COMP 520: Compilers

Sample Solutions – Written Assignment 2
(Note: the rules applied in these solutions can be found in lecture 4, slides 6 – 19)

1. For the following grammars, find the Nullable nonterminals, and determine the Starters and Followers sets for all nonterminals. Next, determine whether the predict function is uniquely defined for each choice point. Finally, answer whether or not the grammar meets the LL(1) condition, and if not, explain why.

Grammar (1)  
\[ S ::= A B c \]
\[ A ::= a \mid B \mid \epsilon \]
\[ B ::= b \mid \epsilon \]

First determine the properties of the nonterminals \{A, B, S\} by finding the fixpoint of the relevant equations.

1. Nullable
\[ N_0(S) = N_0(A) = N_0(B) = \text{false} \]
\[ N_{i+1}(S) = N_i(A) \land N_i(B) \land N(c) = N_i(A) \land N_i(B) \land \text{false} = \text{false} \]
\[ N_{i+1}(A) = (N(a) \land N_i(B)) \lor N(\epsilon) = \left(\text{false} \land N_i(B)\right) \lor \text{true} = \text{true} \]
\[ N_{i+1}(B) = N(b) \lor N(\epsilon) = \text{false} \lor \text{true} = \text{true} \]

No iteration is needed since the equations are at the fixpoint solution

2. Starters
\[ ST_0(S) = \{\}, \quad ST_0(A) = ST_0(B) = \{\epsilon\} \]
\[ ST_{i+1}(S) = ST_i(A) \oplus ST_i(B) \oplus ST(c) = (ST_i(A) - \{\epsilon\}) \cup (ST_i(B) - \{\epsilon\}) \cup \{c\} \]
\[ ST_{i+1}(A) = (ST(a) \oplus ST_i(B)) \cup ST(\epsilon) = \{a\} \cup \{\epsilon\} = \{a, \epsilon\} \text{ (fixpoint)} \]
\[ ST_{i+1}(B) = ST(b) \cup ST(\epsilon) = \{b\} \cup \{\epsilon\} = \{b, \epsilon\} \text{ (fixpoint)} \]

Iterate: \[ ST_1(S) = \{c\}, \quad ST_2(S) = \{a, b, c\} \text{ (fixpoint)} \]

3. Followers
\[ FL_0(S) = \{\} \quad \text{S not on RHS of any rule} \]
\[ FL_0(A) = ST(Bc) - \{\epsilon\} = \left(\{b, \epsilon\} \oplus ST(c)\right) = \{b, c\} \quad \text{A on RHS rule 1} \]
\[ FL_0(B) = (ST(c) \cup ST(\epsilon)) - \{\epsilon\} = \{c\} \quad \text{B on RHS rule 1,2} \]

Since \(A ::= a B\) in rule 2 is of the form \(A ::= a B\gamma\), where \(\gamma\) is nullable, we must add the followers of A into the followers of B:
\[ FL_{i+1}(B) = FL_i(B) \cup FL_i(A) = \{c\} \cup \{b, c\} = \{b, c\} \]

As \(FL_1(B)\) is fully defined, no further iteration is needed and \(FL(B) = FL_1(B) = \{b, c\}\).

Finally we need to add \(\epsilon\) to \(FL(N)\) for any nonterminal \(N\) that can appear as the last symbol in a derivation starting from \(S\). As this is an augmented grammar in which (a) the start symbol \(S\) does not appear on the RHS of any rule and (b) the rule for \(S\) ends in a terminal (\(c\) in this case), we can be sure only \(FL(S)\) includes \(\{\epsilon\}\). So \(FL(S) = \{\epsilon\}\).
In summary we have

<table>
<thead>
<tr>
<th></th>
<th>Nullable</th>
<th>Starters</th>
<th>Followers</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>F</td>
<td>{a, b, c}</td>
<td>{ε}</td>
</tr>
<tr>
<td>A</td>
<td>T</td>
<td>{a, ε}</td>
<td>{b, c}</td>
</tr>
<tr>
<td>B</td>
<td>T</td>
<td>{b, ε}</td>
<td>{b, c}</td>
</tr>
</tbody>
</table>

To determine whether the LL(1) condition is met for this grammar, inspect all choice points to determine whether they define the predict function unambiguously. There are only two choice points: the alternatives in the second and third rules.

- **A ::= aB | ε**
  
  \[
  \text{Predict}(aB) = \text{Starters}(aB) \oplus \text{Followers}(A) = \{a\} \\
  \text{Predict}(ε) = \text{Starters}(ε) \oplus \text{Followers}(A) = \{b, c\}
  \]
  
  The predict sets are disjoint, so the LL(1) condition is met.

- **B ::= b | ε**
  
  \[
  \text{Predict}(b) = \text{Starters}(b) \oplus \text{Followers}(B) = \{b\} \\
  \text{Predict}(ε) = \text{Starters}(ε) \oplus \text{Followers}(B) = \{b, c\}
  \]
  
  The predict sets are **not** disjoint, so the LL(1) condition is **not** met.

Since the LL(1) condition is not met by at least one choice point in the grammar, this grammar is **not** LL(1).
Grammar (2) \[ S ::= A \$ \]
\[ A ::= \text{a} \text{A} \text{x} \mid B \]
\[ B ::= \text{b} \text{B} \text{x} \mid D \]
\[ D ::= \text{d} \text{D} \text{x} \mid \text{f} \]

First determine the grammar properties by finding the fixpoint of the relevant equations.

1. Nullable
   
   \[ N_0(S) = N_0(A) = N_0(B) = N_0(D) = \text{false} \]
   
   \[ N_{i+1}(S) = N_i(A) \land N($) = \text{false} \]
   
   \[ N_{i+1}(A) = N(\text{aA}x) \lor N_i(B) = N_i(B) \]
   
   \[ N_{i+1}(B) = N(\text{bB}x) \lor N_i(D) = N_i(D) \]
   
   \[ N_{i+1}(D) = N(\text{dD}x) \lor N(f) = \text{false} \]

   After two iterations we have a fixpoint \[ N(S) = N(A) = N(B) = N(D) = \text{false} \]

2. Starters
   
   \[ ST_0(S) = ST_0(A) = ST_0(B) = ST_0(D) = \{ \} \text{ since no nullable NTs in this grammar} \]
   
   \[ ST_{i+1}(S) = ST_i(A) \oplus ST($) = ST_i(A) \]
   
   \[ ST_{i+1}(A) = ST(\text{aA}x) \lor ST(B) = \{ \text{a} \} \lor ST_i(B) \]
   
   \[ ST_{i+1}(B) = ST(\text{bB}x) \lor ST_i(D) = \{ \text{b} \} \lor ST_i(D) \]
   
   \[ ST_{i+1}(D) = ST(\text{dD}x) \lor ST(f) = \{ \text{d, f} \} \]

   Iterate to fixpoint:
   
   \[ ST(S) = \{ \text{a, b, d, f} \}, \quad ST(A) = \{ \text{a, b, d, f} \}, \quad ST(B) = \{ \text{b, d, f} \}, \quad ST(D) = \{ \text{d, f} \} \]

3. Followers
   
   \[ FL_0(S) = \{ \} \text{ S not on RHS of any rule} \]
   
   \[ FL_0(A) = \{ \text{x, $} \} \text{ A on RHS rule 2 and rule 1} \]
   
   \[ FL_0(B) = \{ \text{x} \} \text{ B on RHS rule 3, and at end of rule 2 RHS} \]
   
   \[ FL_0(D) = \{ \text{x} \} \text{ D on RHS rule 4 and at end of rule 3 RHS} \]

   In this grammar \text{B and D occur at the RHS of a rules 2 and 3, respectively. Thus we have}

   \[ FL_{i+1}(B) = FL_i(B) \lor FL_i(A) \]
   
   \[ FL_{i+1}(D) = FL_i(D) \lor FL_i(B) \]

   and iterate to fixpoint. This yields \[ FL(B) = FL(D) = \{ \text{x, $} \}. \]

   Finally we need to add \( \epsilon \) to \( FL(N) \) for any nonterminal \( N \) that can appear as the last symbol in a derivation starting from \( S \). As this is an augmented grammar, we can be sure only \( FL(S) \) includes \{\( \epsilon \)\}. So \( FL(S) = \{\( \epsilon \)\}. \]
In summary we have

<table>
<thead>
<tr>
<th></th>
<th>Nullable</th>
<th>Starters</th>
<th>Followers</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>F</td>
<td>{a, b, d, f}</td>
<td>{ε}</td>
</tr>
<tr>
<td>A</td>
<td>F</td>
<td>{a, b, d, f}</td>
<td>{x, $}</td>
</tr>
<tr>
<td>B</td>
<td>F</td>
<td>{b, d, f}</td>
<td>{x, $}</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>{d, f}</td>
<td>{x, $}</td>
</tr>
</tbody>
</table>

To determine whether the LL(1) condition is met for this grammar, inspect all choice points to determine if they define the predict function unambiguously. There are three choice points: the

- **A ::= a A x | B**
  
  Predict(a A x) = \{a\}
  
  Predict(B) = Starters(B) = \{b, d, f\}
  
  The predict sets are disjoint, so the LL(1) condition is met.

- **B ::= b B x | D**
  
  Predict(b B x) = \{b\}
  
  Predict(D) = Starters(D) = \{d, f\}
  
  The predict sets are disjoint, so the LL(1) condition is met.

- **D ::= d D x | f**
  
  Predict(d D x) = \{d\}
  
  Predict(f) = \{f\}
  
  The predict sets are disjoint, so the LL(1) condition is met.

Since the LL(1) condition is met at all choice points in the grammar, this grammar is LL(1).
Grammar (3)  

\[
\begin{align*}
S & ::= A \, c \\
A & ::= a \mid A^* \, b \\
& \quad | \quad b
\end{align*}
\]

The Nullable and Starters functions are straightforward to compute. In this grammar \(A\) appears on the RHS in rule 1 where it is followed by \(c\), and it appears on the RHS in rule 2 where it is either followed by \(b\) or by \(ST(A)\) (since \(A^*\) can yield consecutive occurrences of \(A\)). So we have

\[
FL_0(S) = \{ \}
\]

\[
FL_0(A) = (\{c\} \cup ST(A^* b)) - \{\varepsilon\} = \{c\} \cup \{a,b\} \cup \{b\} = \{a,b,c\}
\]

Since every occurrence of \(A\) in a rule is followed by a terminal, there are no other followers to include. As this grammar is in augmented form, \(FL(S) = \{\varepsilon\}\). In summary:

<table>
<thead>
<tr>
<th></th>
<th>Nullable</th>
<th>Starters</th>
<th>Followers</th>
</tr>
</thead>
<tbody>
<tr>
<td>(S)</td>
<td>(F)</td>
<td>({a,b})</td>
<td>({\varepsilon})</td>
</tr>
<tr>
<td>(A)</td>
<td>(F)</td>
<td>({a,b})</td>
<td>({a,b,c})</td>
</tr>
</tbody>
</table>

**LL(1) Condition:** The first rule has no choice points. There are two choice points in the second rule:

- \(a \, A^* \, b \mid b\)
  
  \[
  \text{Predict}(a \, A^* \, b) = \{a\} \text{ and Predict}(b) = \{b\}
  \]
  
  The predict sets are disjoint, so the LL(1) condition is met at this choice point.

- \(A^* \, b\)
  
  First we need to check that \(A\) is not nullable - this is the case from the table above. Second we need to be able to predict repetition or termination of the Kleene star.
  
  For repetition: \(\text{Predict}(A) = \text{Starters}(A) = \{a,b\}\). To predict end of the repetition: \(\text{Predict}(b) = \{b\}\). The predict sets are **not** disjoint, so the LL(1) condition is **not** met at this choice point.

Since the LL(1) condition is not met by at least one choice point in the grammar, this grammar is **not** LL(1).