Langages écrits par un code infinitaire. Théorème du défaut

By Do Long Van

1. Notations et définitions

Soit A un alphabet non-vide. On note A^* le monoïde libre engendrè par A, i.e. l'ensemble de tous les mots finis sur A, y compris le mot vide noté ε, muni de l'opèration de concaténation. La longueur d'un mot f de A^* est noté $|f|$, et pour tout n, $1 \leqq n \leqq |f|$, $f(n)$ désigne la n-ième lettre du mot f. L'ensemble des mots infinis sur A est noté A^ω. Chaque mot u de A^ω est de longueur $|u| = \omega = \mathrm{Card}\,\mathbf{N}$ et est un application $u = \mathbf{N}^+ \to A$ qu'on ècrit souvent sous la forme $u = u(1)u(2)\ldots$. On pose $A^\infty = A^* \cup A^\omega$ et on appelle *langage infinitaire* (resp. *finitaire, purement infinitaire*) toute partie X de A^∞ (resp. A^*, A^ω). Si $X \subseteqq A^*$, X^ω désigne l'ensemble des mots infinis de la forme $x_1 x_2 \ldots$ avec $x_i \in X$ $(i = 1, 2, \ldots)$. En particulier, pour $f \in A^*$, $\{f\}^* = ff\ldots$. Pour rendre plus clair, dans la suite on notera souvent par $f, g, h, \ldots$ les mots finis, par $u, v, w, \ldots$ les mots infinis, et par $\alpha, \beta, \gamma, \ldots$ les mots dont la longueur est finie ou infinie.

On munit A^∞ d'un produit prolongeant celui de A^* de la manière suivante :

$$\forall u \in A^\omega \forall \alpha \in A^\infty : \; u\alpha = u;$$

$$\forall f \in A^* \forall u \in A^\omega : \; (fu)(n) = \begin{cases} f(n) & \text{pour} \quad 1 \leqq n \leqq |f|, \\ u(n - |f|) & \text{pour} \quad n > |f|. \end{cases}$$

On vérifie sans peine que A^∞ est alors un monoïde.

Pour toute partie X de A^∞ on note $X_{fin} = X \cap A^*$, $X_{inf} = X \cap A^\omega$ et on définit:

$$\begin{cases} X^{(0)} = \{\varepsilon\}, \\ X^{(1)} = X, \\ X^{(k)} = X_{fin} X^{(k-1)}, \quad k \geqq 2. \end{cases}$$

Alors, pour $k \geqq 1$, $X^{(k)} = X_{fin}^k \cup X_{fin}^{k-1} X_{inf}$, et par conséquent chaque élément α de $X^{(k)}$ peut se présenter sous l'une des deux formes :

(i) $\alpha = x_1 \ldots x_k$ avec $x_i \in X_{fin}$ $(i = 1, \ldots, k)$;

(ii) $\alpha = x_1 \ldots x_k$ avec $x_i \in X_{fin}$ $(i = 1, \ldots, k-1)$, $x_k \in X_{inf}$.

Comme d'habitude on note X^* le sous-monoïde de A^∞ engendré par X et pose $X^+ = X^* - \{\varepsilon\}$. On a évidemment

$$X^+ = \bigcup_{k=1}^{\infty} X^{(k)}.$$

Une partie X de A^∞ est un *code infinitaire* sur A (cf. [3]) si chaque élément α de X^+ peut se présenter uniquement sous l'une des deux formes (i) et (ii) pour un certain k, *ou*, d'une façon équivalente, si pour tous $x_1 \ldots x_n \in X^{(n)}$, $x'_1 \ldots x'_m \in X^{(m)}$, l'égalité

$$x_1 \ldots x_n = x'_1 \ldots x'_m$$

implique $n=m$ et $x_i = x'_i$ $(i=1, \ldots, n)$. Dans la suite, sauf spécification contraire, le mot « code » désignera un code infinitaire.

On appele *quasi-libre* (cf [4]) tout sous-monoïde M de A^∞ engendré par un code. Le code qui engendre M est appelé la *base* de M. L'ensemble de tous les sous-monoïdes quasi-libres est noté QL.

Étant donné un sous-monoïde M de A^∞ nous introduisons sur M_{inf} une relation linéaire transitive, notée « $\prec$ », de la manière suivante :

$$u \prec v \Leftrightarrow \exists f \in (M_{fin} - \varepsilon) : v = fu.$$

Si $u \prec v$ nous disons que u *est contenu* dans v ou v *contient* u. Un élément u de M_{inf} est dit *maximal* s'il n'existe aucun élément v de M_{inf} tel que $u \prec v$. On dit que le sous-monoïde M satisfait à la *condition de maximalité* si tout élément non-maximal de M_{inf} est contenu dans un certain élément maximal de M_{inf}. On appelle *chaîne* toute suite croissante $u_1 \prec u_2 \prec \ldots$ d'éléments de M_{inf} ordonnée par « $\prec$ ». Une chaîne peut être finie ou infinie. On dit que le sous-monoïde M satisfait à la *condition de chaîne finie* si toute chaîne d'éléments de M_{inf} est finie. La condition de chaîne finie implique évidemment celle de maximalité, mais l'implication inverse est fausse ([6], Exemple 2).

Un ensemble générateur d'un monoïde M est *minimum* s'il est inclus dans tout ensemble générateur de M. L'ensemble générateur minimum d'un monoïde, s'il existe, est clairement unique.

On appelle *distincte* toute partie X de A^∞ telle que $X_{inf} \cap X^+_{fin} X_{inf} = \emptyset$. Clairement sont distinctes toute partie finitaire ainsi que tout code infinitaire.

2. Langages écrits par un code infinitaire

Étant donnée une classe C de codes on dit qu'un langage X est *écrit par un code de la classe* C ou *C-éctit par un code* s'il existe un $Y \in C$ tel que $X \subseteq Y^*$. Dans le cas ou C coïncide avec la classe de tous les codes on dit simplement que X est *écrit par un code*.

Tout langage finitaire est clairement écrit par un code tandis qu'il existe des langages infinitaires qui ne peuvent écrit par aucun code. L'exemple d'un tel langage est A^∞ pour A non-vide (cf [5], Corollaire 2). Nous caractérisons dans cette section les langages qui sont écrits par un code (code préfixe, code suffixe, code bipréfixe, code normal).

Rappelons que pour toutes parties X et Y de A^∞ on définit

$$Y^{-1}X = \{\alpha \in A^\infty \mid \exists \beta \in Y\colon (\beta\alpha \in X) \,\&\, (|\beta| = \omega \to \alpha = \varepsilon\},$$

$$XY^{-1} = \{\alpha \in A^\infty \mid \exists \beta \in Y\colon \alpha\beta \in X\}.$$

Maintenant, pour toute partie X de A^∞, on pose

$$LB(X) = X^{-1}X \cap XX^{-1};$$

$$UD(X) = X^{-1}X;$$

$$UG(X) = XX^{-1};$$

$$BU(X) = X^{-1}X \cup XX^{-1}.$$

Un sous-monoïde M de A^∞ est par définition *libérable* (resp. *unitaire à droite, unitaire à gauche, biunitaire*) ssi $LB(M) \subseteqq M$ (resp. $UD(M) \subseteqq M$, $UG(M) \subseteqq M$, $BU(M) \subseteqq M$). Notant par LB (resp. UD, UG, BU) la classe de tous les sous-monoïdes libérables (resp. unitaires à droite, unitaires à gauche, biunitaires) on a donc

$$M \in Z \Leftrightarrow Z(M) \subseteqq M \quad \text{pour} \quad Z \in \{LB, UD, UG, BU\}.$$

On vérifie sans peine que, pour toute famille $\{X_i \mid i \in I\}$ de parties de A^∞ et pour tout $Z \in \{LB, UD, UG, BU\}$, $Z(\bigcap_{i\in I} X_i) \subseteqq \bigcap_{i\in I} Z(X_i)$. Les classes de sous-monoïdes LB, UD, UG, BU sont donc fermées par intersection. Comme A^∞ est biunitaire et par conséquent unitaire à droite, unitaire à gauch, libérable, il existe donc, pour toute partie X de A^∞, un plus petit sous-monoïde libérable (unitaire à droite, unitaire à gauche, biunitaire) de A^∞ contenant X qu'on note par $\overline{LB}(X)$ (resp. $\overline{UD}(X)$, $\overline{UG}(X)$, $\overline{BU}(X)$).

On associe maintenant à chaque Z de $\{LB, UD, UG, BU\}$ et chaque partie X de A^∞ une suite croissante $M_n^Z(X)$ de sous-monoïdes de A^∞ ainsi définie :

$$M_0^Z(X) = X^*, \quad M_{n+1}^Z(X) = [Z(M_n^Z(X))]^*, \quad n \geqq 0.$$

Posons:

$$M^Z(X) = \bigcup_{n\geqq 0} M_n^Z(X).$$

Proposition 2.1. *Pour tout* $Z \in \{LB, UD, UG, BU\}$ *et pour toute* X *de* A^∞, *on a :*

$$M^Z(X) = \overline{Z}(X).$$

Preuve. On a $M_n^Z(X) \subseteqq \overline{Z}(X)$ pour tout n. En effet, ceci est évident pour $n=0$. Puis, si $M_n^Z(X) \subseteqq \overline{Z}(X)$, on a :

$$M_{n+1}^Z(X) = [Z(M_n^Z(X))]^* \subseteqq [Z(\overline{Z}(X))]^* \subseteqq [\overline{Z}(X)]^* = \overline{Z}(X)$$

car $\overline{Z}(X) \in Z$. Donc $M^Z(X) \subseteqq \overline{Z}(X)$. D'autre part, comme la suite $M_n^Z(X)$ est croissante, on a :

$$Z(M^Z(X)) = \bigcup_{n\geqq 0} Z(M_n^Z(X)) \subseteqq \bigcup_{n\geqq 0} [Z(M_n^Z(X))]^* = \bigcup_{n\geqq 0} M_{n+1}^Z(X) = M^Z(X),$$

ce qui montre que $M^Z(X) \in Z$. Par minimalité de $\overline{Z}(X)$, il en resulte $\overline{Z}(X) \subseteqq M^Z(X)$. Ainsi $M^Z(X) = \overline{Z}(X)$.

Maintenant, si un langage X est écrit par un code, alors, du fait que QL est fermée par intersection ([6], Corollaire 6), on peut parler du plus petit sous-monoïde quasi-libre contenant X que l'on note $\overline{QL}(X)$.

Le théorème suivant caractérise les langages qui sont écrits par un code :

Théorème 2.2. *Pour tout langage infinitaire X, les conditions suivantes sont équivalentes :*

(i) *X est écrit par un code ;*
(ii) *Il existe $\overline{QL}(X)$ et $\overline{QL}(X)=M^{LB}(X)$;*
(iii) *$M^{LB}(X)$ satisfait à la condition de chaîne finie ;*
(iv) *$M^{LB}(X)$ satisfait à la condition de maximalité ;*
(v) *$M^{LB}(X)$ possède un ensemble générateur minimum distinct.*

Preuve. (i)⇒(ii) : supposons que X soit écrit par un code. Comme il a été dit plus haut, il existe $\overline{QL}(X)$. $\overline{QL}(X)$ est alors libérable car tout sous-monoïde quasi-libre est libérable ([6], Proposition 2). Utilisant la Proposition 2.1 avec $Z=LB$ et la minimalité de $\overline{LB}(X)$ on a $M^{LB}(X)=\overline{LB}(X)\subseteq\overline{QL}(X)$. Mais alors, par le Corollaire 4 en [6], $M^{LB}(X)$ est quasi-libre. Par minimalité de $\overline{QL}(X)$ on en déduit $\overline{QL}(X)\subseteq M^{LB}(X)$. Ainsi $\overline{QL}(X)=M^{LB}(X)$.

(ii)⇒(iii) puisque tout sous-monoïde quasi-libre satisfait à la condition de chaîne finie ([6], Proposition 3).

(iii)⇔(iv)⇔(v) est immédiat du Corollaire 1 en [6] et du fait que $M^{LB}(X)$, par la Proposition 2.1, est libérable.

(v)⇒(i) : Supposons $M^{LB}(X)$ possède un ensemble générateur minimum distinct. En vertu du Théoréme 1 en [6], $M^{LB}(X)$ est quasi-libre. X est donc écrit par un code qui est la base de $M^{LB}(X)$.

Une partie X de A^∞ est *préfixe* (*suffixe*) si aucun mot de X n'est facteur gauche (resp. facteur droit) propre d'un mot de X. La partie X est *bipréfixe* si elle est à la fois préfixe et suffixe. Toute partie préfixe (suffixe, bipréfixe) $X\neq\{\varepsilon\}$ est un code appelé *code préfixe* (resp. *suffixe, bipréfixe*).

Un code X est *normal* si $X_{fin}^+X_{inf}\cap X_{fin}^\omega=\emptyset$. Sont normaux évedemment tout code finitaire ainsi que tout code préfixe.

Un sous-monoïde M de A^∞ est dit *régulier* si $M_{inf}\cap M_{fin}^\omega=\emptyset$. Tout sous-monoïde régulier satisfait à la condition de chaîne finie, mais la réciproque n'est pas vraie (cf. [6], Exemple 4 (suite)).

Si un langage X est écrit par un code préfixe (suffixe, bipréfixe, normal), alors, parce que la classe des sous-monoïdes engendrés par codes préfixes (suffixes, bipréfixes, normaux) est fermée par intersection, on peut parler du plus petit sous-monoïde de cette classe qui contient X. Celui-ci est noté $\overline{P}(X)$ (resp. $\overline{S}(X)$, $\overline{BP}(X)$, $\overline{N}(X)$).

Remarque. Pour toute partie X de A^* on a

$$\overline{LB}(X)=\overline{QL}(X)=\overline{N}(X)=\overline{L}(X);$$

$$\overline{UD}(X)=\overline{P}(X)=\overline{PF}(X);\quad \overline{UG}(X)=\overline{S}(X)=\overline{SF}(X);$$

$$\overline{BU}(X)=\overline{BP}(X)=\overline{BPF}(X),$$

où $\overline{L}(X)$ est le plus petit sous-monoïde libre contenant X ; $\overline{PF}(X)$ (resp. $\overline{SF}(X)$, $\overline{BPF}(X)$) est le plus petit sous-monoïde qui est engendré par un code finitaire préfixe (resp. suffixe, bipréfixe) et qui contient X.

Théorème 2.3. *Pour tout langage infinitaire X, les conditions suivantes sont équivalentes :*
(i) *X est écrit par un code normal ;*
(ii) *Il existe $\overline{N}(X)$ et $\overline{N}(X)=M^{LB}(X)$;*
(iii) *$M^{LB}(X)$ est régulier.*

Preuve. (i)⇒(ii) : si X est écrit par un code normal, alors il existe $\overline{N}(X)$ qui, par le Théorème 2 en [6], est libérable. Par la Proposition 2.1 et par minimalité de $\overline{LB}(X)$, $M^{LB}(X)=\overline{LB}(X)\subseteq\overline{N},(X)$ d'où, par le Corollaire 8 en [6], $M^{LB}(X)$ est aussi engendré par un code normal. Donc $\overline{N}(X)=M^{LB}(X)$.

(ii)⇒(iii) est évident.

(iii)⇒(i) : supposons $M^{LB}(X)$ régulier. Par la Proposition 2.1, $M^{LB}(X)$ est libérable. En vertu du Théorème 2 en [6], $M^{LB}(X)$ est engendré par un code normal. Donc X est écrit par un code normal qui est la base de $M^{LB}(X)$.

Théorème 2.4. *Pour tout langage infinitaire X, les conditions suivantes sont équivalentes :*
(i) *X est écrit par un code préfixe (bipréfixe, suffixe) ;*
(ii) *Il existe $\overline{P}(X)$ (resp. $\overline{BP}(X)$, $\overline{S}(X)$) et $\overline{P}(X)=M^{UD}(X)$ (resp. $\overline{BP}(X)=M^{BU}(X)$, $\overline{S}(X)=M^{UG}(X)$) ;*
(iii) *$M^{UD}(X)$ (resp. $M^{BU}(X)$) est régulier ;*
(iv) *$M^{UD}(X)$ (resp. $M^{BU}(X)$, $M^{UG}(X)$) satisfait à la condition de chaîne finie ;*
(v) *$M^{UD}(X)$ (resp. $M^{BU}(X)$, $M^{UG}(X)$) satisfait à la condition de maximalité ;*
(vi) *$M^{UD}(X)$ (resp. $M^{BU}(X)$, $M^{UG}(X)$) possède un ensemble générateur minimum distinct.*

Preuve. Nous ne traitons que le cas de codes préfixes.

(i)⇒(ii) : si X est écrit par un code préfixe, il existe $\overline{P}(X)$ qui, par le Théorème 3 en [6], est unitaire à droite. En vertu de la Proposition 2.1 et de la minimalité de $\overline{UD}(X)$, $M^{UD}(X)=\overline{UD}(X)\subseteq\overline{P}(X)$. Alors, par le Corollaire 10 en [6], $M^{UD}(X)$ est aussi engendré par un code préfixe. D'où $\overline{P}(X)=M^{UD}(X)$.

(ii)⇔(iii) est immédiat du Théorème 3 en [6].

(iii)⇔(iv)⇔(v)⇔(vi) résulte immédiatement du fait que $M^{UD}(X)$, par la Proposition 2.1, est unitaire à droite et du Corollaire 9 en [6].

(vi)⇒(i) : supposons que $M^{UD}(X)$ possède un ensemble générateur minimum distinct. Alors, étant unitaire à droite, $M^{UD}(X)$, par le Théorème 3 en [6], est engendré par un code préfixe. Par conséquent X est écrit par un code préfixe.

3. Théorème du défaut

On montre dans cette section que le théorème du défaut énoncé sous la forme du Théorème 3.2 en [1] est encore valide pour le cas des langages et codes infinitaires.

Nous avons besoin du lemme suivant :

Lemme 3.1. *Soit X un langage infinitaire écrit par un code et soit Y la base*

de $\overline{QL}(X)$. *Alors tout élément de* Y *est initiale et terminale d'au moins un mot dans* X *; c'est a dire que l'on a ;*

$$Y \subseteqq X(Y^*)^{-1} \cap (Y^*_{fin})^{-1} X.$$

Preuve. Démontrons $Y \subseteqq (Y^*_{fin})^{-1}X$. Supposons l'inclusion fausse, et prenons un $y \in Y - (Y^*_{fin})^{-1}X$. Posons

$$Z = \begin{cases} y^*(Y-y) & \text{si} \quad |y| < \omega, \\ Y-y & \text{si} \quad |y| = \omega. \end{cases}$$

Il est facile de vérifier $Z^+ = Y^*_{fin}(Y-y)$. D'où $X \subseteqq Z^* \subsetneqq Y^*$. Montrons que Z est un code. En effet, chaque élément α de Z^+ possède une factorisation unique en éléments de Y :

$$\alpha = y_1 \ldots y_n \quad \text{avec} \quad y_1 \ldots y_n \in Y^{(n)} \quad \text{et} \quad y_n \neq y.$$

Par conséquent, selon que $|y| < \omega$ ou $|y| = \omega$, α se présente uniquement sous la forme

$$\alpha = y^{p_1} y'_1 y^{p_2} y'_2 \ldots y^{p_r} y'_r \quad \text{avec} \quad y'_1 y'_2 \ldots y'_r \in (Y-y)^{(r)}, \quad p_i \geqq 0$$

$(i = 1, \ldots, r)$ où

$$\alpha = y_1 \ldots y_n \quad \text{avec} \quad y_1 \ldots y_n \in (Y-y)^{(n)},$$

c'est à dire α se factorise uniquement en éléments de Z. Donc Z^* est aussi un sous-monoïde quasi-libre contenant X, contrairement à la minimalité de $\overline{QL}(X) = Y^*$. L'inclusion $Y \subseteqq X(Y^*)^{-1}$ est démontrée d'une façon similaire en posant $Z = (Y-y)y^*$.

Théorème 3.2 (Théorème du défaut). *Soit* X *un langage infinitaire écrit par un code et soit* Y *la base de* $\overline{QL}(X)$. *Si* X *n'est pas un code, alors*

$$Card\,(Y) \leqq Card\,(X) - 1.$$

Preuve. Traitons tout d'abord le cas ou $\varepsilon \notin X$. Soit $\alpha : X \to Y$ l'application définie par :

$$\alpha(x) = y \quad \text{si} \quad x \in yY^*.$$

Elle est partout définie parce que $X \subseteqq Y^*$, et elle est univoque car Y est un code. Le Lemme 3.1 dit alors que α est surjective. X n'etant pas un code, il existe donc $x_1 \ldots x_n \in X^{(n)}$, $x'_1 \ldots x'_m \in X^{(m)}$ vérifiant

$$x_1 \ldots x_n = x'_1 \ldots x'_m, \quad x_1 \neq x'_1.$$

On en tire $\alpha(x_1) = \alpha(x'_1)$, donc α n'est pas injective, ce qui prouve l'inégalite annoncée.

Si $\varepsilon \in X$, on pose $X' = X - \varepsilon$. Clairement $\overline{QL}(X) = \overline{QL}(X')$. Si X' est un code, alors $X' = Y$ et on a $Card\ Y = Card\ X' = Card\ X - 1$; si X' n'est pas un code, alors par dessus $Card\ Y \leqq Card\ X' - 1 < Card\ X - 1$.

Corollaire 3.3. *Soit* $X = \{x_1, x_2\}$ *un langage infinitaire composé de deux mots. Alors* X *n'est pas un code ssi ou tous les deux mots de* X *sont puissances d'un même mot :*

$$x_1 = y^p, \quad x_2 = y^q \quad (p, q \geqq 0),$$

ou l'un déux est puissance ω *de l'autre :*

$$x_1 = x_2^{\omega} \quad \text{où} \quad x_2 = x_1^{\omega}.$$

Preuve ($\Leftarrow$) est évidente. Démontrons ($\Rightarrow$). Supposons que X ne soit pas un code. Trois cas sont possibles :

a) Au mois l'un des deux mots de X, disons x_1, est ε. Alors en prenant pour y le mot x_2 on a

$$x_1 = y^0, \quad x_2 = y.$$

b) x_1 et x_2 sont des mots finis non-vides. Alors X est écrit par un code, donc la base Y de $\overline{QL}(X)$, par le Théorème 3.2, se compose d'un seul mot, $Y=\{y\}$. Nous avons donc

$$x_1 = y^p, \quad x_2 = y^q$$

pour certains $p, q>0$.

c) L'un des deux mot de X, disons x_2, est un mot fini non-vide et l'autre est un mot infini. Alors $x_1=x_2^n x_1$ pour un $n>0$, d'où $x_1=x_2^\omega$.

4. Cas des langages infinitaires reconnaissable

Une partie X de A^∞ est *reconnaissable* s'il existe un morphisme $\varphi: A^\infty \to F$ de A^∞ sur un monoïde fini F qui sature X :

$$\varphi^{-1}\varphi(X) = X,$$

ou, d'une façon équivalente, s'il existe une congruence d'index fini θ de A^∞ qui sature X: X est union de classes de θ.

Le but de cette section est de montrer que si une partie infinitaire reconnaissable est écrite par un code, elle est aussi écrite par un code reconnaissable. Plus précisement nous établissons le résultat suivant qui est généralisation du Théorème 6.1 en [1] :

Théorème 4.1. *Soit X une partie de A^∞ qui est écrite par un code et soit Y la base de $\overline{QL}(X)$. Alors, si X est reconnaissable, Y l'est encore.*

La démonstration du Théorème 4.1 repose sur certaines propositions.

Proposition 4.2. *Si M est un sous-monoïde de A^∞ qui possède un ensemble générateur minimum distinct Y, et en particulier, si M est un sous-monoïde quasi-libre avec la base Y, alors M est reconnaissable ssi Y l'est encore.*

Preuve. Rappelons que la famille de parties reconnaissables de A^∞ est fermée par les opérations booléennes, par le produit et l'étoile (cf [2]). Donc M est reconnaissable si Y l'est. Pour démontrer la réciproque nous utilisons le Théorème 3 en [5] d'après lequel $Y=(M-\varepsilon)-(M_{fin}-\varepsilon)(M-\varepsilon)$. Il est facile de vérifier que pour toute X de A^∞, X_{fin} est reconnaissable si X l'est. Enfin $\{\varepsilon\}$ est clairement reconnaissable. Donc Y est reconnaissable si M l'est.

On associe maintenant à chaque partie Z de A^∞ deux parties U_Z et V_Z ainsi définies :

$$U_0 = V_0 = \{\varepsilon\}; \quad U_{j+1} = U_j^{-1}Z \cup Z^{-1}U_j; \quad V_{j+1} = ZV_j^{-1} \cup V_j Z^{-1} \quad j \geqq 0$$

$$U_Z = \bigcup_{j\geqq 0} U_j, \quad V_Z = \bigcup_{j\geqq 0} V_j.$$

Une congruence θ de A est *standard* si

$$\forall \alpha \in A^{\infty} \quad (\alpha \theta \varepsilon \rightarrow \alpha = \varepsilon).$$

À chaque congruence τ de A^{∞} on associe une congruence standard notée $\bar{\tau}$ définie par

$$\bar{\tau} = \tau \cap \mu,$$

où μ est la congruence dont les classes sont $\{\varepsilon\}$ et $A^{\infty} - \{\varepsilon\}$.

Proposition 4.3. *Soient Z une partie de A^{∞} et θ une congruence standard de A^{∞}. Si Z est saturée par θ, alors U_Z et V_Z sont également saturées par θ.*

Preuve. Montrons tout d'abord que pour toutes parties X et Y de A^{∞}, si X est saturée par θ, alors $Y^{-1}X$ et XY^{-1} le sont encore. En effet, soient $\alpha \theta \beta$ et $\alpha \in Y^{-1}X$. Montrons $\beta \in Y^{-1}X$. D'après la définition, il existe $\gamma \in Y$ tel que

$$\gamma\alpha \in X \quad \text{et} \quad |\gamma| = \omega \rightarrow \alpha = \varepsilon.$$

Comme X est saturée par θ et θ est standard, il en résulte

$$\gamma\beta \in X \quad \text{et} \quad |\gamma| = \omega \rightarrow \beta = \varepsilon,$$

ce qui signifie $\beta \in Y^{-1}X$. Donc $Y^{-1}X$ est saturé par θ. Pour XY^{-1} le raisonnement est similaire.

Évidemment U_0, V_0 sont saturés par θ. Par récurrence,

$$U_{j+1} = U_j^{-1}Z \cup Z^{-1}U_j \quad \text{et} \quad V_{j+1} = ZV_j^{-1} \cup V_j Z^{-1} \quad (j \geqq 0)$$

sont saturés par θ. Par conséquent

$$U_Z = \bigcup_{j \geqq 0} U_j \quad \text{et} \quad V_Z = \bigcup_{j \geqq 0} V_j$$

sont saturés par θ.

La proposition suivante dont la démonstration fait appel à plusieurs lemmes sera démontrée dans la section prochaine

Proposition. 4.4 *Si Z est une partie de A^{∞} qui contient le mot vide ε, alors*

$$[LB(Z^*)]^* = (U_Z \cap V_Z)^*.$$

Demonstration du Théorème 4.1 : Notons τ_X la congruence syntactique de X. Construisons une suite X_i de parties de A^{∞} comme suit :

$$\begin{cases} X_0 = X \cup \{\varepsilon\} \\ X_{i+1} = U_{X_i} \cap V_{X_i}, \quad i \geqq 0. \end{cases}$$

Alors, par la Proposition 4.4, $X_i^* = M_i^{LB}(X)$.

En vertu de la Proposition 4.3, chaque X_i est union de classes de la congruence standard $\bar{\tau}_X$. Si X est reconnaissable, alors $\bar{\tau}_X$ est d'index fini. Par conséquent le nombre des parties X_i différentes est fini. Donc il existe un n tel que $M^{LB}(X) = M_n^{LB}(X) = X_n^*$. D'où, par le Théorème 2.2, $Y^* = X_n^*$. En vertu de la Proposition 4.2, il en résulte que Y est reconnaissable.

5. Démonstration de la Proposition 4.4

Nous donnons d'abord quelques règles de calcul qui seront utilisées constamment dans la suite :

Lemme 5.1. *Soient R, S, T des parties de A^∞, et soit M un sous-monoïde de A^∞. Alors*

$$R \subseteq S \Rightarrow T^{-1}R \subseteq T^{-1}S, \quad R^{-1}T \subseteq S^{-1}T \tag{1}$$

$$R(ST)^{-1} = (RT^{-1})S^{-1} \tag{2}$$

$$(RS)^{-1}T = S^{-1}(R^{-1})T \quad \text{si} \quad \varepsilon \in S \tag{3}$$

$$(R^{-1}S)T^{-1} = R^{-1}(ST^{-1}) \quad \text{si} \quad \varepsilon \in T \tag{4}$$

$$(M^{-1}M)^{-1}M = M^{-1}M = (M^{-1}M)M \tag{5}$$

$$RM \cap SM^{-1} \subseteq (R \cap SM^{-1})M; \quad MR \cap M^{-1}S \subseteq M(R \cap M^{-1}S) \tag{6}$$

Preuve. Prouvons par exemple la formule (3). Soit $\alpha \in (RS)^{-1}T$. Ceci, par la définition, équivant à

$$\exists \beta \in R \exists \gamma \in S\colon\ (\beta\gamma)\alpha \in T \ \&\ (|\beta\gamma| = \omega \to \alpha = \varepsilon).$$

Deux cas sont possibles:

a) $|\beta| < \omega$. Alors la formule dernière implique

$$\exists \gamma \in S\colon\ (\exists \beta \in R\colon\ \beta(\gamma\alpha) \in T) \ \&\ (|\gamma| = \omega \to \alpha = \varepsilon)$$

ce qui équivant à

$$\exists \gamma \in R\colon\ (\gamma\alpha \in R^{-1}T) \ \&\ (|\gamma| = \omega \to \alpha = \varepsilon)$$

ce qui signifie que $\alpha \in S^{-1}(R^{-1}T)$.

b) $|\beta| = \omega$. Alors la formule indiquée plus haut implique $\alpha \in R^{-1}T$. Puisque $\varepsilon \in S$, ceci signifie que $\alpha \in S^{-1}(R^{-1}T)$.

Réciproquement, soit $\alpha \in S^{-1}(R^{-1}T)$. Il est facile de vérifier

$$\begin{aligned}
\alpha \in S^{-1}(R^{-1}T) &\Leftrightarrow \exists \beta \in S\colon\ (\beta\alpha \in R^{-1}T) \ \&\ (|\beta| = \omega \to \alpha = \varepsilon) \\
&\Leftrightarrow \exists \beta \in S\colon\ (\exists \gamma \in R\colon\ \gamma(\beta\alpha) \in T) \ \&\ (|\gamma| = \omega \to \beta\alpha = \varepsilon) \\
&\quad \&\ (|\beta| = \omega \to \alpha = \varepsilon) \\
&\Rightarrow \exists \beta \in S \exists \gamma \in R\colon\ (\gamma\beta)\alpha \in T \ \&\ (|\gamma\beta| = \omega \to \alpha = \varepsilon) \\
&\Leftrightarrow \alpha \in (RS)^{-1}T.
\end{aligned}$$

Posons, pour alléger l'écriture,

$$M = Z^*;\ \overline{U}_j = U_0 \cup U_1 \cup \ldots \cup U_j;\ \overline{V}_j = V_0 \cup V_1 \cup \ldots \cup V_j.$$

Lemme 5.2. *Pour tout $j \geqq 0$,*

$$U_{j+1} \subseteq (\overline{U}_j M)^{-1}Z \quad \text{et} \quad V_{j+1} \subseteq Z(M\overline{V}_j)^{-1}.$$

Preuve. Par récurrence sur j. Pour $j=0$, l'inclusion est vraie puisque $U_1 = Z \subseteq M^{-1}Z = (\overline{U}_0 M)^{-1}Z$. Si $j>0$, alors, par l'hypothèse d'induction et les formules

(1), (3), on a

$$U_{j+1} = U_j^{-1}Z \cup Z^{-1}U_j \subseteqq U_j^{-1}Z \cup Z^{-1}((\overline{U}_{j-1}M)^{-1}Z) =$$
$$= U_j^{-1}Z \cup (\overline{U}_{j-1}MZ)^{-1}Z = (U_j \cup \overline{U}_{j-1}MZ)^{-1}Z \subseteqq$$
$$\subseteqq (U_jM \cup \overline{U}_{j-1}M)^{-1}Z = (\overline{U}_jM)^{-1}Z,$$

ce qui prouve la première inclusion. La deuxième se démontre de la même façon en utilisant (2) au lieu de (3).

Lemme 5.3. *On a*

$$U_Z = M^{-1}U_Z, \quad V_Z = V_ZM^{-1},$$
$$M^{-1}M = U_ZM, \quad MM^{-1} = MV_Z.$$

Preuve. Par définition, $Z^{-1}U_j \subseteqq U_{j+1}$ pour $j \geqq 0$, donc $Z^{-1}U_Z \subseteqq U_Z$. De la même manière, $U_j^{-1}Z \subseteqq U_{j+1}$ $(j \geqq 0)$ implique

$$U_Z^{-1}Z \subseteqq U_Z. \tag{7}$$

De l'inclusion $Z^{-1}U_Z \subseteqq U_Z$, on obtient par récurrence sur n en utilisant (3) et (1).

$$(Z^{(n+1)})^{-1}U_Z = (Z^{(n)})^{-1}(Z_{fin}^{-1}U_Z) \subseteqq (Z^{(n)})^{-1}U_Z \subseteqq U_Z \quad n \geqq 0,$$

d'où, puisque $M = Z^* = \bigcup_{n \geqq 0} Z^{(n)}$, on a $M^{-1}U_Z \subseteqq U_Z^*$. L'inclusion $U_Z \subseteqq M^{-1}U_Z$ résulte de ce que $\varepsilon \in M$. La deuxième se démontre de la même façon.

Pour établir la troisième formule, nous vérifions par récurrence les inclusions

$$U_j \subseteqq M^{-1}M, \quad j \geqq 0.$$

Le cas $j=0$ étant evident. Par le Lemme 5.2 et par l'hypothèse d'induction, on obtient

$$U_{j+1} \subseteqq (\overline{U}_jM)^{-1}Z \subseteqq ((M^{-1}M)M)^{-1}Z.$$

Par (5) et (1)

$$U_{j+1} \subseteqq ((M^{-1}M)M)^{-1}Z = (M^{-1}M)^{-1}Z \subseteqq (M^{-1}M)^{-1}M = M^{-1}M.$$

D'où, $U_Z \subseteqq M^{-1}M$. Par conséquent

$$U_ZM \subseteqq (M^{-1}M)M = M^{-1}M.$$

Pour démontrer l'inclusion inverse $M^{-1}M \subseteqq U_ZM$ nous vérifions tout d'abord l'inclusion $(M^{-1}M)_{fin} \subseteqq U_ZM$. Supposons en effet $f \in (M^{-1}M)_{fin}$. Alors il existe $m \in M$ tel que

$$mf = m' \in M \quad \text{et} \quad (|m| = \omega) \to (f = \varepsilon).$$

Puisque $\varepsilon \in U_ZM$ on peut supposer $f \neq \varepsilon$ ce qui implique $|m| < \omega$.

Nous vérifions par récurrence sur $|mm'|$ que $f \in U_ZM$. Si $|mm'| = 0$, $f \in M \subseteqq U_ZM$; si $|mm'| \neq 0$, alors, puisque $M = Z^*$, il existe h, h' tels que

$$m = m_1h, \quad f = h'm_2 \quad \text{avec} \quad m_1, m_2 \in M, \quad hh' \in Z.$$

Alors $h \in (M^{-1}M)_{fin}$ avec $|mm_1| < |mm'|$. Par l'hypothèse d'induction, $h \in U_ZM$.

En utilisant (3) et (7) on obtient

$$h' \in (U_Z M)^{-1} Z = M^{-1}(U_Z^{-1} Z) \subseteq M^{-1} U_Z = U_Z.$$

Par conséquent $f = h' m \in U_Z M$.

Nous vérifions maintenant $(M^{-1}M)_{inf} \subseteq U_Z M$. Soit $u \in (M^{-1}M)_{inf}$. Il existe alors $m \in M_{fin}$ tel que $mu \in M$. D'une façon similaire au dessus, il existe h, h' tels que

$$m = m_1 h, \quad u = h' m_2 \quad \text{avec} \quad m_1, m_2 \in M, \quad hh' \in Z.$$

Alors $h \in (M^{-1}M)_{fin}$ ce qui implique $h \in U_Z M$. D'une façon similaire au dessus on a $h' \in U_Z$, par conséquent $u = h' m_2 \in U_Z M$.

Démonstration de la Proposition 4.4.

D'après le Lemme 5.3,

$$U_Z \cap V_Z \subseteq U_Z M \cap M V_Z = M^{-1}M \cap MM^{-1} = LB(M) = LB(Z^*).$$

Donc $(U_Z \cap V_Z)^* \subseteq [LB(Z^*)]^*$.

Réciproquement,

$$LB(M) = M^{-1}M \cap MM^{-1} = U_Z M \cap MM^{-1} \subseteq (U_Z \cap MM^{-1}) M.$$

Puis, par les formules deuxième et quatrième du Lemme 5.3 et par (6), on a

$$U_Z \cap MM^{-1} = U_Z \cap M V_Z = M^{-1} U_Z \cap M V_Z \subseteq M(M^{-1} U_Z \cap V_Z) = M(U_Z \cap V_Z).$$

D'où $LB(M) \subseteq M(U_Z \cap V_Z)M$. Comme $Z \subseteq U_Z \cap V_Z$, $M \subseteq (U_Z \cap V_Z)^*$, par $LB(M) \subseteq (U_Z \cap V_Z)^*$. Donc $[LB(Z^*)]^* \subseteq (U_Z \cap V_Z)^*$.

Abstract

We give necessary and sufficient conditions for an infinitary language to be written by an infinitary code. It is shown that the "Theorem of defect" remains valid for the case of infinitary languages and codes, and that if a recognizable infinitary language is written by an infinitary code, it can be also written by a recognizable infinitary code.

INSTITUTE OF MATHEMATICS
P/0 BOX 631 BO HO
HANOI, VIETNAM

References

[1] J. Berstel, D. Perrin, J. F. Perrot, A. Restivo, Sur le théorème du défaut, Journal of Algebra, 60 (1979), 169—180.

[2] S. Eilenberg, "Automata Languages and Machines", Vol. A, Academic Press, New York—London, 1974.

[3] Do Long Van, Codes avec des mots infinis, R.A.I.R.O. Informatique théorique, 16 (1982), 371—386.

[4] Do Long Van, Sous-monoïdes et codes avec des mots infinis, Semigroup Forum, 26 (1983), 75—87.

[5] Do Long Van, Sur les ensembles générateurs minimums des sous-monoïdes de A^{∞}, Preprint Series, Hanoi, 1984, №21.

[6] Do Long Van, Caractérisations combinatoires des sous-monoïdes engendrés par un code infinitaire, Preprint Series, Hanoi—1984, № 6.

(Received Febr. 3, 1985)