Informatique théorique III
(Automates, langages & calculabilité)

Formulaire du Cours 2004-2005
Partie 14
EPFL – I&C

Uwe Nestmann
Sébastien Briais
Daniel C. Bünzli

14 février 2005

Bibliographie

[HMU03] John Hopcroft, Rajeev Motwani, and Jeffrey Ullman. *Introduction to Automata Theory, Languages and Computation*. Pearson Education International, 2003. ISBN 0321210298.

Table des matières

$Id: notes-14.tex,v 1.5 2005/01/29 19:17:44 uwe Exp $

semaine 13

Voir [HMU03] §9.4.

4.10 Problèmes pratiques ... et indécidables

4.10.1 Définition (Jeu de Domino) Soit Σ un alphabet.
Une *jeu de domino* (A, B) de taille $k \in \mathbb{N}$
est une paire d'applications $A, B : [1, k] \to \Sigma^*$.

4.10.2 Notation Un jeu de domino (A, B) de taille $k \in \mathbb{N}$ définit une liste finie de paires que nous notons :

$$\left[\frac{A(1)}{B(1)}\right], \left[\frac{A(2)}{B(2)}\right], \dots, \left[\frac{A(k)}{B(k)}\right]$$

4.10.3 Définition Soit Σ un alphabet et $f : [1, k] \to \Sigma^*$.
Soit $I = i_1 \cdot i_2 \dots \cdot i_m \in [1, k]^+$ un mot non vide. On définit :

$$f\langle I\rangle \triangleq f(i_1) \cdot f(i_2) \cdot \dots \cdot f(i_m)$$

4.10.4 Définition (PCP) Soit Σ un alphabet.
Soit (A, B) un jeu de domino de taille $k \in \mathbb{N}$ et $I \in [1, k]^+$.

1. I est une *solution* de (A, B) si $A\langle I\rangle = B\langle I\rangle$.
2. I est une *solution partielle* de (A, B) si
   (a) $A\langle I\rangle$ est préfixe de $B\langle I\rangle$, ou
   (b) $B\langle I\rangle$ est préfixe de $A\langle I\rangle$.

Le *problème de correspondance de Post* (PCP) consiste à déterminer si un jeu de domino (A, B) donné a une solution ou non. Un jeu (A, B) est une *instance du PCP*.

4.10.5 Lemme Soit Σ un alphabet, (A, B) une instance de taille k du PCP et $I \in [1, k]^+$ une solution de (A, B).
Alors tout préfixe de I est une solution partielle de (A, B).

4.10.6 Définition (PCPM) Soit Σ un alphabet, (A, B) un jeu de domino de taille k et $I \in [1, k]^*$. I est *solution* de (A, B) selon le *PCP modifié* (PCPM) si :

$$A\langle 1 \cdot I\rangle = B\langle 1 \cdot I\rangle$$

(A, B) est une *instance du PCPM*.

4.10.7 Définition Toute instance (A, B) du PCP(M) peut être encodé en binaire, et en utilisant un symbole séparateur de la façon suivante :
– prendre la taille de l'alphabet ;
– déterminer le nombre de bits nécessaire ;
– commencer par coder ce nombre en binaire ;
– puis, en binaire, toutes les paires de l'instance du PCP ;

– séparer à chaque fois la suite avec le séparateur.

Les langages PCP et PCPM sont ainsi définis (de maniére injective) par des ensembles de mots sur un alphabet fini (ici, de taille 3).

$$\begin{array}{lcl} \mathrm{PCP}^{\oplus} & \triangleq & \{\, w \in \mathrm{PCP} \mid w \text{ est soluble} \,\} \\ \mathrm{PCPM}^{\oplus} & \triangleq & \{\, w \in \mathrm{PCPM} \mid w \text{ est soluble} \,\} \end{array}$$

4.10.8 Théorème $\mathrm{PCPM}^{\oplus} \leq_{\mathrm{red}} \mathrm{PCP}^{\oplus}$.

4.10.9 Théorème $\mathrm{PCP}^{\oplus}$ est semi-décidable.

4.10.10 Définition Soit $M = (Q, \{0,1\}, \Gamma, \delta, q_0, \mathrm{B}, F)$ une MT.

Soit $c = (q, i, \mathrm{rub})$ une configuration de M.

Si $\mathrm{rub}(z) = \mathrm{B}$ pour tout $z \in \mathbb{Z}$, alors $w_c \triangleq \epsilon$.

S'il existe $z \in \mathbb{Z}$ tel que $\mathrm{rub}(z) \neq \mathrm{B}$, alors soit z_d, z_f définis par :

- $\mathrm{rub}(z_d) \neq \mathrm{B} \neq \mathrm{rub}(z_f)$, et
- $\forall j \in \mathbb{Z} \setminus [z_d, z_f]\,.\ \mathrm{rub}(z) = \mathrm{B}$.

Soit $w_c \triangleq \mathrm{rub}(z_d) \cdot \ldots \cdot \mathrm{rub}(z_f)$.

Nous représontons c aussi comme élément de $\Gamma^* Q \Gamma^+$ par :

$$\Theta(c) \triangleq \begin{cases} q \cdot \mathrm{B} & \text{si } w_c = \epsilon \\ \mathrm{rub}(z_d) \cdot \ldots \cdot \mathrm{rub}(i{-}1) \cdot q \cdot \mathrm{rub}(i) \cdot \ldots \cdot \mathrm{rub}(z_f) & \text{si } i \in [z_d, z_f] \\ q \cdot \mathrm{B}^{z_d - i} \cdot w_c & \text{si } i < z_d \\ w_c \cdot \mathrm{B}^{i-1-z_f} \cdot q \cdot \mathrm{B} & \text{si } z_f < i \end{cases}$$

Comme inverse, nous définissons pour le cas « standard » dans lequel la tête se trouve à l'intérieur du mot commençant en position 1 :

$$\Theta^{-1}(X_1 \cdot \ldots \cdot X_{i-1} \cdot q \cdot X_i \cdot \ldots \cdot X_n) \triangleq (q, i, \mathrm{rub})$$

où $X_1 \neq \mathrm{B} \neq X_n$ et tel que $\mathrm{rub}(j) \triangleq \begin{cases} X_j & \text{si } j \in [1, n] \\ \mathrm{B} & \text{sinon} \end{cases}$

4.10.11 Théorème $L_{\mathrm{U}} \leq_{\mathrm{red}} \mathrm{PCPM}^{\oplus}$.

4.10.12 Corollaire $\mathrm{PCPM}^{\oplus}$ et $\mathrm{PCP}^{\oplus}$ sont indécidable.

4.10.13 Définition

Soit (A, B) une instance de taille k du PCP sur l'alphabet Σ.

Soit, pour tout $1 \leq j \leq k$: $w_j \triangleq A(j)$ et $x_j \triangleq B(j)$.

Soit $a_1, \ldots, a_k$ des nouveaux symboles. Alors,

$$G_{AB} \triangleq \big(\{\, S, S_A, S_B \,\}, \Sigma \cup \{\, a_1, \ldots, a_k \,\}, P, S\big)$$

est la grammaire (non-contextuelle) définie par :

$$\begin{array}{lll} S & \to_G & S_A \mid S_B \\ S_A & \to_G & w_1 a_1 \mid \cdots \mid w_k a_k \mid w_1 S_A a_1 \mid \cdots \mid w_k S_A a_k \\ S_B & \to_G & x_1 a_1 \mid \cdots \mid x_k a_k \mid x_1 S_B a_1 \mid \cdots \mid x_k S_B a_k \end{array}$$

4.10.14 Théorème Soit (A, B) une instance du PCP.
(A, B) est soluble ssi G_{AB} est ambiguë.

4.10.15 Corollaire Le problème qui consiste à décider l'ambiguïté d'une grammaire quelconque est indécidable.