Parsing

Homework 5 (LL(1)), due 08 June 2020

Laura Kallmeyer

SS 2020, Heinrich-Heine-Universität Düsseldorf

Question 1 (Shift-reduce parsing)

Consider the deduction-based definition of shift-reduce parsing:

- *Item form* $[\Gamma, i]$ *(w has been shifted up to position i).*
- *Axiom:* $\dfrac{}{[\epsilon, 0]}$
- *Shift:* $\dfrac{[\Gamma, i]}{[\Gamma a, i+1]} \quad w_{i+1} = a$
- *Reduce:* $\dfrac{[\Gamma \alpha, i]}{[\Gamma A, i]} \quad A \to \alpha \in P$
- *Goal item* $[S, n]$.

Show the soundness of the algorithm, i.e., show that:
(S) If $[\Gamma, i]$ *can be deduced then* $\Gamma \stackrel{*}{\Rightarrow} w_1 \dots w_i$ *holds.*
This can be shown with an induction over the deduction rules. In other words, you have to

1. *Show that (S) holds for any axiom.*
2. *Show for the shift rule that, assuming that (S) holds for the antecedent item, (S) necessarily also holds for the consequent item.*
3. *Show the same for reduce.*

Note that $w_1 \dots w_0$ *is considered to be the empty word preceding the first terminal in the input.*

Solution:

- Axiom: $[\varepsilon, 0]$ holds and the part of the input from position 0 to position 0 is just ε. Therefore, $\varepsilon \stackrel{*}{\Rightarrow} w_1 \dots w_0 = \varepsilon$ holds trivially.
- Reduce: We have to show that, assuming that our claim holds for the antecedent item $[\Gamma\alpha, i]$ of a reduce rule, it also holds for the consequent item $[\Gamma A, i]$. Because of our induction assumption, we know that $\Gamma\alpha \stackrel{*}{\Rightarrow} w_1 \dots w_i$ and since this reduction was possible, it follows that $A \to \alpha \in P$ (side condition). Consequently $\Gamma A \stackrel{A \to \alpha}{\Rightarrow} \Gamma\alpha \stackrel{*}{\Rightarrow} w_1 \dots w_i$ and therefore, more generally, $\Gamma A \stackrel{*}{\Rightarrow} w_1 \dots w_i$.
- Shift: We have to show that, assuming that our claim holds for the antecedent item $[\Gamma, i]$ of a shift rule, it also holds for the consequent item $[\Gamma a, i+1]$. The side condition tells us that $a = w_{i+1}$, and our induction assumption yields $\Gamma \stackrel{*}{\Rightarrow} w_1 \dots w_i$. If we append the terminal a to both sides in this derivation, we obtain $\Gamma a \stackrel{*}{\Rightarrow} w_1 \dots w_i w_{i+1}$, which holds trivially.

Since all items generated by the parser are either the axiom or obtained from the axiom by a sequence of shift/reduce steps, every item necessarily satisfies our soundness claim.

Question 2 (LL(1) grammar)

Consider a CFG with $N = \{S, A, B, C\}$*,* $T = \{a, b, c\}$*, start symbol* S *and the following productions:* $S \to ACB, A \to aAa, A \to C, C \to cC \mid \varepsilon, B \to b.$

1. *Compute the* $First$ *sets for all non-terminals.*
2. *Compute the* $Follow$ *sets of all non-terminals.*
3. *Is this grammar LL(1)? Explain your answer.*

Solution:

1. $First(A) = \{\epsilon, a, c\}$, $First(B) = \{b\}$, $First(C) = \{c, \varepsilon\}$, $First(S) = \{a, c, b\}$.
2. $Follow(S) = \{\$\}$, $Follow(A) = \{a, b, c\}$, $Follow(B) = \{\$\}$, $Follow(C) = \{a, b, c\}$.
3. We need to check whether for all $A \in N$ with $A \to \alpha_1 | \dots | \alpha_n$ being all A-productions in G, the following holds: a) $First(\alpha_1), \dots, First(\alpha_n)$ are pairwise disjoint, and b) if $\epsilon \in First(\alpha_j)$ for some $j \in [1..n]$, then $Follow(A) \cap First(\alpha_i) = \emptyset$ for all $1 \leq i \leq n, j \neq i$ (see slide 6).

   Check of the conditions:

   - For S, the condition is trivially fulfilled since there is only one S-production.
   - For A, $First(aAa) = \{a\}$ and $First(C) = \{c, \epsilon\}$ are disjoint.
     But: $First(aAa) = \{a\}$ and $Follow(A) = \{a, b, c\}$ are not disjoint. Therefore the grammar is not LL(1).
   - For C, $First(cC) = \{c\}$ and $Follow(C) = \{a, b, c\}$ are not disjoint. This gives another reason why the grammar is not LL(1).