Caractérisation de premier en analyse LL(1)

Référence : 131 Développements pour l'oral.

Définition de premier.

Définition 1. *Soit G une grammaire, et $w \in (V \cup \Sigma)^*$, on définit :*

$$\mathbf{premier}(w) = \begin{cases} \mathbf{premier}'(w) & si\ w \not\Rightarrow^* \epsilon \\ \mathbf{premier}'(w) \cup \{\epsilon_0\} & si\ w \Rightarrow^* \epsilon \end{cases}$$

où :

$$\mathbf{premier}'(w) = \{a \in \Sigma | w \Rightarrow^* aw', w' \in (V \cup \Sigma)^*\}$$

Remarque 1. *ϵ_0 est un symbole frais qui peut être interprété comme « première lettre de ϵ ».*

Définition axiomatique. On considère une vision ensembliste des fonctions. On veut montrer que la fonction **premier** est la plus petite partie P de $(V \cup \Sigma)^* \times (\Sigma \cup \{\epsilon_0\})$ telle que

1. $a \in P(a)$;
2. $\epsilon_0 \in P(\epsilon)$;
3. si $N \to w_1...w_n$, alors $P(w_1....w_n) \subset P(N)$;
4. si $N \to \epsilon$, alors $\epsilon_0 \in P(N)$;
5. $(P(w_1)\backslash\{\epsilon_0\}) \subset P(w_1....w_n)$;
6. si $\epsilon_0 \in P(w_1)$, alors $P(w_2...w_n) \subset P(w_1....w_n)$

Lemme 1. *La fonction* **premier** *vérifie les axiomes 1 à 6.*

Démonstration.

1. Pour $a \in \Sigma$, on a $a \Rightarrow^* a$ et donc $a \in \mathbf{premier}(a)$.
2. De même, $\epsilon \Rightarrow^* \epsilon$, et donc $\epsilon_0 \in \mathbf{premier}(\epsilon)$;
3. Si $N \to w_1...w_n$. Soit $a \in \mathbf{premier}(w_1...w_n)$.
   Si $a = \epsilon_0$, alors $w_1...w_n \Rightarrow^* \epsilon$, et donc par transitivité $N \Rightarrow^* \epsilon$. Ainsi, $\epsilon_0 \in \mathbf{premier}(N)$.
   Si $a \neq \epsilon_0$, on a $w_1...w_n \Rightarrow aw'$, et donc encore par transitivité, $N \Rightarrow^* aw'$, et finalement $a \in \mathbf{premier}(N)$.
4. Si $N \Rightarrow \epsilon$, alors $N \Rightarrow^* \epsilon$ et donc $\epsilon_0 \in \mathbf{premier}(N)$.
5. Soit $w_1...w_n \in (\Sigma \cup V)^*$. Soit $a \in \mathbf{premier}(w_1)\backslash\{\epsilon_0\}$. Ainsi, $a \in \mathbf{premier}'(w_1)$, et donc $w_1 \Rightarrow^* aw_1'$. En appliquant les mêmes règles au mot $w_1....w_n$, on a $w_1....w_n \Rightarrow^* aw_1'w_2...w_n$, et donc $a \in \mathbf{premier}(w_1...w_n)$.
6. Soit $w_1...w_n \in (\Sigma \cup V)^*$, avec $\epsilon_0 \in \mathbf{premier}(w_1)$. Alors, on a $w_1 \Rightarrow^* \epsilon$. En appliquant les même règle, on a $w_1...w_n \Rightarrow^* \epsilon w_2...w_n = w_2...w_n$. On en déduit directement $\mathbf{premier}(w_2...w_n) \subset \mathbf{premier}(w_1...w_n)$.

□

Lemme 2. *Soit P la plus petite partie vérifiant les axiomes 1 à 6. Montrer que pour tout $n \in \mathbb{N}$, pour tout $a \in \Sigma$, pour tout $w, w' \in (\Sigma \cup V)^*$, on a*

— *si* $w \Rightarrow^n aw'$, *alors* $a \in P(w)$ *;*
— *si* $w \Rightarrow^n \epsilon$, *alors* $\epsilon_0 \in P(w)$.

En admettant ce lemme pour l'instant, on peut montrer le résultat voulu.

Théorème 1. premier *est la plus petite partie* P *vérifiant les propriétés 1 à 6.*

Démonstration. Soit P la plus petite partie vérifiant les axiomes 1 à 6.
Puisque **premier** vérifie ces axiomes d'après le lemme 1, on a $P \subset$ **premier**.
Le lemme 2 nous permet de montrer l'autre inclusion. En effet, si $a \in$ **premier**(w). Ainsi, il existe $w' \in (\Sigma \cup V)^*$ tel que $w \Rightarrow^* aw'$ (avec possiblement $a = \epsilon_0$ et donc $w' = \epsilon$). Par le lemme 2, on a directement $a \in P(w)$, et donc pour tout w, on a **premier**$(w) \subset P(w)$.
Finalement, **premier** $= P$. □

Algorithme. Cette définition axiomatique nous donne un moyen de calculer **premier** par saturation.

Algorithme 1 : Calcul_Premier(w)

```
pour a ∈ Sigma faire
    premier(a) ← {a} ;
premier(ϵ) ← {ϵ0} ;
pour N → ϵ ∈ R faire
    premier(N) ← {ϵ0} ;
tant que premier est modifié faire
    pour N → w1...wn ∈ R faire
        premier(N) ← premier(N) ∪ premier(w1...wn) ;
        si ϵ0 ∉ premier(w1) alors
            premier(w1...wn) ← premier(w1...wn) ∪ (premier(w1)\{ϵ0}) ;
        sinon
            premier(w1...wn) ← premier(w1...wn) ∪ premier(w2...wn) ;
```

Démonstration du lemme 2. On montre le résultat par récurrence forte sur n.
Pour $n = 0$, soit $a \in \Sigma$, $w, w' \in (\Sigma \cup V)^*$.
— Si $w = aw'$, alors on a $a \in P(a)$ d'après 1, et en particulier $a \in P(a)\backslash\{\epsilon_0\}$. D'après 5, on a $a \in P(aw')$, et donc $a \in P(w)$.
— Si $w = \epsilon$, alors par 2, on a directement $\epsilon_0 \in P(w)$.
On suppose la propriété vraie pour tout $k < n$ pour un certain $n > 0$. soit $a \in \Sigma$, $w, w' \in (\Sigma \cup V)^*$.
On note $w = w_1...w_m$.
— Si $w \Rightarrow^n aw'$. On veut montrer que $a \in P(w)$. On distingue deux cas
  — Si $w_1 \in \Sigma$, alors $w_1 = a$. Ainsi par 1 et 5, on a directement $a \in P(w)$.
  — Si $w_1 \notin \Sigma$, ie $w_1 = N \in V$. On distingue deux cas :
    — S'il existe $j < n$ tel que

$$Nw_2...w_m \Rightarrow^j w_2...w_m \Rightarrow^{n-j} aw'$$

On a alors $\epsilon_0 \in P(N)$ par HR au rang j, et $a \in P(w_2...w_m)$ par HR au rang $n-j$. De plus, par l'axiome 6, on a $P(w_2...w_m) \subset P(w_1...w_m)$ et donc $a \in P(w)$.

  — Sinon, la première règle appliquée à N est de la forme $N \Rightarrow x_1...x_k$. De plus, puisque $x_1...x_k \not\Rightarrow^* \epsilon$, on prend i l'indice minimal tel que $x_i \Rightarrow^* a$. Il existe alors $j < n$ tel que
  $$Nw_2...w_n \Rightarrow^j x_i...x_k w_2...w_m \Rightarrow aw'$$
  Par application successive de 6, on a
  $$P(x_i...x_k) \subset P(x_1...x_k)$$
  Or par HR, on a $a \in P(x_i) \subset P(x_i...x_k)$ par 5. En combinant, on a $a \in P(x_1...x_k)$, puis $a \in P(N)$ par 3, et finalement par 5, on a $a \in P(Nw_2...w_m)$, c'est-à-dire $a \in P(w)$.

— Si $w \Rightarrow^n \epsilon$, on recommence avec w qui commence par un non terminal.

$\square$