Parsing Beyond CFG
Homework 6: TAL, formal properties

Laura Kallmeyer

Heinrich-Heine-Universität Düsseldorf

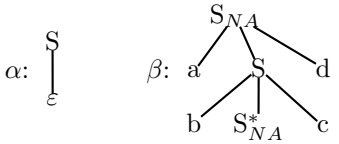
Question 1 (Pumping Lemma)

$L_4 = \{a^n b^n c^n d^n \mid n \geq 0\}$, $L_5 = \{a^n b^n c^n d^n e^n \mid n \geq 0\}$

1. *Give a TAG generating L_4.*
2. *Show that L_5 is not a TAL using the weak pumping lemma.*
   Hint: Consider the word $w = a^{c+1}b^{c+1}c^{c+1}d^{c+1}e^{c+1}$ with c being the constant from the pumping lemma.

Solution:

1. TAG for L_4:



2. Assume that L_5 is a TAL and satisfies the weak pumping lemma with some constant c. Take $w = a^{c+1}b^{c+1}c^{c+1}d^{c+1}e^{c+1}$. According to the pumping lemma one can find $w_1, \ldots w_4$, at least one of them not empty, such that they can be inserted repeatedly at four positions into w yielding a new word in L_5. At least one of the $w_1, \ldots w_4$ must contain two different terminal symbols since they altogether must contain equal numbers of as, bs, cs, ds and es. Then, when doing a second insertion of the $w_1, \ldots w_4$, the as, bs, cs, ds and es get mixed and the resulting word is not in L_5. Contradiction.

Question 2 (Closure Properties)

1. *Show that $L_{MIX5} = \{w \mid w \in \{a,b,c,d,e\}^*, |w|_a = |w|_b = |w|_c = |w|_d = |w|_e\}$ is not a TAL.*
   Hint: Use the closure of TALs under intersection with regular language and the result about L_5 shown above.
2. *Show that $L = \{w \mid w \in \{a,b,c,d,e,f,g\}^*, |w|_a = |w|_b = |w|_c = |w|_d = |w|_e\}$ is not a TAL.*
   Hint: Use the closure of TALs under homomorphisms and the result about L_{MIX5} shown in 1.

Solution:

1. We assume that $L_{MIX5} = \{w \mid w \in \{a,b,c,d,e\}^*, |w|_a = |w|_b = |w|_c = |w|_d = |w|_e\}$ is a TAL.
   Then $L_{MIX5} \cap L(a^*b^*c^*d^*e^*) = L_5$ must also be a TAL.
   Contradiction, consequently L_{MIX5} is not a TAL.
2. We assume that $L = \{w \mid w \in \{a,b,c,d,e,f,g\}^*, |w|_a = |w|_b = |w|_c = |w|_d = |w|_e\}$ is a TAL.
   Then its image under a homomorphism h with $h(a) = a, h(b) = b, h(c) = c, h(d) = d, h(e) = e, h(f) = \varepsilon, h(g) = \varepsilon$ must also be a TAL.
   Contradiction since $h(L) = L_{MIX5}$ is not a TAL. Consequently, L is not a TAL.