq k$
If either v or x has both a y and an r, we can pump $i = 2$ and get something clearly wrong, since we are only allowed to change symbols 3 times throughout the word, but if say $x = y^a r^b$ then $x^2 = y^a r^b y^a r^b$ for non zero a and b, and when we put them into the expression $uv^2\epsilon x^2\zeta$ we change symbols atleast 4 times, which is not allowed in L. So for the remainder of the case assume v and x only has at most 1 type of letter. Then either v has atleast one y or x has atleast one r, so when we pump $i = 2$ we have an inconsistency with the number of r's in the front i.e.

$$uv^2\epsilon x^2\zeta = r^k w y^p r^q, \qquad p \geq 2k, q \geq k$$

Where if v has a y we get $p > 2k$ or if x has an r we get $q > k$ and since we dont change the front most r's we get that $uv^2\epsilon x^2\zeta \notin L$ as wanted.

Case 6: $u = r^k w y^{2k} r^{k-l-j}, v\epsilon x = r^l, \zeta = r^j$ for $0 < l$ and $l + j \leq k$
Once more unto the breach, the condition $|vs| > 0$ saves the day, because then v or x has an r, and when we pump $i = 2$ we get too many r's in the last batch, compared to the first. $\square$

In conclusion we chose $z = r^k w y^{2k} r^k$ and applied the contraposition of the pumping lemma for context free languages, to show that $L = \{r^n w y^{2n} r^n \mid n \geq 0\} \ni z$ is not context free.

Question 2.

We wish to construct an unrestricted grammar to describe $L = \{r^n w y^{2n} r^n \mid n \geq 0\}$, consider the following

$$S \to rRYA \mid w \tag{0.1}$$
$$R \to w \mid rRY \tag{0.2}$$
$$Y \to yyZ \tag{0.3}$$
$$Zyy \to yyZ \tag{0.4}$$
$$ZA \to Ar \tag{0.5}$$
$$A \to \Lambda \tag{0.6}$$

The idea is to start with S and then we can just end with w to represent that word, if however we want an r we need to place 2 y's and finish with another r. To accomplish this we put a Y for every r in front, and later use the Y to place 2 y's and a Z that we in turn destroy with A to get the r's in the rear of the word, so the grammar accepts all of L. The grammar also doesnt accept any words not in L since the only variables that maps to a terminal without another variable are S and A. If we perfom the $S \to w$ we get the word $w \in L$ so no trouble. If we perfom the derivation $A \to \Lambda$ early, there will still be variables within the parse that now cannot be resolved and the parseing attempt fails.
Consider the following conceptual derivation of $r^n w y^{2n} r^n$

$$\begin{aligned}
S &\xrightarrow{(0.1)} rRYA \xrightarrow{(0.2),n-1} r^n RY^n A \\
&\xrightarrow{(0.3)} r^n wyyZY^{n-1}A \xrightarrow{(0.3),n-1} r^n w(yyZ)^n A \\
&\xrightarrow{(0.4)} r^n w(yyZ)^{n-2}yyyyZZA \xrightarrow{(0.4),n(n-1)/2} r^n 2y^{2n} Z^n A \\
&\xrightarrow{(0.5)} r^n 2y^{2n} Z^{n-1} Ar \xrightarrow{(0.5),n-1} r^n 2y^{2n} Ar^n \\
&\xrightarrow{(0.6)} r^n wy^{2n} r^n
\end{aligned}$$

Note that the $\xrightarrow{(0.4),n(n-1)/2}$ step means applying rule (0.4) on each Z the maximum amount of times possible, the last Z is already in place so we apply this step $\sum_{k=0}^{n-1} k = \frac{n(n-1)}{2}$ times. Note that we should start from the back since we dont have any rules allowing us to move the first of the two Z's in expressions like $yyZZyy$, but of course the sequencencing of the operations is unimportant for the purposes of this exercise, this observation is solely meant for writing an efficient parsing algorithm.

Question 3.

Claim: $\forall L_1 \in$DCFL, $\forall L_2 \in$REG : $L_1 \cap L_2 \in$DCFL
Proof: Let $L_1 \in$DCFL and $L_2 \in$REG, then we know that there exists a DPDA say $M = (Q_N, \Sigma_M, \Gamma_M, q_{M,0}, A_M, \delta_M)$ and a DFA say $N = (Q_N, \Sigma_N, q_{N,0}, A_N, \delta_N)$ such that $\mathcal{L}(M) = L_1$ and $\mathcal{L}(N) = L_2$. We now construct

$$M \times N = (Q_m \times Q_N, \Sigma_M \cup \Sigma_N, \Gamma_M, (q_{M,0}, q_{N,0}), A_M \times A_N, \delta)$$

Where δ is constructed by the following rules

$$\text{if } q_M \xrightarrow{a\ B/W} q'_M \text{ and } q_N \xrightarrow{a} q'_N \text{ then} \qquad (q_M, q_N) \xrightarrow{a\ B/W} (q'_M, q'_N) \tag{0.7}$$
$$\text{if } q_M \xrightarrow{\Lambda\ B/W} q'_M \text{ then} \qquad (q_M, q_N) \xrightarrow{a\ B/W} (q'_M, q_N) \tag{0.8}$$

We wish to prove that $M \times N$ is a DPDA, because then $\mathcal{L}(M \times N) = L_1 \cap L_2$ is a Deterministic Context Free Language, as wanted. We know from theorem Lecture 10 slide 6 that the intersection of a CFL and a REG is a CFL, so since DCFL$\subsetneq$CFL the theorem proves that $M \times N$ is a PDA, since we used exactly the same construction as the theorem. So we are left to prove that this construction complies with the added demands of a DPDA, recall that a PDA is a DPDA if the following two criterion are met

$$\forall q, a, A : |\delta(q, a, A)| \leq 1 \tag{0.9}$$
$$\text{at most of of } \delta(q, a, A) \text{ and } \delta(q, \Lambda, A) \text{ is non empty} \tag{0.10}$$

Assume for contradiction that this is not the case.
Part 1: If rule (0.9) is not true then we must have a tripple

$$((q_{(M,1)}, q_{(N,1)}), a, A) \in (Q_M \times Q_N) \times (\Sigma_M \cup \Sigma_N) \times \Gamma_M : |\delta((q_M, q_N), a, A)| > 1$$

This means that we have atleast two **different** pairs in the set, say $((q_{(M,2)}, q_{(N,2)}), B)$ and $((q_{(M,3)}, q_{(N,3)}), D)$.
We further subdivide in the cases $a = \Lambda$ and $a \neq \Lambda$.

Part 1.1: $a \neq \Lambda$
Now the rules state

$$\begin{aligned}
(q_{(M,1)}, q_{(N,1)}) &\xrightarrow{a\ A/B} (q_{(M,2)}, q_{(N,2)}) &\overset{(0.7)}{\iff} q_{(M,1)} \xrightarrow{a\ A/B} q_{(M,2)} \text{ and } q_{(N,1)} \xrightarrow{a} q_{(N,2)} \\
(q_{(M,1)}, q_{(N,1)}) &\xrightarrow{a\ A/D} (q_{(M,3)}, q_{(N,3)}) &\overset{(0.7)}{\iff} q_{(M,1)} \xrightarrow{a\ A/D} q_{(M,3)} \text{ and } q_{(N,1)} \xrightarrow{a} q_{(N,3)}
\end{aligned}$$

The implication in the above follows from (0.7) since we have full knowledge about the way δ was constructed.
Recall that N is a DFA so $q_{(N,1)} \xrightarrow{a} q_{(N,2)}$ and $q_{(N,1)} \xrightarrow{a} q_{(N,3)}$ implies that $q_{(N,2)} = q_{(N,3)}$.
Note also that

$$\begin{aligned}
q_{(M,1)} \xrightarrow{a\ A/B} q_{(M,2)} &\iff (q_{(M,2)}, B) \in \delta(q_{(M,1)}, a, A) \\
q_{(M,1)} \xrightarrow{a\ A/D} q_{(M,3)} &\iff (q_{(M,3)}, D) \in \delta(q_{(M,1)}, a, A)
\end{aligned}$$

But M is a DPDA, so $|\delta(q_{(M,1)}, a, A)| \leq 1$ and then we must have $(q_{(M,3)}, D) = (q_{(M,2)}, B) \iff q_{(M,2)} = q_{(M,3)} \wedge B = D$.
In summary we have shown that

$$((q_{(M,2)}, q_{(N,2)}), a, B) = ((q_{(M,3)}, q_{(N,3)}), a, D)$$

Which is a contradiction since we assumed that the two pairs were different.

Part 1.2: $a = \Lambda$
With $a = \Lambda$ we used (0.8) to get our pairs, and from the get go $q_{(N,1)} = q_{(N,2)} = q_{(N,3)}$

$$\begin{aligned}
(q_{(M,1)}, q_{(N,1)}) &\xrightarrow{\Lambda\ A/B} (q_{(M,2)}, q_{(N,1)}) &\overset{(0.8)}{\iff} q_{(M,1)} \xrightarrow{a\ A/B} q_{(M,2)} \\
(q_{(M,1)}, q_{(N,1)}) &\xrightarrow{\Lambda\ A/D} (q_{(M,3)}, q_{(N,1)}) &\overset{(0.8)}{\iff} q_{(M,1)} \xrightarrow{a\ A/D} q_{(M,3)}
\end{aligned}$$

Now a very similar argument to that of case 1.1 shows that $q_{M,2} = q_{M,3}$, due to (0.9) and we obtain the contradiction that the two pairs are not different.

Part 2: In part 1 we have seen that $M \times N$ satisfies the first DPDA criteria (0.9), now we move on to show the second criteria (0.10), we continue to use contradiciton.
If (0.10) does not hold we must have a tripple $((q_{(M,1)}, q_{(N,1)}), a, A)$ such that both $\delta((q_{(M,1)}, q_{(N,1)}), a, A)$ and $\delta((q_{(M,1)}, q_{(N,1)}), \Lambda, A)$ are non empty. We then take elements $((q_{(M,2)}, q_{(N,2)}), B) \in \delta((q_{(M,1)}, q_{(N,1)}), a, A)$ and $((q_{(M,3)}, q_{(N,3)}), D) \in \delta((q_{(M,1)}, q_{(N,1)}), \Lambda, A)$, schematically this is of course equivivalent to

$$\begin{aligned}
(q_{(M,1)}, q_{(N,1)}) &\xrightarrow{a\ A/B} (q_{(M,2)}, q_{(N,2)}) &\overset{(0.7)}{\iff} q_{(M,1)} \xrightarrow{a\ A/B} q_{(M,2)} \text{ and } q_{(N,2)} \xrightarrow{a} q_{(N,2)} \\
(q_{(M,1)}, q_{(N,1)}) &\xrightarrow{\Lambda\ A/D} (q_{(M,3)}, q_{(N,3)}) &\overset{(0.8)}{\iff} q_{(M,1)} \xrightarrow{a\ A/D} q_{(M,3)} \text{ and } q_{(N,1)} = q_{(N,3)}
\end{aligned}$$

We obtain our contradiction entirely from the fact that M is a DPDA, because the situation described above exactly means that both $\delta_M(q_{(M,1)}, a, A)$ and $\delta_M(q_{(M,1)}, \Lambda, A)$ each contain an element, which is not allowed for DPDA's by (0.10). □

Question 4.

Given that DCFL is closed under complements we wish to prove that for all $L_1 \in$DCFL and $L_2 \in$REG we have $L_1 \cup L_2 \in$DCFL.

Proof:
For $X \subseteq \Sigma^*$ denote by X^c the complement $\Sigma^* \setminus L$ and consider

$$(L_1 \cup L_2)^c \overset{\text{De Moivre}}{=} L_1^c \cap L_2^c \overset{Q3}{\in} \text{DCFL} \tag{0.11}$$

$$L_1 \cup L_2 = (L_1^c \cap L_2^c)^c \overset{(0.11)}{\in} \text{DCFL} \tag{0.12}$$

The "element of" statement of (0.11) is due to the result of question 3 applied to an intersection of a regular language and a deterministic context free language, since both of these categories are closed under complements. In (0.12) we once again use the closure property of DCFL's. $\square$