JEAN-MARIE BOE

Un problème combinatoire de la théorie des codes

Publications du Département de Mathématiques de Lyon, 1984, fascicule 6B
« Théorie des langages et complexité des algorithmes », , p. 115-120

<http://www.numdam.org/item?id=PDML_1984___6B_A6_0>

UN PROBLEME COMBINATOIRE DE LA THEORIE DES CODES

par Jean-Marie BOE

Le but de cet article est de montrer l'équivalence entre un problème combinatoire et une conjecture de la théorie des codes à longueur variable, puis d'examiner l'algorithmique du problème combinatoire.

A^* désigne l'ensemble des mots (monoide libre) sur l'alphabet A. Pour une partie X de A^*, X^* désigne le sous-monoide engendré par X: ensemble des mots écrits avec les mots de X.

X est un code si tout produit de mots de X se factorise de manière unique en ces memes mots:

$x_1 x_2 \ldots x_n = x'_1 x'_2 \ldots x'_m$ où $x_i, x'_j \in X$ entraine n=m et $x_i = x'_i$ pour tout i.

En d'autres termes, X^*est un monoide libre sur X.

Nous ne considèrerons ici que des codes finis maximaux (pour l'inclusion).

Pour étudier les propriétés de X, nous représentons le monoide libre A par des relations sur un ensemble E = [1..n]. μ_X désignera un morphisme (voir exemple plus loin) qui associe à tout mot de A une relation sur E. La condition de codicité se traduit par une condition de non-ambiguité du produit des relations définie ci-aprés:

Un produit de deux relations r et s (matrices booléennes) est non-ambigu si $(i,j)\in rs$ entraine l'existence d'un unique k tel que $(i,k)\in r$ et $(k,j)\in s$. Le lecteur remarquera que la condition suivante est équivalente: r et s étant des matrices à coefficients entiers 0 ou 1, le produit rs est aussi à coefficients 0 ou 1.
Un monoide de relations non-ambigu est un monoide où le produit est non-ambigu.

La condition de finitude nécessite la définition suivante: Une relation sur [1..n] est 1-triangulaire si les éléments (i,j) de la relation vérifient $i<j$ ou $j=1$.

Enfin nous ne considèrerons que des monoides de relations transitifs: tout couple (i,j) est présent dans une des relations du monoide. Il vient:

Théorème 1 [BePe] : Une partie X de A^* est un code maximal fini ssi il existe une représentation μ_X de A^* par des relations sur $[1..n]$ vérifiant:
i- X^* est l'ensemble des mots qui envoient 1 sur 1:
$$X^* = \{x \in A^* \mid (1,1) \in \mu_X(x)\}$$
ii- $\mu_X(A^*)$ est un monoide de relations non-ambigu.
iii- la relation vide n'appartient pas à $\mu_X(A^*)$.
iv- $\mu_X(a)$ est 1-triangulaire pour $a \in A$.

Indications de preuve: Soit X un code maximal fini, posons:
$$S = \{(u,v) \in A^* \times A^* \mid uv \in X,\ u \neq 1 \neq v\} \cup \{(1,1)\}$$
On identifie S à E en identiant (1,1) à 1 et en numérotant les (u,v) de telle sorte que (u,v) précède (u',v') si uv=u'v' et u plus court que v.
Pour $m \in A^*$, soit $\mu_X(m) = \{((u,v),(u',v')) \in S \times S \mid um \in X^* u'$ et $mv' \in vX^*\}$
On vérifie alors que μ_X satisfait les conditions i-iv du théorème.

Réciproquement, les conditions i et ii assurent que X^* est un sous-monoide libre de A^*, la condition iii la maximalité du code et la condition iv sa finitude.

Par la suite nous désignerons par m.r.n.a. un monoide de relations non-ambigu transitif ne contenant pas la relation vide. Par ailleurs, étant donné un code maximal fini X, μ_X désignera une représentation vérifiant les propriétés du théorème.

I-BOITES ET MONOIDE DE RELATIONS NON-AMBIGU.

Etant donnée une famille L de parties d'un ensemble E, une section c de L est une partie de E qui intersecte tous les éléments de L en un point et un seul.
On désigne par boite un couple (L,C) de familles de parties de E tel que:
i- L et C recouvrent E;
ii- toute section de L appartient à C; toute section de C appartient à L.

Il est commode de représenter une boite par un tableau T dont l'élément T(i,j) est l'intersection du ième élément de L et du jème élément de C.

exemple: $E = \{1,2,3\}$ $L = \{\{1,3\},\{2\}\}$ $C = \{\{1,2\},\{3\}\}$
ce qui se représente par la boite suivante, indicée par LxC:

	{1,2}	{2,3}
{1,3}	1	3
{2}	2	2

Un tableau d'éléments de E est une boite s'il vérifie:
i- si un élément est présent en (i,j) et (k,l) alors il apparait en (i,l) (condition de section);
ii- toutes lignes et toutes colonnes sont différentes (lignes et colonnes représentent des ensembles qui seraient égaux si cela n'était);
iii- le tableau est maximal, i.e. on ne peut rajouter ni ligne ni colonne.

exemple: | 1 | 2 |
| --- | --- |
| 3 | 4 |

n'est pas une boite sur $\{1,2,3,4\}$ car on peut la compléter en:

1	2
3	4
1	4
3	2

ou

1	2	1	2
3	4	4	3

Considérons alors l'ensemble $L.C = \{ lxc \mid l \in L \quad c \in C \}$. On a:

Proposition 1: L'ensemble des sections de L.C est un m.r.n.a.

Pour prouver cette proposition, établissons le

Lemme: L'ensemble des sections de L.C est égal à l'ensemble des relations qui envoient L dans L et C dans C:

$$Sect(L.C) = \{ r \mid lr \in L,\ rc \in C \}$$

Preuve: Soit r dans Sect(L.C) et l dans L; pour tout c de C, il existe un unique (i,j) de $lxc \cap r$, d'où lr intersecte c en un seul point j et appartient donc à L. De manière analogue, on montre que $rc \in C$

Réciproquement, soient l dans L et c dans C; lr étant dans L, il intersecte c en un seul point j; de même rc et l s'intersectent en i; (i,j) est donc l'unique point d'intersection de r et lxc.

Preuve de la proposition 1: Le lemme prouve que Sect(L.C) est un monoide. Montrons que le produit est non-ambigu: soient r et s dans Sect(L.C) et (i,j) un élément de rs; prenons un l contenant i et un c contenant j; lr et sc s'intersectent en k qui est l'unique point tel que $(i,k) \in r$ et $(k,j) \in s$.
Par ailleurs la relation vide ne peut appartenir à Sect(L.C).
Enfin remarquons que les relations cxl pour c dans C et l dans L appartiennent à Sect(L.C); ceci assure la transitivité du monoide puisque L et C recouvrent E.

Considérons maintenant le treillis T des m.r.n.a. sur un ensemble E, ordonnés par inclusion. On a:

Proposition 2: L'ensemble des élémennts maximaux du treillis T est en bijection avec l'ensemble des boites sur E.

Eléments de preuve: Remarquons tout d'abord que les relations cxl, appartenant à Sect(L.C), en constituent l'idéal minimal car elles sont de rang 1. Un tel idéal est maximal pour l'inclusion d'aprés la maximalité des boites. Une étude plus détaillée des m.r.n.a. [BOE 76] montre que tout monoide de ce type se plonge dans un m.r.n.a. possédant un idéal minimal de rang 1 dont les éléments sont du type cxl avec $c \in C'$ $l \in L'$. On plonge alors L' et C' dans des familles L et C constituant une boite L.C et on vérifie que les relations du monoide appartiennent à Sect(L.C).

II-CODES ET BOITES

Les propositions précédentes fournissent une méthode de construction de codes finis maximaux à partir d'une boite: on recherche les relations 1-triangulaires sections de la boite et on associe à chaque lettre de l'alphabet une de ces relations.

exemple: soit la boite

1	3
2	2

$L.C = \{\{(1,1),(1,2),(3,1),(3,2)\},\{(1,2),(1,3),(3,2),(3,3)\},$
$\{(2,1),(2,2)\},\{(2,2),(2,3)\}\}$

Les sections 1-triangulaires de L.C sont:
$r_1 = \{(1,1),(1,3),(2,1),(2,3)\}$
$r_2 = \{(1,2),(2,1),(2,3)\}$
$r_3 = \{(3,1),(1,3),(2,1),(2,3)\}$

En choisissant $\mu_X(a) = r_1$, $\mu_X(b) = r_2$ on obtient:
$X = \{aa, ab, bb, aab, abb\}$

On remarquera que $r_3 = \{1,2\}x\{1,3\}$, c'est à dire le produit d'une colonne et d'une ligne de la boite.

III-CODES ET FACTORISATION

Etant données deux parties X et Y de A , on dit que le produit XY est non-ambigu si pour tout mot m de XY il existe un unique couple $(x,y)\in XxY$ tel que $m=xy$.

Conjecture de la factorisation:
Tout code fini maximal factorise le monoide libre, i.e. il existe deux parties finies P et Q de A^* telles que $A^* = QX^*P$ où le produit est non-ambigu.

Si à toute partie L de A on associe sa série caractéristique $\underline{L}$, élément de $Z\langle\langle A\rangle\rangle$, ceci s'écrit $\underline{A}^* = \underline{Q}\,\underline{X}^*\,\underline{P}$ et en observant que $\underline{X}^* = 1/(1-\underline{X})$ si X est un code, on obtient la formulation équivalente:

$$\underline{X}-1 = \underline{P}(\underline{A}-1)\underline{Q}$$

Récemment, Reutenauer [Reu 84] a montré que le polynome $\underline{X}-1$ se factorise en $R(\underline{A}-1)S$, R et S étant des polynomes à coefficients entiers (la conjecture spécifie que les coefficients sont égaux à 0 ou 1). En fait, son résultat est plus précis: il étend aux polynomes en variables non-commutatives un résultat de Schutzenberger [Sch 65] concernant les polynomes commutatifs associés aux codes.

IV-BOITES ET FACTORISATION

On dira qu'un m.r.n.a. est 1-triangulable s'il est plongeable dans un m.r.n.a. possédant une relation 1-triangulaire de rang 1.

Théorème 2: Un code maximal X fini factorise le monoide libre ssi il existe une représentation μ_X telle que $\mu_X(A^*)$ soit 1-triangulable.

Preuve: Soit X un code factorisant: $\underline{X}-1 = \underline{P}(\underline{A}-1)\underline{Q}$
Soit w une lettre extérieure à A et posons $\underline{X'}-1 = \underline{P}(\underline{A}+w-1)\underline{Q}$.
X' est un code vérifiant $X' = X \cup PwQ$.
L'image syntaxique de ww appartient à l'idéal minimal du monoide syntaxique de X'^* car tout mot de X' possède au plus une occurence de w. Mais il en est de même de w car il possède les mêmes contextes que ww puisque $PwQ\subset X'$. La représentation standard du monoide syntaxique de X'^* [Boe 76] possède donc une relation 1-triangulaire de rang 1 (l'image de w) et sa restriction μ_X à A^* est 1-triangulable.

Réciproquement, soit μ_X une représentation telle que $M = \mu_X(A^*)$ soit 1-triangulable. Définissons une extension $\mu_{X'}$ de μ_X en ajoutant une lettre w à A et en choisissant pour $\mu_{X'}(w)$ la matrice 1-triangulaire du plongement de M. Cette extension définit un code X' sur l'alphabet A+w. Le mot w est un mot synchronisant coupant [Boe 81] qui permet de factoriser $\underline{X'}-1$ en $\underline{P}(\underline{A}+w-1)\underline{Q}$ et donc $\underline{X}-1$ en $\underline{P}(\underline{A}-1)\underline{Q}$.

Ce résultat amène à poser la conjecture suivante: Etant donnée une boite (L,C), si L.C admet une section 1-triangulaire, il existe l de L et c de C tels que cxl soit 1-triangulaire.

En effet cette conjecture implique celle de la factorisation: Soit X un code maximal fini et μ_X une représentation. D'aprés la proposition 2, $\mu_X(A^*)$ est plongé dans Sect(L.C) pour une certaine boite (L,C). $\mu_X(A^*)$ est donc 1-triangulable et X factorise A^* d'aprés le théorème 2.

V-BOITES ET ALGORITHMIQUE

On a ramené le problème de la factorisation à un problème combinatoire se prétant mieux à la programmation pour la recherche de contre-exemple. Le but est de trouver des boites (L,C) telles qu'aucune relation cxl ne soit 1-triangulaire mais possédant toutefois des sections 1-triangulaires.

Il faut remarquer tout d'abord que la recherche des sections d'une famille de parties est un problème NP-complet. On peut donc s'attendre à ce que les temps d'éxécution croissent rapidement avec la taille des ensembles.

Le programme s'organise alors comme suit:

1- Construction de "modèles" de boites pour un ensemble donné: les modèles de boites sont les représentants des classes d'isomorphie des boites, deux boites étant isomorphes si elles s'échangent par permutation.
Ceci pose deux problèmes: peut-on construire toutes les boites sur n éléments à partir de boites sur n-1 éléments?
Quelles sont les procédures rapides permettant de vérifier l'isomorphisme de deux boites?

2- Rechercher les permutations de E telles que la boite associée (L,C) ne possède pas de parties l et c pour lesquelles cxl est 1-triangulable.

3- Rechercher les sections 1-triangulaires d'une telle boite. Apparait ici une autre question: caractériser les familles de parties ne possédant pas de sections. On observe cependant qu'une telle situation est souvent due au fait que deux parties disjointes sont incluses dans une meme troisième, ce qui se teste en un temps polynomial, alors que la recherche des sections est exponentielle.

REFERENCES

[BePe] J.Berstel, D.Perrin: The theory of codes, à paraitre.

[Boe 76] J.M. Boe: Représentations des monoides; applications à la théorie des codes. Thèse 3ème cycle Montpellier (1976).

[Boe 81] J.M. Boe: Sur les codes synchronisants coupants, Proc. Colloquim Arco Felice, Naples, C.N.R. Roma (1981) 7-10.

[Reu 84] C. Reutenauer: Non commutative factorization of variable length codes, à paraitre.

[Sch 65] M.P. Schutzenberger: Sur certains sous-monoides libres, Bull. Soc. Math. France 93(1965) 209-223.