

Langages et Traducteurs

Devoir Surveillé du jeudi 5 novembre 2009

Durée : une heure – sans documents – Barème indicatif

On s'intéresse dans ce sujet à la syntaxe (simplifiée) de certaines constructions du langage C.

On note tout d'abord A le langage décrivant un accès à un élément de tableau multi-dimensionnel. Ce langage est défini sur l'ensemble de lexèmes $\{i, [,], \#\}$, où i dénote un identificateur et $\#$ est une marque de fin de séquence. Le langage A contient l'ensemble des séquences de la forme " $i[e_1][e_2] \dots [e_n]$ " où les e_i sont soit des identificateurs (i), soit eux-même des accès à des éléments de tableau (donc des séquences de A).

On donne ci-dessous quelques exemples d'éléments de A :

```
i[i]#  
i[i][i]#  
i[i[i][i]][i]#  
i[i[i[i]]][i][i]#  
...
```

Remarque : la séquence $i\#$ n'appartient pas à A .

Q1 (1 point). D'après vous le langage A est-il un *langage régulier*? Justifiez votre réponse en quelques lignes.

Q2 (4 points). Proposez une grammaire G_1 , qui est LL(1), et qui décrit le langage A . Vous justifierez le fait que G_1 est LL(1) et vous donnerez l'arbre de dérivation obtenu pour la séquence " $i[i[i]][i]\#$ ".

Q3 (4 points). Proposez une grammaire G_2 , qui est LR(0) ou SLR(1), et qui décrit le langage A . Vous justifierez le fait que G_2 est LR(0) ou SLR(1).

Rappel : Une grammaire est SLR(1) ssi elle satisfait l'une des deux conditions suivantes

1. soit elle est LR(0) ;
2. soit, dans l'analyseur LR(0) obtenu :
 - pour tout conflit lecture/réduction de la forme $([X \rightarrow \alpha.a\beta], [Y \rightarrow \gamma.])$ alors $a \notin \text{Suivants}(Y)$
 - pour tout conflit réduction/réduction de la forme $([X \rightarrow \alpha.], [Y \rightarrow \beta.])$ alors $\text{Suivants}(X) \cap \text{Suivants}(Y) = \emptyset$

On ajoute maintenant le fait que tout identificateur i peut être précédé d'une séquence (éventuellement vide) d'opérateurs $*$ (opérateur de déréférencement). On note B ce nouveau langage, qui contient ainsi les séquences suivantes :

```
*i[i]#
i[*i]#
**i[*i[***i]][*i]
...
```

Remarque : le langage A est strictement inclus dans le langage B .

Q4 (3 points).

On suppose que l'opérateur $*$ est *plus prioritaire* que l'opérateur $[]$. Ainsi, la séquence $*i[i]$ doit être interprétée comme $(*i)[i]$ et non comme $*(i[i])$.

Sous cette hypothèse donnez une grammaire G_3 qui décrit le langage B . Donnez l'arbre de dérivation obtenu pour " $**i[*i[i]]\#$ ". Cet arbre doit justifier le fait que les priorités sont correctement prises en compte par G_3 .

Q5 (4 points). La grammaire G_3 est-elle LL(1)? Si non, transformez-la en une grammaire G_4 , LL(1), décrivant le même langage.

On s'intéresse maintenant pour finir à un langage D constitué d'expressions arithmétiques construites à l'aide :

- de l'opérateur de multiplication noté $*$
- d'identificateurs i précédés d'une séquence (éventuellement vide) d'opérateurs de déréférencement, également noté $*$

Ainsi, ce langage D contient par exemple les séquences suivantes :

```
i#
*i#
i*i#
*i*i#
*i**i#
*i**i***i#
...
```

Q6 (4 points). On suppose que l'opérateur de multiplication est associatif à gauche et que l'opérateur de déréférencement est plus prioritaire que l'opérateur de multiplication.

Proposez une grammaire G_5 , LL(1), qui décrive le langage D sous ces hypothèses. Justifiez que cette grammaire est LL(1).