4
Les automates à pile.

1 – Généralités.

Les automates finis reconnaissent exactement les langages réguliers : si l'on veut construire des automates adaptés aux langages algébriques (qui pour la plupart ne sont pas réguliers) il faut introduire, sinon un ensemble infini d'états, du moins un ensemble non borné. Ceci se fait au moyen d'une pile qui, associée à un ensemble d'états (au sens précédent du terme) simule bien un ensemble non borné.

Configurations.

Une configuration d'un automate à pile, est un triplet constitué :

- d'une pile ;
- d'un état ;
- de la partie du mot restant à reconnaître.

Il y a donc $3! = 6$ façons de présenter une configuration.

De plus, on peut convenir de mettre le sommet de la pile à gauche ou à droite, et de lire le mot à reconnaître de gauche à droite ou de droite à gauche !

Dans ce qui suit, nous adopterons généralement les conventions suivantes :

- nous écrirons les configurations sous la forme (π, q, u), c'est–à–dire (pile, état, mot) ;
- lorsqu'il n'y a qu'un seul état, celui–ci ne figurera pas : les configurations s'écriront donc simplement (π, u) ;
- nous placerons le sommet de la pile π à droite (sauf dans la démonstration du théorème) ;
- nous lirons le mot u de gauche à droite.

1.1 – Les automates à pile.

Un automate à pile (AP en abrégé) sur l'alphabet $\mathcal{A}$ est la donnée d'un 7–uplet $\mathbf{A} = (\Pi, Q, \mathcal{A}, \delta, Z_0, q_0, F)$ où

- Π est un alphabet fini, dit *alphabet de pile* ;
- Q est un ensemble fini d'états ;
- $\delta : \Pi^* \times Q \times (\varepsilon + \mathcal{A}) \to \mathcal{P}(\Pi^* \times Q)$ est une fonction finie ;
- $Z_0 \in \Pi$ est le symbole initial de la pile ;
- $q_0 \in Q$ est l'état initial ;
- $F \subseteq Q$ est l'ensemble des états finaux.

La propriété de finitude de δ signifie que, pour tout élément (α, q, x) de son domaine :

- $\delta(\alpha, q, x)$ est fini ;
- $\delta(\alpha, q, x) = \emptyset$ sauf pour un ensemble fini de triplets (α, q, x).

δ peut se représenter de deux façons équivalentes :

- par une table exprimant $\delta(\alpha, q, x)$ en fonction de (α, q) et de x (une table de dimension 3 serait préférable !) ;
- par un graphe de transition, dont les arêtes sont ici

$$\boxed{q} \xrightarrow{(\alpha \multimap \beta; x)} \boxed{r}$$

pour chaque $(\beta, r) \in \delta(\alpha, q, x)$.

Le sens de $\alpha \multimap \beta$ est

"dépiler α, puis empiler β"

c'est–à–dire réécrire la pile de la façon suivante

$$\pi\alpha \Longrightarrow \pi\beta$$

Un AP est dit *restreint* (APR en abrégé) lorsque

$$\delta(\alpha, q, x) \neq \emptyset \text{ implique } \alpha \in \Pi$$

c'est–à–dire que α est réduit à un seul caractère.

Dérivation dans un AP.

$\alpha, \beta \in \Pi^*$, $q, r \in Q$ et $x \in \varepsilon + \mathcal{A}$ tels que $(\beta, r) \in \delta(\alpha, q, x)$ définissent un changement de configuration :

$$(\alpha, q, x) \underset{\mathbf{A}}{\overset{1}{\vdash\!\!-}} (\beta, r, \varepsilon)$$

qui, pour tout $\pi \in \Pi^*$ et tout $u \in \mathcal{A}^*$ induit *une transition* :

$$(\pi\alpha, q, xu) \underset{\mathbf{A}}{\overset{1}{\vdash\!\!-}} (\pi\beta, r, u)$$

La définition d'*une dérivation de longueur* i

$$(\pi, q, u) \underset{\mathbf{A}}{\overset{i}{\vdash\!\!-}} (\rho, r, v)$$

et celle de la relation $\underset{\mathbf{A}}{\overset{*}{\vdash\!\!-}}$ entre configurations, sont analogues à celles qui ont été données pour les AF.

Langage reconnu par un AP.

Les configurations initiales et des configurations finales sont généralement choisies de la façon suivante.

Configurations initiales.

Voici deux types classiques de configurations initiales :

- (Z_0, q_0, u) ;
- (ε, q_0, u), (la donnée de Z_0 est alors inutile !) ;

pour tout $u \in \mathcal{A}^*$.

Configurations finales.

Voici trois types classiques de configurations finales :

- **Etat final** : (π, s, ε) pour toute pile π et tout état final $s \in F$;
- **Pile vide** : $(\varepsilon, q, \varepsilon)$ pour tout état q, (la donnée de F est alors inutile !) ;
- **Etat final et pile vide** : $(\varepsilon, s, \varepsilon)$ pour tout état final $s \in F$.

Dans tous les cas, **le mot doit être complètement consommé**, c'est–à–dire réduit à ε dans une configuration finale.

Remarque.

D'autres choix sont possibles, car des artifices simples permettent de se ramener aux cas précédents : nous utiliserons souvent cette facilité.

Ceci conduit aux deux définitions suivantes d'un langage reconnu par un AP $\mathbf{A}$:

$\mathcal{L}(\mathbf{A})$

Le langage $\mathcal{L}(\mathbf{A}) \subseteq \mathcal{A}^*$ reconnu par $\mathbf{A}$, par **état final** est défini par : $u \in \mathcal{L}(\mathbf{A})$ ssi il existe $\pi \in \Pi^*$ et $s \in F$ tels que

$$(Z_0, q_0, u) \underset{\mathbf{A}}{\overset{*}{\vdash}} (\pi, s, \varepsilon)$$

---

et

$\mathcal{L}_\varepsilon(\mathbf{A})$

Le langage $\mathcal{L}_\varepsilon(\mathbf{A}) \subseteq \mathcal{A}^*$ reconnu par $\mathbf{A}$, par **pile vide** est défini par : $u \in \mathcal{L}_\varepsilon(\mathbf{A})$ ssi il existe $q \in Q$ tel que

$$(Z_0, q_0, u) \underset{\mathbf{A}}{\overset{*}{\vdash}} (\varepsilon, q, \varepsilon)$$

Notons un résultat, important et facile à vérifier, qui confirme la remarque de la page précédente :

$\mathcal{L}(\mathbf{A}) = \mathcal{L}_\varepsilon(\mathbf{A}')$

Pour tout AP $\mathbf{A}$ il existe un AP $\mathbf{A}'$ tel que $\mathcal{L}(\mathbf{A}) = \mathcal{L}_\varepsilon(\mathbf{A}')$, et réciproquement.

Exemple 1.

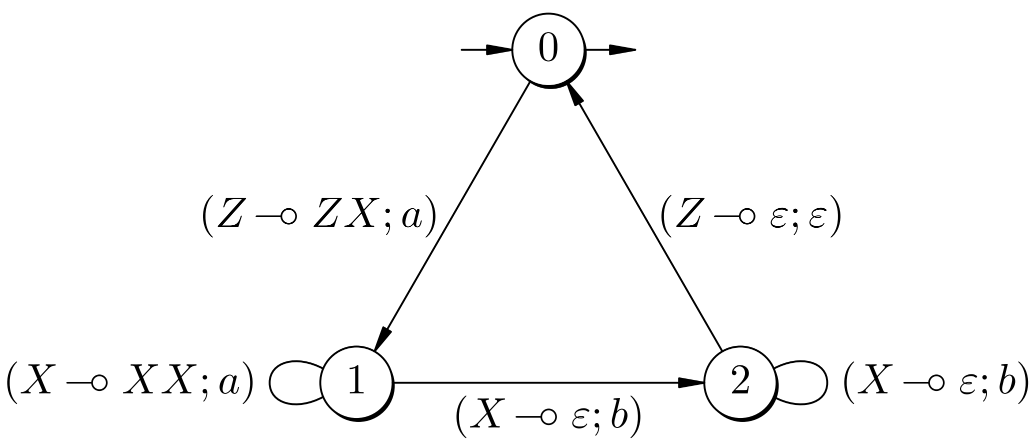
Soit $\mathbf{A}_1$ l'APR défini par les données :

$\Pi = Z + X$, $Q = 0 + 1 + 2$, $\mathcal{A} = a + b$, δ est définie par

(P, q)	$\delta(P, q, \varepsilon)$	$\delta(P, q, a)$	$\delta(P, q, b)$
$(Z, 0)$	$\emptyset$	$(ZX, 1)$	$\emptyset$
$(X, 1)$	$\emptyset$	$(XX, 1)$	$(\varepsilon, 2)$
$(X, 2)$	$\emptyset$	$\emptyset$	$(\varepsilon, 2)$
$(Z, 2)$	$(\varepsilon, 0)$	$\emptyset$	$\emptyset$

$Z_0 = Z$, $q_0 = 0$ et $F = 0$.

Le graphe de transition de $\mathbf{A}_1$ est le suivant :



Il est facile de construire les dérivations suivantes dans $\mathbf{A}_1$, pour tout entier naturel n :

$$
\begin{aligned}
(X, 1, a^n) &\vdash\!\!\frac{n}{\quad} (X^{n+1}, 1, \varepsilon) \\
(X^n, 2, b^n) &\vdash\!\!\frac{n}{\quad} (\varepsilon, 2, \varepsilon)
\end{aligned}
$$

Nous allons vérifier que $\mathcal{L}(\mathbf{A}_1) = \{a^n b^n \mid n \geq 0\}$.

☞ $\{a^n b^n \mid n \geq 0\} \subseteq \mathcal{L}(\mathbf{A}_1)$:

- $\varepsilon \in \mathcal{L}(\mathbf{A}_1)$:

$$(Z, 0, \varepsilon) \vdash\!\!\frac{0}{\quad} (Z, 0, \varepsilon)$$

- $a^{n+1} b^{n+1} \in \mathcal{L}(\mathbf{A}_1)$ pour tout $n > 0$:

$$
\begin{aligned}
(Z, 0, a^{n+1} b^{n+1}) &\vdash\!\!\frac{1}{\quad} (ZX, 1, a^n b^{n+1}) \\
&\vdash\!\!\frac{n}{\quad} (ZX^{n+1}, 1, b^{n+1}) \\
&\vdash\!\!\frac{1}{\quad} (ZX^n, 2, b^n) \\
&\vdash\!\!\frac{n}{\quad} (Z, 2, \varepsilon) \\
&\vdash\!\!\frac{1}{\quad} (\varepsilon, 0, \varepsilon)
\end{aligned}
$$

☞ $\mathcal{L}(\mathbf{A}_1) \subseteq \{a^n b^n \mid n \geq 0\}$: la suite des états d'une dérivation complète non triviale forme un mot $01^m 2^n 0$.

Or il faut dépiler autant d'occurrences de X que l'on en a empilé : on a donc $n = m$, d'où on peut conclure.

Exemple 2.

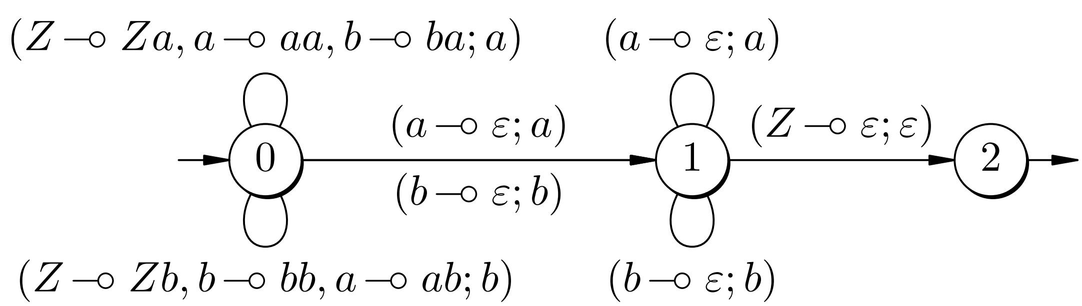
Soit $\mathbf{A}_2$ l'APR défini par les données :

$\Pi = Z + a + b$, $Q = 0 + 1 + 2$, $\mathcal{A} = a + b$, δ est définie par

(P,q)	$\delta(P,q,\varepsilon)$	$\delta(P,q,a)$	$\delta(P,q,b)$
$(Z,0)$	$\emptyset$	$(Za,0)$	$(Zb,0)$
$(a,0)$	$\emptyset$	$(aa,0)+(\varepsilon,1)$	$(ab,0)$
$(b,0)$	$\emptyset$	$(ba,0)$	$(bb,0)+(\varepsilon,1)$
$(a,1)$	$\emptyset$	$(\varepsilon,1)$	$\emptyset$
$(b,1)$	$\emptyset$	$\emptyset$	$(\varepsilon,1)$
$(Z,1)$	$(\varepsilon,2)$	$\emptyset$	$\emptyset$

$Z_0 = Z$, $q_0 = 0$ et $F = 2$.

Le graphe de transition de $\mathbf{A}_2$ est le suivant :



On peut vérifier que $\mathcal{L}(\mathbf{A}_2) = \{u\tilde{u} \mid u \in \mathcal{A}^* \textit{ et } |u| \neq 0\}$.

Exemple 3

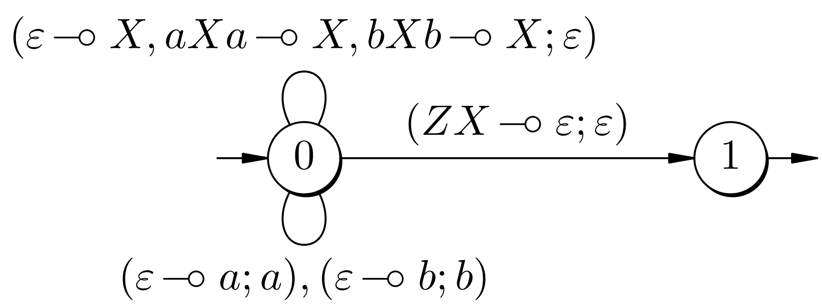
Soit $\mathbf{A}_3$ l'AP défini par les données :

$\Pi = Z + X + a + b$, $Q = 0 + 1$, $\mathcal{A} = a + b$, δ est définie par

(α, q)	$\delta(\alpha, q, \varepsilon)$	$\delta(\alpha, q, a)$	$\delta(\alpha, q, b)$
$(\varepsilon, 0)$	$(X, 0)$	$(a, 0)$	$(b, 0)$
$(aXa, 0)$	$(X, 0)$	$\emptyset$	$\emptyset$
$(bXb, 0)$	$(X, 0)$	$\emptyset$	$\emptyset$
$(ZX, 0)$	$(\varepsilon, 1)$	$\emptyset$	$\emptyset$

$Z_0 = Z$, $q_0 = 0$ et $F = 1$.

Le graphe de transition de $\mathbf{A}_3$ est le suivant :



On peut vérifier que $\mathcal{L}(\mathbf{A}_3) = \{u\tilde{u} \mid u \in \mathcal{A}^*\}$.

Tout AP est équivalent à un APR.

Plus précisément, les deux propriétés suivantes sont équivalentes pour tout langage L :

- il existe un AP $\mathbf{A}$ tel que $L = \mathcal{L}(\mathbf{A})$;
- il existe un APR $\mathbf{A}'$ tel que $L = \mathcal{L}(\mathbf{A}')$.

Un APR étant un cas particulier d'AP, l'implication réciproque de l'énoncé ci–dessus est parfaitement triviale.

Soit $\mathbf{A} = (\Pi, Q, \mathcal{A}, \delta, Z_0, q_0, F)$ un AP, il faut construire un APR $\mathbf{A}' = (\Pi', Q', \mathcal{A}, \delta', Z_0', q_0', F')$.

Les états de $\mathbf{A}'$ comportent un "tampon" : la configuration $(\sigma, \tau q, u)$ de $\mathbf{A}'$ simule la configuration $(\sigma\tau, q, u)$ de $\mathbf{A}$.
Soit m la plus grande des longueurs de ϱ et π tels que $(\varrho, r) \in \delta(\pi, q, \xi)$ (on suppose que $m > 0$). On définit :

- $Q' = (\varepsilon + \Pi)^m Q$;
- $\Pi' = \Pi + Z_0'$ où $Z_0' \notin \Pi$ est un nouveau symbole ;

- $(P\alpha, \varrho r) \in \delta'(P, \alpha\pi q, \xi)$ ssi $(\varrho, r) \in \delta(\pi, q, \xi)$,
- $(\varepsilon, P\alpha q) \in \delta'(P, \alpha q, \varepsilon)$ pour tous α, π et $\varrho \in \Pi^*$, $P \in \Pi$, q et $r \in Q$, et $\xi \in \varepsilon + \mathcal{A}$;

- $q_0' = Z_0 q_0$;
- pour tout $\alpha q \in Q'$: $\alpha q \in F'$ ssi $q \in F$.

Il est clair que $\mathcal{L}(\mathbf{A}') = \mathcal{L}(\mathbf{A})$.

THÉORÈME.

Les langages reconnus par les AP sont exactement les langages algébriques.

ATTENTION : **ici, le sommet de la pile est placé à gauche** : tout $(\beta, r) \in \delta(\alpha, q, x)$ définit des transitions élémentaires qui s'écrivent donc

$$(\alpha\pi, q, xu) \underset{\mathbf{A}}{\overset{1}{\vdash}} (\beta\pi, r, u).$$

Les AP considérés simulent des **dérivations à gauche** dans une grammaire, (enchaînements de dérivations élémentaires de la forme $vX\pi \underset{G}{\overset{1}{\Longrightarrow}} v\alpha\pi$ où $v \in \mathcal{A}^*$).

☞ Soit $G = (\mathcal{V}, \mathcal{A}, R)$ une grammaire et soit $S \in \mathcal{V}$: nous allons construire un APR $\mathbf{A}$ tel que $\mathcal{L}_\varepsilon(\mathbf{A}) = \mathcal{L}(G, S)$.
$\mathbf{A} = (\mathcal{V} + \mathcal{A}, 0, \mathcal{A}, \delta, S, 0, \emptyset)$ a un seul état 0 (qui n'apparaît donc pas dans les configurations). δ est définie par les transitions suivantes :

- pour toute $(X \longrightarrow \alpha) \in R$: $0 \xrightarrow{(X \multimap \alpha; \varepsilon)} 0$
- pour tout $x \in \mathcal{A}$: $0 \xrightarrow{(x \multimap \varepsilon; x)} 0$

On vérifie que, pour toute $X \in \mathcal{V}$ et tout $u \in \mathcal{A}^*$:

$$X \underset{G}{\overset{*}{\Longrightarrow}} u \text{ ssi } (X, u) \underset{\mathbf{A}}{\overset{*}{\vdash}} (\varepsilon, \varepsilon)$$

ce qui permet de conclure.

Exemple 3.1 (suite).

Avec la grammaire $S \longrightarrow SbS + \varepsilon + a + aa$, on obtient :

P	$\delta(P,\varepsilon)$	$\delta(P,a)$	$\delta(P,b)$
S	$SbS + \varepsilon + a + aa$	$\emptyset$	$\emptyset$
a	$\emptyset$	ε	$\emptyset$
b	$\emptyset$	$\emptyset$	ε

Dans l'illustration qui suit, $\cdot$ est un symbole étranger qui n'est utilisé que pour "suivre la lecture du doigt" :

$$
\begin{array}{llll}
\cdot S & \overset{1}{\Longrightarrow} \cdot SbS & (S, abbaa) & \overset{1}{\longmapsto} (SbS, abbaa) \\
 & \overset{1}{\Longrightarrow} \cdot abS & & \overset{1}{\longmapsto} (abS, abbaa) \\
 & \quad a \cdot bS & & \overset{1}{\longmapsto} (bS, bbaa) \\
 & \quad ab \cdot S & & \overset{1}{\longmapsto} (S, baa) \\
 & \overset{1}{\Longrightarrow} ab \cdot SbS & & \overset{1}{\longmapsto} (SbS, baa) \\
 & \overset{1}{\Longrightarrow} ab \cdot bS & & \overset{1}{\longmapsto} (bS, baa) \\
 & \quad abb \cdot S & & \overset{1}{\longmapsto} (S, aa) \\
 & \overset{1}{\Longrightarrow} abb \cdot aa & & \overset{1}{\longmapsto} (aa, aa) \\
 & \quad abba \cdot a & & \overset{1}{\longmapsto} (a, a) \\
 & \quad abbaa \cdot & & \overset{1}{\longmapsto} (\varepsilon, \varepsilon)
\end{array}
$$

☞ Réciproquement, si $\mathbf{A} = (\Pi, Q, \mathcal{A}, \delta, Z_0, q_0, \emptyset)$ est un APR, on peut construire une grammaire $G = (\mathcal{V}, \mathcal{A}, R)$ et trouver $S \in \mathcal{V}$ telles que $\mathcal{L}_\varepsilon(\mathbf{A}) = \mathcal{L}(G, S)$.

On considère l'alphabet des variables $\mathcal{V} = S + Q \times \Pi \times Q$ où S est un nouveau symbole.

On définit alors les ensembles de mots sur $\mathcal{V}$ suivants :

- $[qPr] = (q, P, r)$ pour tout $(q, P, r) \in Q \times \Pi \times Q$;
- $[q\alpha Pr] = \sum_{s \in Q} [q\alpha s][sPr]$ pour tout $\alpha \in \Pi^* - \varepsilon$.

Les règles de la grammaire G sont construites comme suit :

- $S \longrightarrow [q_0 Z_0 q]$ pour tout $q \in Q$.
- pour tout élément $(\alpha, r) \in \delta(P, q, x)$ et tout $s \in Q$, on ajoute l'ensemble des règles suivantes

$$[qPs] \longrightarrow \begin{cases} x & \text{si } \alpha = \varepsilon, \\ x[r\alpha s] & \text{sinon.} \end{cases}$$

On vérifie que, pour tout $P \in \Pi$, tout q, tout $s \in Q$ et tout $u \in \mathcal{A}^*$:

$$[qPs] \xRightarrow[G]{*} u \text{ ssi } (P, q, u) \vdash_{G}^{*} (\varepsilon, s, \varepsilon)$$

ce qui permet de conclure :

$$S \xRightarrow[G]{1} [q_0 Z_0 q] \xRightarrow[G]{*} u \text{ ssi } (Z_0, q_0, u) \vdash_{G}^{*} (\varepsilon, q, \varepsilon)$$

Exemple.

Soit $\mathbf{A}_4$ l'APR, avec une pile dont le sommet est placé à gauche, défini par les données :

$\Pi = Z + X$, $Q = 0 + 1$, $\mathcal{A} = a + b$, δ définie par

(P, q)	$\delta(P, q, \varepsilon)$	$\delta(P, q, a)$	$\delta(P, q, b)$
$(Z, 0)$	$\emptyset$	$(XZ, 0)$	$\emptyset$
$(X, 0)$	$\emptyset$	$(XX, 0)$	$(\varepsilon, 1)$
$(X, 1)$	$(\varepsilon, 1)$	$\emptyset$	$(\varepsilon, 1)$
$(Z, 1)$	$(\varepsilon, 1)$	$\emptyset$	$\emptyset$

$Z_0 = Z$, $q_0 = 0$ et $F = \emptyset$.

En ne considérant que les variables accessibles à partir de S, on a successivement :

Productions initiales :

$$S \longrightarrow [0Z0] + [0Z1]$$

$\delta(Z, 0, a) = (XZ, 0)$ ajoute les règles :

$$\begin{aligned} [0Z0] &\longrightarrow a[0X0][0Z0] + a[0X1][1Z0] \\ [0Z1] &\longrightarrow a[0X0][0Z1] + a[0X1][1Z1] \end{aligned}$$

$\delta(X, 0, a) = (XX, 0)$ ajoute les règles :

$$\begin{aligned} [0X0] &\longrightarrow a[0X0][0X0] + a[0X1][1X0] \\ [0X1] &\longrightarrow a[0X0][0X1] + a[0X1][1X1] \end{aligned}$$

$\delta(X, 0, b) = (\varepsilon, 1)$ ajoute la règle :

$$[0X1] \longrightarrow b$$

$\delta(X, 1, \varepsilon) = (\varepsilon, 1)$ ajoute la règle :

$$[1X1] \longrightarrow \varepsilon$$

$\delta(X, 1, b) = (\varepsilon, 1)$ ajoute la règle :

$$[1X1] \longrightarrow b$$

$\delta(Z, 1, \varepsilon) = (\varepsilon, 1)$ ajoute la règle :

$$[1Z1] \longrightarrow \varepsilon$$

En ne conservant que les variables productives, on obtient :

$$\begin{aligned} S &\longrightarrow [0Z1] \\ [0Z1] &\longrightarrow a[0X1][1Z1] \\ [0X1] &\longrightarrow a[0X1][1X1] + b \\ [1X1] &\longrightarrow \varepsilon + b \\ [1Z1] &\longrightarrow \varepsilon \end{aligned}$$

ce qui, par élimination de l'ε–production donne la grammaire G :

$$\begin{aligned} S &\longrightarrow a[0X1][1Z1] \\ [0X1] &\longrightarrow a[0X1][1X1] + b \\ [1X1] &\longrightarrow \varepsilon + b \\ [1Z1] &\longrightarrow \varepsilon \end{aligned}$$

La comparaison de ces règles avec la table de transition de $\mathbf{A}_4$ ne laisse que peu de doute sur le fait que l'on a effectivement $\mathcal{L}(G, S) = \mathcal{L}_\varepsilon(\mathbf{A}_4)$.

APR déterministe

Un APR **A** est déterministe ssi pour tout $P \in \Pi$, tout $q \in Q$ et tout $x \in \varepsilon + \mathcal{A}$:

1) $\delta(P, q, x)$ comporte **au plus un élément** ;
2) si $\delta(P, q, \varepsilon) \neq \emptyset$ alors $\delta(P, q, x) = \emptyset$ pour $x \neq \varepsilon$.

Plus généralement :

AP déterministe

Un AP **A** est dit *déterministe* ssi pour tous $\alpha, \beta \in \Pi^*$, tout $q \in Q$ et tout $x \in \varepsilon + \mathcal{A}$:

1) $\delta(\alpha, q, x)$ comporte **au plus un élément** ;
2) lorsque $\alpha \neq \beta$ et
   - ou bien $\delta(\alpha, q, x) \neq \emptyset$ et $\delta(\beta, q, x) \neq \emptyset$,
   - ou bien $\delta(\alpha, q, x) \neq \emptyset$ et $\delta(\beta, q, \varepsilon) \neq \emptyset$,

   alors aucun des mots α et β n'est un facteur droit de l'autre.

Contrairement aux AF, un AP n'est pas nécessairement équivalent à un AP déterministe. En fait, *les langages déterministes*, c'est–à–dire, les langages reconnaissables par des AP déterministes forment une classe **strictement** comprise entre les langages réguliers et les langages algébriques.

Il est facile de vérifier, par exemple, que le complémentaire d'un langage déterministe est déterministe.

2 – Dérivations à droite et analyse ascendante.

Nous allons construire des AP qui permettent l'analyse (*parsing*) d'un mot par une grammaire $G = (\mathcal{V}, \mathcal{A}, R)$.

Dérivations à droite.

• Une dérivation élémentaire dans G est dite *à droite* ssi elle se présente sous la forme $\pi X v \underset{G}{\overset{1}{\Longrightarrow}} \pi\alpha v$ pour $v \in \mathcal{A}^*$.

• Une dérivation *à droite* est un enchaînement de dérivations élémentaires à droite : il est clair que **toute dérivation est équivalente à une dérivation à droite.** On écrira $\beta \underset{d,G}{\overset{k}{\Longrightarrow}} \gamma$ ou plus simplement $\beta \underset{d}{\overset{k}{\Longrightarrow}} \gamma$.

Exemple 3.1 (suite).

Numérotons les règles de notre grammaire :

$1 : S \longrightarrow SbS \qquad 2 : S \longrightarrow \varepsilon \qquad 3 : S \longrightarrow a \qquad 4 : S \longrightarrow aa$

La dérivation à droite que voici :

$$S \underset{d}{\overset{1}{\Longrightarrow}} SbS \underset{d}{\overset{1}{\Longrightarrow}} Sbaa \underset{d}{\overset{1}{\Longrightarrow}} SbSbaa \underset{d}{\overset{1}{\Longrightarrow}} Sbbaa \underset{d}{\overset{1}{\Longrightarrow}} abbaa$$

est entièrement décrite par la suite $1, 4, 1, 2, 3$ des règles qui ont été utilisées pour la construire.

Les AP que nous considérons, produisent des dérivations qui vont du mot à analyser vers l'axiome de la grammaire (aussi parle–t–on **d'analyse ascendante**) : la liste précédente sera donc obtenue *à l'envers* !

Soit $G = (\mathcal{V}, \mathcal{A}, R)$ une grammaire et soit $S \in \mathcal{V}$: nous allons construire un APR $\mathbf{A}$ reconnaissant le langage $\mathcal{L}(G, S)$. Cette construction sera améliorée par la suite.

$\mathbf{A} = (\mathcal{V} + \mathcal{A}, 0, \mathcal{A}, \delta, S, 0, \emptyset)$ a un seul état 0 (qui n'apparaît donc pas dans les configurations).

δ est définie par les transitions suivantes :

– Pour toute $(X \longrightarrow \alpha) \in R$: $\circled{0} \xrightarrow{(\alpha \multimap X; \varepsilon)} \circled{0}$

– Pour tout $x \in \mathcal{A}$: $\circled{0} \xrightarrow{(\varepsilon \multimap x; x)} \circled{0}$

On peut vérifier que : pour tout $v \in \mathcal{A}^*$

$$\text{il existe } w \in \mathcal{A}^* \text{ tel que } \pi \overset{*}{\underset{G}{\Longrightarrow}} w \text{ et } u = wv$$

$$\text{ssi}$$

$$(\varepsilon, u) \overset{*}{\underset{\mathbf{A}}{\vdash}} (\pi, v).$$

En particulier : $S \overset{*}{\underset{G}{\Longrightarrow}} u$ ssi $(\varepsilon, u) \overset{*}{\underset{\mathbf{A}}{\vdash}} (S, \varepsilon)$.

Cet AP produit ce que l'on peut appeler une *analyse partielle d'un mot* u, c'est–à–dire, une dérivation $\pi \underset{d}{\Longrightarrow} w$ et un mot $v \in \mathcal{A}^*$ tels que $u = wv$: en ceci, on ne s'est intéressé qu'à la fin d'une dérivation à droite !

Exemple 3.1 (suite).

Voici un exemple de dérivation dans l'AP obtenu en appliquant la construction précédente à notre grammaire. On a mis en regard de cette dérivation ce qui lui correspond dans la grammaire : les règles sont appliquées "à l'envers" et on a encore utilisé un point qui permet de suivre la lecture "avec le doigt".

$$
\begin{array}{lll}
\cdot\, abbaa & a \cdot bbaa & (\varepsilon, abbaa) \overset{1}{\longmapsto} (a, bbaa) \\
& \overset{1}{\Longleftarrow} S \cdot bbaa & \overset{1}{\longmapsto} (S, bbaa) \\
& Sb \cdot baa & \overset{1}{\longmapsto} (Sb, baa) \\
& \overset{1}{\Longleftarrow} SbS \cdot baa & \overset{1}{\longmapsto} (SbS, baa) \\
& \overset{1}{\Longleftarrow} S \cdot baa & \overset{1}{\longmapsto} (S, baa) \\
& Sb \cdot aa & \overset{1}{\longmapsto} (Sb, aa) \\
& Sba \cdot a & \overset{1}{\longmapsto} (Sba, a) \\
& Sbaa \cdot & \overset{1}{\longmapsto} (Sbaa, \varepsilon) \\
& \overset{1}{\Longleftarrow} SbS \cdot & \overset{1}{\longmapsto} (SbS, \varepsilon) \\
& \overset{1}{\Longleftarrow} S \cdot & \overset{1}{\longmapsto} (S, \varepsilon)
\end{array}
$$

2.1 – Préfixes viables.

Une analyse partielle, comme ci–dessus, n'est vraissemblable que s'il peut aussi exister une dérivation $S \underset{d}{\Longrightarrow} \pi v$, car une dérivation, fût–elle à droite, a aussi un commencement !

Il serait intéressant de trouver une condition utilisable, que doit nécessairement satisfaire un "préfixe" $\pi \in (\mathcal{A}+\mathcal{V})^*$, pour qu'il existe une dérivation à droite $S \underset{d}{\Longrightarrow} \pi v$.

Forme générale des dérivations à droite.

Une dérivation $X \underset{d}{\Longrightarrow} \alpha$ de longueur > 0 se décompose en : $X \overset{1}{\Longrightarrow} \alpha'\beta \underset{d}{\Longrightarrow} \alpha' v$ où $v \in \mathcal{A}^*$ et $\alpha' v = \alpha$. Si $\beta = Y\alpha''$ pour $Y \in \mathcal{V}$, on peut préciser de la façon suivante :

$$X \overset{1}{\Longrightarrow} \alpha' Y \alpha'' \underset{d}{\Longrightarrow} \alpha' Y v'' \underset{d}{\Longrightarrow} \alpha' \gamma v''$$

Cette dérivation est *la composée* de

$$X \overset{1}{\Longrightarrow} \alpha' Y \alpha'' \underset{d}{\Longrightarrow} \alpha' Y v'' \quad \text{et de} \quad Y \underset{d}{\Longrightarrow} \gamma$$

Toute dérivation à droite, non triviale, est la composée de dérivations du type précédent :

$$\begin{array}{lllll}
X_0 & \overset{1}{\Longrightarrow} & \alpha'_1 X_1 \alpha''_1 & \underset{d}{\Longrightarrow} \alpha'_1 X_1 v_1 & \text{pour } v_1 \in \mathcal{A}^* \\
\ldots & \ldots & \ldots & \ldots \quad \ldots & \\
X_{n-1} & \overset{1}{\Longrightarrow} & \alpha'_n X_n \alpha''_n & \underset{d}{\Longrightarrow} \alpha'_n X_n v_n & \text{pour } v_n \in \mathcal{A}^* \\
X_n & \overset{1}{\Longrightarrow} & \alpha'_{n+1} \alpha''_{n+1} & \underset{d}{\Longrightarrow} \alpha'_{n+1} v_{n+1} & \text{pour } v_{n+1} \in \mathcal{A}^*
\end{array}$$

Commentaires.

- La composée de ces dérivations a la forme

$$X_0 \underset{d}{\Longrightarrow} \pi X_n v \overset{1}{\Longrightarrow} \pi\alpha'_{n+1}\alpha''_{n+1}v \underset{d}{\Longrightarrow} \pi\alpha'_{n+1}v_{n+1}v$$

où $\pi = \alpha'_1 \dots \alpha'_n$ et $v = v_n \dots v_1$.

- Les mots $\pi\alpha'_{n+1}$ obtenus ainsi sont les “préfixes” que nous cherchons à caractériser.

- Lorsque $\alpha''_{n+1} = \varepsilon$, on a $v_{n+1} = \varepsilon$ et $X_n \overset{1}{\Longrightarrow} \alpha'_{n+1}$ est la dernière dérivation élémentaire : $X_n \longrightarrow \alpha'_{n+1}$ est la dernière règle de la dérivation qui nous intéresse.

Augmentation d’une grammaire.

$S \overset{0}{\Longrightarrow} S$ ne rentre pas dans ce cadre ! La désignation effective de l’axiome S, se fait en *augmentant* G d’une nouvelle règle $\emptyset \longrightarrow S$ qui “produit” S à partir de l’ensemble vide ($\emptyset \notin \mathcal{V}$! ! !). Toute $S \underset{d}{\overset{k}{\Longrightarrow}} \gamma$ s’écrit alors $\emptyset \overset{1}{\Longrightarrow} S \underset{d}{\overset{k}{\Longrightarrow}} \gamma$.

Préfixes viables

$\pi\alpha' \in (\mathcal{A}+\mathcal{V})^*$ est appelé un *préfixe viable* ssi il existe une dérivation à droite

$$\emptyset \underset{d}{\Longrightarrow} \pi X v \overset{1}{\Longrightarrow} \pi\alpha'\alpha'' v$$

où v est nécessairement un élément de $\mathcal{A}^*$.

La règle $X \longrightarrow \alpha'\alpha''$ définit alors un *item valide* pour ce préfixe viable, que l’on notera $\cdot X \longrightarrow \alpha' \cdot \alpha''$, où $\cdot$ est un nouveau symbole.

2.2 – Item d'une grammaire.

On introduit donc un nouveau symbole $\cdot \notin \mathcal{A} + \mathcal{V}$, souvent appelé *pointeur.*

Les item

Toute $X \in \mathcal{V} + \{\emptyset\}$ et tout couple $\alpha \in (\mathcal{A} + \mathcal{V})^*$ et $\beta \in (\mathcal{A} + \mathcal{V})^*$ tels que $X \longrightarrow \alpha\beta$ est une règle de G définissent un *item* $\cdot X \longrightarrow \alpha \cdot \beta$ de G.

Pour chaque règle $X \longrightarrow \alpha$ de G, $\cdot X \longrightarrow \cdot \alpha$ est donc un item et les autres s'obtiennent par des applications de l'opération de *décalage* décrit par la règle

$$\cdot \xi \longrightarrow \xi \cdot$$

au second membre de $\cdot X \longrightarrow \cdot \alpha$.

Exemples.

Compte tenu du fait que $\cdot \emptyset = \emptyset$, la règle $\emptyset \longrightarrow S$ définit deux item :

- *l'item initial* : $\emptyset \longrightarrow \cdot S$
- *l'item final* : $\emptyset \longrightarrow S \cdot$

Une règle $X \longrightarrow \varepsilon$ définit le seul item $\cdot X \longrightarrow \cdot$ dont le second membre est une occurrence du pointeur.

La règle $X \longrightarrow aXaYb$ définit les item suivants :

$$\begin{array}{ll} \cdot X \longrightarrow \cdot aXaYb & \cdot X \longrightarrow aXa \cdot Yb \\ \cdot X \longrightarrow a \cdot XaYb & \cdot X \longrightarrow aXaY \cdot b \\ \cdot X \longrightarrow aX \cdot aYb & \cdot X \longrightarrow aXaYb \cdot \end{array}$$

Un item de la forme $\cdot X \longrightarrow \alpha \cdot$ est dit **complet**.

La description de la forme générale des dérivations à droite permet d'observer qu'**une dérivation à droite détermine un enchaînement d'item** :

- ou bien

$$\begin{array}{ccc} \emptyset & \longrightarrow & \cdot S \\ \cdot S & \longrightarrow & \alpha'_1 \cdot X_1 \alpha''_1 \\ \cdot X_1 & \longrightarrow & \alpha'_2 \cdot X_2 \alpha''_2 \\ \dots & & \dots \\ \cdot X_{n-1} & \longrightarrow & \alpha'_n \cdot X_n \alpha''_n \\ \cdot X_n & \longrightarrow & \alpha'_{n+1} \cdot \alpha''_{n+1} \end{array}$$

- ou bien, le cas spécial qui vient de l'augmentation

$$\emptyset \quad \longrightarrow \quad S \cdot$$

Réciproquement (dans la mesure où toute variable de G est productive), tout enchaînement d'item comme ci–dessus permet de construire des dérivations à droite assurant la viabilité d'un préfixe donné.

Finalement, avec ces notations, tout préfixe viable a la forme $\alpha'_1 \dots \alpha'_{n+1}$ (ou S), et peut s'obtenir par une succession de décalages et d'enchaînements d'item.
Cette observation nous mène directement à la construction de l'ε–AF suivant.

2.2.1 – AFD des item.

Considérons l'ε–AF $\mathbf{A}$, sur l'alphabet $\mathcal{A}+\mathcal{V}$, défini par :

- Les états sont les item ;
- l'entrée est l'item $\emptyset \longrightarrow \cdot S$;
- une transition(*) par $\xi \in \mathcal{A}+\mathcal{V}$ est définie sur les états de la forme $\cdot X \longrightarrow \alpha \cdot \xi\beta$ par un décalage

$$\boxed{\cdot X \longrightarrow \alpha \cdot \xi\beta} \xrightarrow{\quad\xi\quad} \boxed{\cdot X \longrightarrow \alpha\xi \cdot \beta}$$

- une ε–transition(*) est définie sur les états de la forme $\cdot X \longrightarrow \alpha \cdot Y\beta$, par

$$\boxed{\cdot X \longrightarrow \alpha \cdot Y\beta} \xrightarrow{\quad\varepsilon\quad} \boxed{\cdot Y \longrightarrow \cdot \gamma}$$

  pour toute règle $Y \longrightarrow \gamma$;
- les sorties ne jouent aucun rôle ici.

En considérant un enchaînement d'item comme ci–dessus, on obtient :

$$(\emptyset \longrightarrow \cdot S, \pi) \mathrel{\underset{\mathbf{A}}{\overset{*}{\vdash}}} (\cdot X \longrightarrow \alpha \cdot \beta, \varepsilon)$$

ssi

π est un préfixe viable et $\cdot X \longrightarrow \alpha \cdot \beta$ est valide pour π.

(*) On représente $r \in \delta(q, \xi)$ par l'arête $q \xrightarrow{\quad\xi\quad} r$ du graphe de transition.

Un ε–AF **A** est peu maniable et il est préférable de considérer l'AFD équivalent $D(\mathbf{A})$, obtenu en appliquant l'algorithme de détermination et en négligeant l'état vide.

L'AFD équivalent à **A** est donc la partie accessible de l'AFD défini de la façon suivante :

- les états sont des ensembles clos non vides d'item, nous les noterons par des lettres du genre I, J, éventuellement indexées ;
- l'état initial : $I_0 = cl(\emptyset \longrightarrow \cdot S)$;
- le résultat $I \bullet \xi$ de l'action de ξ sur I, est la clôture de la réunion des images des éléments de I,
- les sorties ne jouent aucun rôle ici.

L'AFD ainsi construit à partir d'une grammaire G (où on a fixé un axiome) s'appelle **AFD des item LR de G**.

LR est un sigle angloricain qui vient du fait que dans le type d'analyse où intervient un tel AFD, la lecture du mot à analyser se fait de gauche à droite (*Left to right scanning*) et qu'elle construit une dérivation à droite (*Rightmost derivation*).

Voyons maintenant deux exemples du calcul de l'AFD des item LR d'une grammaire.

L'axiome est la variable de la première règle.

Exemple 1.

Soit G_1 la grammaire :

$$\begin{aligned}
&1 : S \longrightarrow (S \wedge S)\\
&2 : S \longrightarrow \neg S\\
&3 : S \longrightarrow \mathbf{id}
\end{aligned}$$

AFD des item LR de G_1.

Chaque nouvel état I_i est accompagné de l'ensemble Σ_i des symboles ξ tels que $I_i \bullet \xi \neq \emptyset$.

$$\begin{aligned}
I_0 = \;& \{\emptyset \longrightarrow \cdot S, && \Sigma_0 = S + (+\neg + \mathbf{id}\\
& \cdot S \longrightarrow \cdot(S \wedge S), \cdot S \longrightarrow \cdot\neg S, \cdot S \longrightarrow \cdot\,\mathbf{id}\}\\
I_1 = \;& I_0 \bullet S && \Sigma_1 = \emptyset\\
= \;& \{\emptyset \longrightarrow S\,\cdot\}\\
I_2 = \;& I_0 \bullet (&& \Sigma_2 = S + (+\neg + \mathbf{id}\\
= \;& \{\cdot\, S \longrightarrow (\cdot\, S \wedge S),\\
& \cdot S \longrightarrow \cdot(S \wedge S), \cdot S \longrightarrow \cdot\neg S, \cdot S \longrightarrow \cdot\,\mathbf{id}\}\\
I_3 = \;& I_0 \bullet \neg && \Sigma_3 = S + (+\neg + \mathbf{id}\\
= \;& \{\cdot\, S \longrightarrow \neg \cdot S,\\
& \cdot S \longrightarrow \cdot(S \wedge S), \cdot S \longrightarrow \cdot\neg S, \cdot S \longrightarrow \cdot\,\mathbf{id}\}\\
I_4 = \;& I_0 \bullet \mathbf{id} && \Sigma_4 = \emptyset\\
= \;& \{\cdot\, S \longrightarrow \mathbf{id}\,\cdot\}\\
I_5 = \;& I_2 \bullet S && \Sigma_5 = \wedge\\
= \;& \{\cdot\, S \longrightarrow (S \cdot \wedge S)\}\\
I_2 = \;& I_2 \bullet (
\end{aligned}$$

$$I_3 = I_2 \bullet \neg$$

$$I_4 = I_2 \bullet \mathbf{id}$$

$$\begin{aligned} I_6 &= I_3 \bullet S & \Sigma_6 = \emptyset \\ &= \{\cdot S \longrightarrow \neg S \cdot\} \end{aligned}$$

$$I_2 = I_3 \bullet ($$

$$I_3 = I_3 \bullet \neg$$

$$I_4 = I_3 \bullet \mathbf{id}$$

$$\begin{aligned} I_7 &= I_5 \bullet \wedge & \Sigma_7 = S + (+\neg + \mathbf{id} \\ &= \{\cdot S \longrightarrow (S \wedge \cdot S), \\ &\quad \cdot S \longrightarrow \cdot(S \wedge S), \cdot S \longrightarrow \cdot \neg S, \cdot S \longrightarrow \cdot \mathbf{id}\} \end{aligned}$$

$$\begin{aligned} I_8 &= I_7 \bullet S & \Sigma_8 =) \\ &= \{\cdot S \longrightarrow (S \wedge S \cdot)\} \end{aligned}$$

$$I_2 = I_7 \bullet ($$

$$I_3 = I_7 \bullet \neg$$

$$I_4 = I_7 \bullet \mathbf{id}$$

$$\begin{aligned} I_9 &= I_8 \bullet) & \Sigma_9 = \emptyset \\ &= \{\cdot S \longrightarrow (S \wedge S) \cdot\} \end{aligned}$$

Exemple 2.

Soit G_2 la grammaire :

$$\begin{array}{l}
1 : E \longrightarrow E \oplus T \\
2 : E \longrightarrow T \\
3 : T \longrightarrow T * F \\
4 : T \longrightarrow F \\
5 : F \longrightarrow (E) \\
6 : F \longrightarrow \mathbf{id}
\end{array}$$

AFD des item LR de G_2.

$$\begin{array}{rll}
I_0 = & \{\emptyset \longrightarrow \cdot E, & \Sigma_0 = E + T + F + (+\mathbf{id} \\
& \cdot E \longrightarrow \cdot E \oplus T, \cdot E \longrightarrow \cdot T, & \\
& \cdot T \longrightarrow \cdot T * F, \cdot T \longrightarrow \cdot F, & \\
& \cdot F \longrightarrow \cdot (E), \cdot F \longrightarrow \cdot \mathbf{id}\} & \\
I_0 \bullet E = & I_1 & \Sigma_1 = \oplus \\
= & \{\emptyset \longrightarrow E \cdot, & \\
& \cdot E \longrightarrow E \cdot \oplus T\} & \\
I_0 \bullet T = & I_2 & \Sigma_2 = * \\
= & \{\cdot E \longrightarrow T \cdot, & \\
& \cdot T \longrightarrow T \cdot * F\} & \\
I_0 \bullet F = & I_3 & \Sigma_3 = \emptyset \\
= & \{\cdot T \longrightarrow F \cdot\} & \\
I_0 \bullet (= & I_4 & \Sigma_4 = E + T + F + (+\mathbf{id} \\
= & \{\cdot F \longrightarrow (\cdot E), & \\
& \cdot E \longrightarrow \cdot E \oplus T, \cdot E \longrightarrow \cdot T, & \\
& \cdot T \longrightarrow \cdot T * F, \cdot T \longrightarrow \cdot F, & \\
& \cdot F \longrightarrow \cdot (E), \cdot F \longrightarrow \cdot \mathbf{id}\} &
\end{array}$$

$$\begin{aligned} I_0 \bullet \mathbf{id} &= I_5 && \Sigma_5 = \emptyset \\ &= \{\cdot F \longrightarrow \mathbf{id}\cdot\} \end{aligned}$$

$$\begin{aligned} I_1 \bullet \oplus &= I_6 && \Sigma_6 = T + F + (+\mathbf{id} \\ &= \{\cdot E \longrightarrow E \oplus \cdot T, \\ &\quad \cdot T \longrightarrow \cdot T * F, \cdot T \longrightarrow \cdot F, \\ &\quad \cdot F \longrightarrow \cdot(E), \cdot F \longrightarrow \cdot \mathbf{id}\} \end{aligned}$$

$$\begin{aligned} I_2 \bullet * &= I_7 && \Sigma_7 = F + (+\mathbf{id} \\ &= \{\cdot T \longrightarrow T * \cdot F, \\ &\quad \cdot F \longrightarrow \cdot(E), \cdot F \longrightarrow \cdot \mathbf{id}\} \end{aligned}$$

$$\begin{aligned} I_4 \bullet E &= I_8 && \Sigma_8 =) + \oplus \\ &= \{\cdot F \longrightarrow (E\,\cdot), \\ &\quad \cdot E \longrightarrow E \cdot \oplus T\} \end{aligned}$$

$$I_4 \bullet T = I_2$$

$$I_4 \bullet F = I_3$$

$$I_4 \bullet (= I_4$$

$$I_4 \bullet \mathbf{id} = I_5$$

$$\begin{aligned} I_6 \bullet T &= I_9 && \Sigma_9 = * \\ &= \{\cdot E \longrightarrow E \oplus T\,\cdot, \\ &\quad \cdot T \longrightarrow T \cdot *F\} \end{aligned}$$

$$I_6 \bullet F = I_3$$

$$I_6 \bullet (= I_4$$

$$I_6 \bullet \mathbf{id} = I_5$$

$$\begin{aligned} I_7 \bullet F &= I_{10} & \Sigma_{10} = \emptyset \\ &= \{\cdot T \longrightarrow T * F \cdot\} \end{aligned}$$

$$I_7 \bullet (= I_4$$

$$I_7 \bullet \mathbf{id} = I_5$$

$$\begin{aligned} I_8 \bullet) &= I_{11} & \Sigma_{11} = \emptyset \\ &= \{\cdot F \longrightarrow (E) \cdot\} \end{aligned}$$

$$I_8 \bullet \oplus = I_6$$

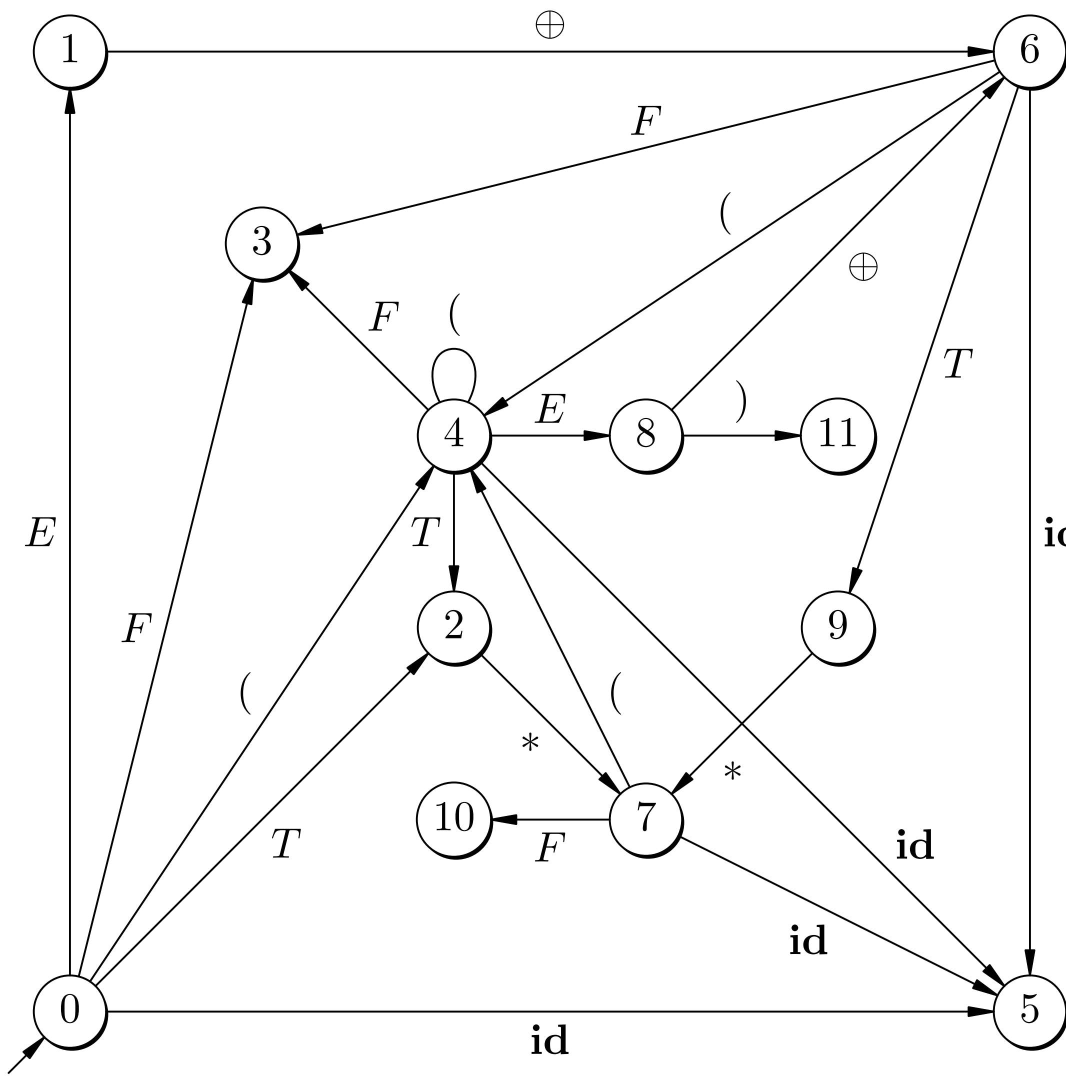
$$I_9 \bullet * = I_7$$

Voici la table de l'AFD des item de G_2 ainsi obtenu.

Pour simplifier, nous avons représenté chaque état I_i par son indice i, et nous n'avons rien mis dans les cases devant contenir l'état vide !

I	$I \bullet \oplus$	$I \bullet *$	$I \bullet ($	$I \bullet)$	$I \bullet \mathbf{id}$	$I \bullet E$	$I \bullet T$	$I \bullet F$
0			4		5	1	2	3
1	6							
2		7						
3								
4			4		5	8	2	3
5								
6			4		5		9	3
7			4		5			10
8	6			11				
9		7						
10								
11								

Et voici, à titre de curiosité, le graphe de transition de cet AFD : on ne manquera pas cependant de s'y reporter pour observer qu'il vérifie bien les quelques propriétés intéressantes qui seront signalées par la suite.



Propriétés de l'AFD des item.

1) $I \bullet \xi$ est défini ssi il existe $(\cdot X \longrightarrow \alpha \cdot \xi\beta) \in I$.
2) $(\cdot X \longrightarrow \alpha \cdot \xi\beta) \in I$ implique $(\cdot X \longrightarrow \alpha\xi \cdot \beta) \in I \bullet \xi$.
3) En conséquence, toutes les transitions aboutissant à un état donné sont étiquetées par le même symbole, c'est–à–dire : si $I \bullet \xi = J \bullet \eta$ alors $\xi = \eta$.
4) Soient I_0 l'état initial et $\pi \in (\mathcal{A} + \mathcal{V})^*$, alors :
   – π est un préfixe viable ssi $I_0 \bullet \pi$ est défini.

5 Lorsque $I = I_0 \bullet \pi$ est défini :
   – I est l'ensemble des item valides pour π ;
   – si $(\cdot X \longrightarrow \alpha \cdot \xi\beta) \in I$ alors $\pi\xi$ est viable ;
   – si $(\cdot X \longrightarrow \alpha \cdot) \in I$ alors $\pi = \pi'\alpha$ pour un $\pi' \in (\mathcal{A} + \mathcal{V})^*$ et $\pi' X$ est un préfixe viable.

2.3 – L'automate à pile LR.

Nous allons définir un AP plus fin que le précédent, reconnaissant $\mathcal{L}(G, S)$ et dans lequel toute dérivation d'un mot $u \in \mathcal{L}(G, S)$ décrit, mais à l'envers, une dérivation à droite de u dans G à partir de S.

L'alphabet de pile, noté Π, est l'ensemble des états de l'AFD des item de la grammaire G, pour l'axiome $S \in \mathcal{V}$.

Or, la pile de cet AP est toujours égale à un tel chemin, issu de l'état initial I_0.
D'après la propriété 3) ci–dessus, un chemin dans l'AFD qui conduit d'un état I à un état J détermine entièrement un mot $\alpha \in (\mathcal{A} + \mathcal{V})^*$ tel que $I \bullet \alpha = J$.

Ces deux données sont donc équivalentes ! mais il est intéressant, pour des raisons pratiques, de les considérer simultanément.

• Pour tout $I \in \Pi$ et tout $\xi \in \mathcal{A} + \mathcal{V}$ tels que $J = I \bullet \xi$ soit défini, on introduit le nouveau symbole noté $[I\xi J]$.

• *La représentation de la pile* s'obtient en considérant la réunion de Π et de l'ensemble des symboles précédents : pour tout $I \in \Pi$ et tout $\alpha \in (\mathcal{A}+\mathcal{V})^*$ tels que $J = I \bullet \alpha$ soit défini, on note $[I\alpha J] \in \Pi^*$ le mot défini par la récurrence suivante :

- $[II] = I$;
- $[I\xi J]$ est un symbole ;
- $[I\alpha\xi J] = [I\alpha K][K\xi J]$, pour $\alpha \neq \varepsilon$ et $K = I \bullet \alpha$.

Lorsque $[I\alpha J]$ est défini, il décrit un chemin dans l'AFD. Par exemple, dans celui de la grammaire G_2 :

$$[I_0 E \oplus T I_9] = [I_0 E I_1][I_1 \oplus I_6][I_6 T I_9]$$

• Lorsque $[I\alpha J]$ est défini, on peut effectivement le décomposer en ses deux constituants :

1) le mot α, qui est appelé *la pile* ;
2) le chemin $ch([I\alpha J]) \in \Pi\Pi^*$ parcouru dans l'AFD, qui est appelé *la pile des états* :
   - $ch(I) = I$;
   - $ch([I\xi J]) = IJ$;
   - $ch([I\alpha\xi J]) = ch([I\alpha K])J$, pour $\alpha \neq \varepsilon$ et $K = I \bullet \alpha$.

   Par exemple $ch([I_0 E \oplus T I_9]) = I_0 I_1 I_6 I_9$.

Voici la description de l'AP d'analyse LR.

- L'alphabet de pile Π est l'ensemble des états de l'AFD des item, mais **nous utiliserons la représentation de la pile qui a été définie ci–dessus** ;
- les états sont ceux de l'AFD des item LR de G : **l'état courant est le sommet de la pile et nous ne noterons donc pas dans la configuration** ;
- l'état initial est $I_0 = cl(\emptyset \to \cdot S)$;
- le seul état final est $I_1 = cl(\emptyset \to S\,\cdot) = I_0 \bullet S$;
- les transitions sont de deux types :

  Les transitions de décalage. Pour tout $x \in \mathcal{A}$ et tout état I tel que $J = I \bullet x$ est défini, on a la transition suivante :

  $$I \xrightarrow{(\varepsilon \multimap [IxJ]; x)} J$$

  dite de *décalage* (mieux, de *lecture*) de x.

  Les transitions de réduction. Pour toute règle $X \to \alpha$ et tout état I tels que $(\cdot\, X \to \alpha\,\cdot) \in I$, pour tout état K tel que $K \bullet \alpha = I$ et tel que $J = K \bullet X$ est défini, on a la transition suivante

  $$I \xrightarrow{([K\alpha I] \multimap [KXJ]; \varepsilon)} J$$

  dite de *réduction* de la règle $X \to \alpha$.

On peut vérifier que

Pour tout $u \in \mathcal{A}^*$: $u \in \mathcal{L}(G, S)$ *ssi* $(I_0, u) \vdash^{*} ([I_0 S I_1], \varepsilon)$

Les données nécessaires à la définition des transitions sont consignées dans une table $Action(I, \xi)$:

Table LR

Les $Action(\_, \_)$ sont les plus petits ensembles tels que

– $J \in Action(I, \xi)$ si $I \bullet \xi = J$ pour $\xi \in \mathcal{A} + \mathcal{V}$
– $(X \longrightarrow \alpha) \in Action(I, \varepsilon)$ si $(\cdot X \longrightarrow \alpha \cdot) \in I$ pour $X \in \mathcal{V}$
– $Acc \in Action(I, \varepsilon)$ si $(\emptyset \longrightarrow S \cdot) \in I$

C'est donc la table de transition de l'AFD des item LR de G enrichi d'une colonne pour ε qui comporte, pour chaque état, l'ensemble des "item complets" qu'il contient.

• Les états sont énumérés $I_0, \ldots, I_n, \ldots$: on note di (décaler en i) au lieu de I_i dans la table de transition.

• Les règles sont codées par des entiers : on note rk (réduire par la règle k), la réduction par $\emptyset \longrightarrow S$, qui n'est pas réellement faite, est codée par Acc.

Ceci donne la description suivante de la table LR :

– si $I_i \bullet \xi = I_j$ alors $dj \in Action(i, \xi)$
– si $(\cdot X \longrightarrow \alpha \cdot) \in I_i$ avec $X \in \mathcal{V}$ alors
$$rk \in Action(i, \varepsilon)$$
où k est le numéro de la règle $X \longrightarrow \alpha$
– si $(\emptyset \longrightarrow S \cdot) \in I_i$ alors $Acc \in Action(i, \varepsilon)$.

Exemple 1 (suite).

Table LR de G_1.

	ε	(	$\wedge$	)	$\neg$	**id**	S
0		$d2$			$d3$	$d4$	$d1$
1	Acc						
2		$d2$			$d3$	$d4$	$d5$
3		$d2$			$d3$	$d4$	$d6$
4	r3						
5			$d7$				
6	r2						
7		$d2$			$d3$	$d4$	$d8$
8				$d9$			
9	r1						

On constate que cette table définit un AP déterministe.

L'analyse LR de $\neg(\mathbf{id} \wedge \neg\mathbf{id})$ dans G_1.

Pile	Pile d'états	Entrée	Act.
ε	0	$\neg(\mathbf{id} \wedge \neg\mathbf{id})$	$d3$
$\neg$	03	$(\mathbf{id} \wedge \neg\mathbf{id})$	$d2$
$\neg($	032	$\mathbf{id} \wedge \neg\mathbf{id})$	$d4$
$\neg(\mathbf{id}$	0324	$\wedge\neg\mathbf{id})$	$r3$
$\neg(S$	0325	$\wedge\neg\mathbf{id})$	$d7$
$\neg(S\wedge$	03257	$\neg\mathbf{id})$	$d3$
$\neg(S \wedge \neg$	032573	$\mathbf{id})$	$d4$
$\neg(S \wedge \neg\mathbf{id}$	0325734	$)$	$r3$
$\neg(S \wedge \neg S$	0325736	$)$	$r2$
$\neg(S \wedge S$	032578	$)$	$d9$
$\neg(S \wedge S)$	0325789	ε	$r1$
$\neg S$	036	ε	$r2$
S	01	ε	Acc

Exemple 2 (suite).

Table LR de G_2.

	ε	$\oplus$	$*$	(	)	**id**	E	T	F
0				$d4$		$d5$	$d1$	$d2$	$d3$
1	Acc	$d6$							
2	$r2$		$d7$						
3	$r4$								
4				$d4$		$d5$	$d8$	$d2$	$d3$
5	$r6$								
6				$d4$		$d5$		$d9$	$d3$
7				$d4$		$d5$			$d10$
8		$d6$			$d11$				
9	$r1$		$d7$						
10	$r3$								
11	$r5$								

Cette table ne définit pas un AP déterministe : un conflit "décalage–réduction" se produit dans les états 1, 2 et 9. Il faudra être plus prévoyant pour les résoudre ...

Une analyse LR de $\mathbf{id} \oplus \mathbf{id} * \mathbf{id}$ dans G_2.

Pile	Pile d'états	Entrée	Act.
ε	0	$\mathbf{id} \oplus \mathbf{id} * \mathbf{id}$	$d5$
$\mathbf{id}$	0 5	$\oplus \mathbf{id} * \mathbf{id}$	$r6$
F	0 3	$\oplus \mathbf{id} * \mathbf{id}$	$r4$
T	0 2	$\oplus \mathbf{id} * \mathbf{id}$	$r2$
E	0 1	$\oplus \mathbf{id} * \mathbf{id}$	$d6$
$E\oplus$	0 1 6	$\mathbf{id} * \mathbf{id}$	$d5$
$E \oplus \mathbf{id}$	0 1 6 5	$* \mathbf{id}$	$r6$
$E \oplus F$	0 1 6 3	$* \mathbf{id}$	$r4$
$E \oplus T$	0 1 6 9	$* \mathbf{id}$	$d7$
$E \oplus T*$	0 1 6 9 7	$\mathbf{id}$	$d5$
$E \oplus T * \mathbf{id}$	0 1 6 9 7 5	ε	$r6$
$E \oplus T * F$	0 1 6 9 7 10	ε	$r3$
$E \oplus T$	0 1 6 9	ε	$r1$
E	0 1	ε	Acc