Théorème de Cook

Langages Formels. Calculabilité et complexité.
Olivier Carton

2011-2012

Théorème 0.1 : de Cook, 1971
Le problème Sat *est NP-complet.*

Démonstration. Soit A un problème de NP, et soit $\mathscr{M}$ une machine de Turing non-déterministe qui décide A en temps polynomial.

Pour chaque entrée w, on va construire une formule φ_w qui sera satisfiable *si et seulement si* $\mathscr{M}$ accepte w.

On note $n = |w|$. On peut supposer que chaque calcul acceptant sur w est de longueur exactement n^k (quitte à rajouter des transitions inutiles).

La machine utilise donc au plus n^k cellules de sa bande de travail, et donc les configurations sont de longueur au plus n^k : de même, on les prendra de longueur exactement n^k, quitte à rajouter des symboles blancs.

On les note dans un tableau :

Conf.	0	1	2	3	$\cdots$	n^k
$C_0 =$	q_0	w_1	w_2	w_3	$\cdots$	$\sharp$
$C_1 =$	w'_1	q_1	w_2	w_3	$\cdots$	$\sharp$
$C_2 =$	w'_1	w_2	q_2	w_3	$\cdots$	$\sharp$
$C_2 =$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\sharp$
$\vdots$						$\vdots$
$C_{n^k} =$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

On va donc coder une formule φ_w qui code l'existence d'un tel tableau.

On définit les variables $x_{i,j,a}$ pour $i, j \in [\![0, n^k]\!]$ et a symbole de $A = \Gamma \cup Q$ qui codent le fait que la variable a se trouve dans la case i, j. Il y a $|A|n^{2k+2}$ telles variables.

On décompose notre formule φ_w en quatre formules φ_0, φ_1, φ_2 et φ_3, qui vont chacune coder une propriété du tableau.

φ_0 : Cette formule code le fait que chaque case du tableau contient un et un seul symbole de A :

$$\varphi_0 = \bigwedge_{0 \leqslant i,j \leqslant n^k} \left[\left(\bigvee_{a \in A} x_{i,j,a} \right) \wedge \left(\bigwedge_{a \neq a' \in A} (\overline{x}_{i,j,a} \vee \overline{x}_{i,j,a'} \right) \right].$$

φ_1 : Cette formule code le fait que la première ligne du tableau est bien $q_0 w$:

$$\varphi_1 = \left(\bigwedge_{0 \leqslant i \leqslant n} x_{0,i,w_i} \right) \wedge \left(\bigwedge_{n+1 \leqslant i \leqslant n^k} x_{0,i,\sharp} \right).$$

φ_2 : Cette formule assure que chaque ligne est obtenue en appliquant une transition valide de $\mathscr{M}$.
Il suffit de remarquer que la valeur d'une case (i, j) ne dépend que des trois cases au-dessus $(i-1, j-1)$, $(i-1, j)$ et $(i-1, j+1)$.
Si dans ces trois cases se trouvent des symboles de bande, alors le contenu de la case (i, j) est le même qu'en $(i-1, j)$.
Si l'état de la configuration se trouve en $(i-1, j)$, alors l'état de C_i se trouve en $(i, j-1)$ ou $(i, j+1)$.
Donc, il suffit bien de regarder les "fenêtres" de taille 2×3 du tableau. L'ensemble des fenêtres possibles ne dépend que de A et des transitions de $\mathscr{M}$, et donc ne dépend pas de la taille de l'entrée n.
Le fait que chaque fenêtre du tableau corresponde bien à une transition s'écrit donc comme une conjonction pour $0 \leqslant i, j \leqslant n^k$ de disjonctions des fenêtres possibles, ce qui est polynomial en n.

φ_3 : Cette formule code le fait que $\mathscr{M}$ accepte w, *i.e* qu'au moins une des cases de la dernière ligne contient un état final :

$$\varphi_3 = \bigvee_{q \in F} \left(\bigvee_{0 \leqslant j \leqslant n^k} x_{n^k, j, q} \right).$$

□