Regularitet og Automater: Assignment 6

Rasmus Dalsgaard (201605295)

May 2017

Introduction

In this assignment, we're tasked to use the pumping lemma on languages such as to show that the given language is not a CFL.
The exercise processed is [Martin][1] 6.2 (a):

$$L = \{a^i b^j c^k | i < j < k\}$$

Pumping lemma

The lemma for context-free languages states that if L is context-free and $n = 2^{p+1}$, then

- $\exists n > 0 : \forall \omega \in L$
- $|\omega| \geq n$
- $\exists v, w, x, y, z \in \Sigma^*$
- $\omega = vwxyz$
- $|wxy| \leq n$
- $|wy| > 0$
- $\forall m \geq 0 : vw^m xy^m z \in L$

To show that a language is not context-free, we must break out of the constraints in a mathematically valid fashion to reach an out-of-bounds state that invalidates the status of CFL such that

$$\forall m \geq 0 : vw^m xy^m z \notin L$$

[1] John Martin: Introduction to Languages and The Theory of Computation, Fourth Edition

Process

Considering that $L = \{a^i b^j c^k | i < j < k\}$ and $n > 0$ as per the lemma, let an arbitrary string $\omega \in L$.

In order to employ the pumping lemma, let an n-bound string be

$$\omega = a^n b^{n+1} c^{n+2} \in L$$

$$|\omega| = 3n + 3 \geq n$$

Splitting ω into $vwxyz$ such that $|wxy| \leq n$ and $|wy| > 0$, according to the lemma, we know that $|wy| > 0$ and that the precursive a^n fills out the n first characters of ω. A visual way of representing this would be writing out generated string from a random n value of, say, $n = 3$

$$\omega = a^n b^{n+1} c^{n+2} = \underbrace{aaa}_{\text{n}}\underbrace{bbbb}_{\text{n+1}}\underbrace{ccccc}_{\text{n+2}}$$

The reach of $|wxy| \leq n$ allows us to reach only a's, b's or c's, or an overlap of a maximum of two of these, but *not* all three symbols. For example

$$\omega = \overbrace{aaa}^{\leq \text{ n}}\overbrace{bbb}^{\leq \text{ n}} b \overbrace{ccc}^{\leq \text{n}} cc$$

$$\omega = a \underbrace{a \overbrace{abb}^{\leq \text{ n}} bbc}_{> |wxy|} cccc$$

Breaking stuff

Considering the structure of $L = \{a^i b^j c^k | i < j < k\}$ we know that if a^i is greater or equal to j or k then the language is not context-free. This effect can be obtained by utilizing the pumping formula

$$\omega = uv^m xy^m z \in L$$

Evaluating the syntax based on $|wxy| \leq n$ and looking at bracings above, we have the cases chalked up: either $wxy \in \omega$ contains a's, a'a and b's, just b's, b's and c's or merely c's. Visually clearer with a defined n, say $n = 3$, merely for examplified clarity:

Case 1: pure a

$\forall \epsilon : 0 < \epsilon \leq n | wxy = a^\epsilon$

$$\omega = a^n b^{n+1} c^{n+2} \rightarrow vwxyz = \underbrace{\quad}_{\text{v}}\underbrace{aaa}_{\text{wxy}}\underbrace{bbbbccccc}_{\text{z}}$$

Having $wxy = a^\epsilon$ and pumping $w^m xy^m$ by letting $m = n = 2^{p+1}$, we get that $a^{n\epsilon + n - \epsilon} b^{n+1} c^{n+2}$ breaks the syntax of $i < j < k$ because $i = n\epsilon + n - \epsilon > j = n + 1$ and thus the pumped $\omega \notin L$

Case 2: a and b

$wxy = a^\epsilon b^\delta$ where $0 < \epsilon < n$ and $n < \delta < n$ and $|wxy| \leq n$

$$\omega = a^n b^{n+1} c^{n+2} \to vwxyz = \underbrace{aa}_{\text{v}} \underbrace{abb}_{\text{wxy}} \underbrace{bbcccccc}_{\text{z}}$$

Having $wxy = a^\epsilon b^\delta$, we may utilize the procedure from above to pump $w^m xy^m$ with an $m = n$ and thus reach a state with $a^{n+(n-1)\cdot\epsilon}$ and/or $b^{2n\cdot\delta}$, smashing the valid syntax of $i < j < k$ and thus the pumped $\omega \notin L$.

Case 3: pure b

$\forall\epsilon : 0 < \epsilon \leq n | wxy = b^\epsilon$

$$\omega = a^n b^{n+1} c^{n+2} \to vwxyz = \underbrace{aaa}_{\text{v}} \underbrace{bbb}_{\text{wxy}} \underbrace{bccccc}_{\text{z}}$$

Having $wxy = b^\epsilon$ and pumping $w^m xy^m$ by letting $m = n = 2^{p+1}$, we get that $a^n b^{n\epsilon+n-\epsilon+1} c^{n+2}$ breaks the syntax of $i < j < k$ because $j = n\epsilon + n - \epsilon + 1 > k = n + 2$ and thus the pumped $\omega \notin L$

Case 4: b and c

$wxy = b^\epsilon c^\delta$ where $0 < \epsilon < n$ and $0 < \delta < n$ and $|wxy| \leq n$

$$\omega = a^n b^{n+1} c^{n+2} \to vwxyz = \underbrace{aaabb}_{\text{v}} \underbrace{bbc}_{\text{wxy}} \underbrace{cccc}_{\text{z}}$$

Considering b^ϵ and c^ϵ Let $m = 0$ such that $\omega = vw^0 xy^0 z$. This pumps down ω and removes $\epsilon + \delta$ symbols from wxy, such that $\omega = a^n b^{n+1-\epsilon} c^{n+2-\delta}$, which leaves us with an amount of b's and c's out of the syntatically valid range, and thus the pumped $\omega \notin L$.

Case 5: pure c

$\forall\epsilon : 0 < \epsilon \leq n | wxy = c^\epsilon$

$$\omega = a^n b^{n+1} c^{n+2} \to vwxyz = \underbrace{aaabbbb}_{\text{v}} \underbrace{ccc}_{\text{wxy}} \underbrace{cc}_{\text{z}}$$

Having $wxy = c^\epsilon$ and pumping $w^m xy^m$ by letting $m = 0$, we get that $a^n b^{n+1} c^{n+2-\epsilon}$ breaks the syntax of $i < j < k$ because $|wxy| \geq 1$ and by removing at least that many c's we get that $j = n + 1 \geq k = n + 2 - \epsilon$ and thus the pumped $\omega \notin L$

Conclusion

$L = \{a^i b^j c^k | i < j < k\}$ is pumped out of bounds by $m = n$ whenever $c \notin L$ and $m = 0$ when $c \in L$, and is not a context-free language.