Numération: Autour des automates non finis et des langages non réguliers

Marion Le Gonidec

26 octobre 2009

- En bases entières :
  - Réels dont les parties fractionnaires forment mots automatiques et mots substitutifs,
  - Ensembles d'entiers reconnaissables,
- En bases non entières et en numération abstraite :
  - β-numération,
  - numérations basées sur des substitutions,
  - numérations basées sur des langages réguliers.

- Problème de Hartmanis et Stearns :
  Peut-on calculer les décimales d'un nombre algébrique en temps réel ?

Les limites...

- Problème de Hartmanis et Stearns :
  Peut-on calculer les décimales d'un nombre algébrique en temps réel ?
  La représentation k-adique d'un nombre algébrique n'est pas un mot automatique.

Les limites...

- Problème de Hartmanis et Stearns :
  Peut-on calculer les décimales d'un nombre algébrique en temps réel ?
  La représentation k-adique d'un nombre algébrique n'est pas un mot automatique.
- Existe-t-il une machine temps réel qui permet de décider si un nombre est premier à partir de sa représentation en base k ?

Les limites...

- Problème de Hartmanis et Stearns :
  Peut-on calculer les décimales d'un nombre algébrique en temps réel ?
  La représentation k-adique d'un nombre algébrique n'est pas un mot automatique.
- Existe-t-il une machine temps réel qui permet de décider si un nombre est premier à partir de sa représentation en base k ?
  L'ensemble des nombres premiers n'est pas reconnaissable ni par un automate fini ni par un automate à pile.

Les limites...

- Problème de Hartmanis et Stearns :
  Peut-on calculer les décimales d'un nombre algébrique en temps réel ?
  La représentation k-adique d'un nombre algébrique n'est pas un mot automatique.
- Existe-t-il une machine temps réel qui permet de décider si un nombre est premier à partir de sa représentation en base k ?
  L'ensemble des nombres premiers n'est pas reconnaissable ni par un automate fini ni par un automate à pile.
- Quels langages permettent de fournir un bon système de représentation des réels ?

Les limites...

- Problème de Hartmanis et Stearns :
  Peut-on calculer les décimales d'un nombre algébrique en temps réel ?
  La représentation k-adique d'un nombre algébrique n'est pas un mot automatique.
- Existe-t-il une machine temps réel qui permet de décider si un nombre est premier à partir de sa représentation en base k ?
  L'ensemble des nombres premiers n'est pas reconnaissable ni par un automate fini ni par un automate à pile.
- Quels langages permettent de fournir un bon système de représentation des réels ?
  Le langage sur lequel est basé la numération en base $\frac{3}{2}$ n'est ni régulier, ni algébrique.

- Extensions de la notion de mot automatique,
- Extensions de la notion d'ensemble reconnaissable,
- Systèmes de numération abstraits basés sur des langages quelconques.

Soit $\mathbf{m} = m_0 m_1 m_2 \dots$ un mot infini sur un alphabet fini A.

Plusieurs caractérisations :

- Les langages $\{[n]_k \mid m_n = a\}$ sont réguliers,
- Il existe un automate fini avec fonction de sortie qui, pour tout n renvoie m_n après lecture de $[n_k]$.
- $\mathbf{m}$ est l'image morphique du point fixe d'une substitution de longueur constante sur un alphabet fini.
- Le k-noyau de $(\{(m_{k^a n+b})_{n\geq 0} \mid a \geq 1, 0 \leq b < a\})$ de $\mathbf{m}$ est fini.

Quelques propriétés

- Stabilité du caractère automatique par extraction de sous suite et opération de décalage,
- Fonction de complexité de l'ordre de 1 ou n,
- Dépendance à la base : Si $\mathbf{m}$ est k- et l-automatique pour k et l multiplicativement indépendants, alors $\mathbf{m}$ est ultimement périodique.

Mots infinis algébriques

Un mot infini $\mathbf{m} = m_0 m_1 m_2 \dots$ sur un alphabet fini A est k-algébrique si les langages $\{[n]_k \mid m_n = a\}$ sont algébriques.

Un mot infini $\mathbf{m} = m_0 m_1 m_2 \dots$ sur un alphabet fini A est k-algébrique si les langages $\{[n]_k \mid m_n = a\}$ sont algébriques.

- Stabilité du caractère algébrique par extraction de sous-suite et opération de décalage,

Mots infinis algébriques

Un mot infini $\mathbf{m} = m_0 m_1 m_2 \ldots$ sur un alphabet fini A est k-algébrique si les langages $\{[n]_k \mid m_n = a\}$ sont algébriques.

- Stabilité du caractère algébrique par extraction de sous-suite et opération de décalage,
- Fonction de complexité :
  Si tous ces langages peuvent être reconnus par un même automate déterministe temps réel, alors la complexité de $\mathbf{a}$ est au plus polynomiale.

Mots infinis algébriques

Un mot infini $\mathbf{m} = m_0 m_1 m_2 \ldots$ sur un alphabet fini A est k-algébrique si les langages $\{[n]_k \mid m_n = a\}$ sont algébriques.

- Stabilité du caractère algébrique par extraction de sous-suite et opération de décalage,
- Fonction de complexité :
  Si tous ces langages peuvent être reconnus par un même automate déterministe temps réel, alors la complexité de **a** est au plus polynomiale.
  Si s est le nombre de sommets de pile de l'automate :
  - $p_{\mathbf{a}}(n) = O(n \log_k^2 n)$ si $s = 1$,
  - $p_{\mathbf{a}}(n) = O(n^{1+4\log_k s})$ si $s \geq 2$.

Ensembles algébriquement reconnaissables

Un ensemble d'entiers E est algébriquement reconnaissable si $\{[n]_k \mid n \in E\}$ est un langage algébrique. Soit $\mathrm{Alg}(k)$ est la famille formée par ces ensembles.

- Si E est algébriquement reconnaissable, alors $E(a, b) = \{an + b \mid n \in E\}$ l'est aussi,
- $\mathrm{Alg}(k) = \mathrm{Alg}(l)$ si et seulement si $k^a = l^b$ pour deux entiers a et b,
- Si elle existe, la densité asymptotique d'un ensemble algébriquement reconnaissable est un nombre algébrique lorsque la grammaire qui l'engendre est non-ambigüe.
- Quelque soit la base choisie, ni l'ensemble des nombres premiers, ni celui des carrés parfaits ne sont algébriquement reconnaissables.

Questions ouvertes...

Mots infinis k-algébriques :

- Déterminer quelle sous-famille de $\mathrm{Alg}(k)$ forme les supports des lettres des mots k-algébriques,
- Dépendance à la base,
- Structure du k-noyau,
- Optimalité de la majoration de la complexité,
- Etude des systèmes dynamiques associés à ces mots.

Mots k^∞-automatiques

Un mot infini $\mathbf{m} = m_0 m_1 m_2 \dots$ sur un alphabet fini A est k^∞-automatique si $\mathbf{m}$ est l'image morphique du point fixe d'une substitution de longueur constante sur un alphabet dénombrable D.

Les morphismes admissibles $D \longrightarrow A$ sont constants sauf sur un nombre fini de lettres de D.

Mots k^∞-automatiques

Un mot infini $\mathbf{m} = m_0 m_1 m_2 \dots$ sur un alphabet fini A est k^∞-automatique si $\mathbf{m}$ est l'image morphique du point fixe d'une substitution de longueur constante sur un alphabet dénombrable D.

Les morphismes admissibles $D \to A$ sont constants sauf sur un nombre fini de lettres de D.

- Les supports des lettres forment des langages dont des automates minimaux sont dénombrables avec un nombre fini d'états terminaux, Certains peuvent être algébriques ou sensibles au contexte.
- Dans le cas où ces automates sont de degré entrant borné, alors la complexité de $\mathbf{m}$ est au plus polynomiale,
- $\mathbf{m}$ ne peut pas être la représentation k-adique d'un nombre algébrique,

Ensembles k^∞-reconnaissables

$\mathcal{A} = (D, \phi, d_0, F)$ un k-automate déterministe, de degré entrant borné, avec D dénombrable et F fini.
E est k^∞-reconnaissable s'il existe un k-automate $\mathcal{A} = (D, \phi, d_0, F)$ déterministe et de degré entrant borné, avec D dénombrable et F finipour lequel :

$$n \in E \iff \phi(d_0, [n]_k) \in F.$$

- Il existe deux manières de reconnaitre des ensembles d'entiers avec des automates dénombrable,
- Les deux familles d'ensembles d'entiers ont des propriétés de stabilité différentes,

Ensembles k^∞-reconnaissables

$\mathcal{A} = (D, \phi, d_0, F)$ un k-automate déterministe, de degré entrant borné, avec D dénombrable et F fini.
E est k^∞-reconnaissable s'il existe un k-automate $\mathcal{A} = (D, \phi, d_0, F)$ déterministe et de degré entrant borné, avec D dénombrable et F finipour lequel :

$$n \in E \iff \phi(d_0, [n]_k) \in F.$$

- Il existe deux manières de reconnaitre des ensembles d'entiers avec des automates dénombrable,
- Les deux familles d'ensembles d'entiers ont des propriétés de stabilité différentes,
- L'ensemble des nombres premiers n'est reconnaissable dans aucune base quelque soit le sens de lecture.
- L'ensemble des factorielles est reconnaissable dans toute base, uniquement en lecture droite$\longrightarrow$gauche.

Questions ouvertes...

- Trouver un ensemble d'entier reconnaissable dans toute base en lecture gauche$\longrightarrow$droite,
- Situer ces automates par rapport à la hiérarchie de Chomsky,

Systèmes de numération basés sur des langages quelconques

Problématique :

On se donne un système de numération abstrait $S = (L, \Sigma, <)$. Les réels sont représentés par les mots infinis limites de mots du langage L sur l'alphabet ordonné $(\Sigma, <)$.

Quelles conditions doit vérifier L pour fournir un bon système de représentation des réels ?

- Représentation d'un intervalle,
- Équivalence entre "convergence en mot" et convergence numérique.
- Respect de l'ordre lexicographique induit sur L par $(\Sigma, <)$.

Si w est limite d'une suite de mots $(w^{(n)})_{n\geq 0}$ de mots de L, on pose :

$$\mathrm{Real(w)} = \lim_{n\to\infty} \frac{\#\{u \in L \mid u \leq w^{(n)}\}}{\#(L \cap \Sigma^{\leq |w^{(n)}|})}$$

Premières remarques :

- Le langage L doit grandir au moins exponentiellement.
- Cette limite peut ne pas exister pour certains langages non préfixiels.

Sous les conditions suivantes :

- L'ensemble des mots limites $\mathrm{Adh}(L)$ est non-dénombrable,
- Pour tout mot v de Σ^*, il existe un réel $r_v \geq 0$ tel que :

$$\lim_{n\to\infty} \frac{\#\{u \in \Sigma^{n-|v|} \mid vu \in L\}}{\#(L \cap \Sigma^{\leq n})} = r_v,$$

- pour tout mot w de $\mathrm{Adh}(L)$, $\lim_{n\to\infty} r_{w[0,n-1]} = 0$,

Sous les conditions suivantes :

- L'ensemble des mots limites $\mathrm{Adh}(L)$ est non-dénombrable,
- Pour tout mot v de Σ^*, il existe un réel $r_v \geq 0$ tel que :

$$\lim_{n\to\infty} \frac{\#\{u \in \Sigma^{n-|v|} \mid vu \in L\}}{\#(L \cap \Sigma^{\leq n})} = r_v,$$

- pour tout mot w de $\mathrm{Adh}(L)$, $\lim_{n\to\infty} r_{w[0,n-1]} = 0$,

La convergence en mot est équivalente à la convergence numérique :

$$\mathrm{Real(w)} = \lim_{n\to\infty} \frac{\#\{u \in L \mid u \leq w^{(n)}\}}{\#(L \cap \Sigma^{\leq |w^{(n)}|})}$$

existe et est indépendant de la suite $\left(w^{(n)}\right)_{n\geq 0}$ de mots de L choisie.

Sous les conditions suivantes :

- L'ensemble des mots limites $\mathrm{Adh}(L)$ est non-dénombrable,
- Pour tout mot v de Σ^*, il existe un réel $r_v \geq 0$ tel que :

$$\lim_{n\to\infty} \frac{\#\{u \in \Sigma^{n-|v|} \mid vu \in L\}}{\#(L \cap \Sigma^{\leq n})} = r_v,$$

- pour tout mot w de $\mathrm{Adh}(L)$, $\lim_{n\to\infty} r_{w[0,n-1]} = 0$,

La convergence en mot est équivalente à la convergence numérique :

$$\mathrm{Real(w)} = \lim_{n\to\infty} \frac{\#\{u \in L \mid u \leq w^{(n)}\}}{\#(L \cap \Sigma^{\leq |w^{(n)}|})}$$

existe et est indépendant de la suite $(w^{(n)})_{n\geq 0}$ de mots de L choisie. L'application $w \in \mathrm{Adh}(L) \mapsto \alpha_w$ est uniformément continue et décrit un intervalle du type $[s_0, 1]$.

Ces conditions sont applicables :

- aux les clotures préfixielles :
  - des langages rationnels,
  - des langages algébriques usuels,
- au langage de la numération en base $\frac{3}{2}$.

Ces conditions sont applicables :

- aux les clotures préfixielles :
  - des langages rationnels,
  - des langages algébriques usuels,
- au langage de la numération en base $\frac{3}{2}$.

Une question ouverte :

- Condition necessaire et suffisante de non-dénombrabilité de $\mathrm{Adh}(L)$.

Merci.