

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](http://www.numdam.org/item?id=PDML_1984__6B_A6_0)

© Université de Lyon, 1984, tous droits réservés.

L'accès aux archives de la série « Publications du Département de mathématiques de Lyon » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## 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 (monoïde libre) sur l'alphabet  $A$ . Pour une partie  $X$  de  $A^*$ ,  $X^*$  désigne le sous-monoïde 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:

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

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

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

Pour étudier les propriétés de  $X$ , nous représentons le monoïde libre  $A$  par des relations sur un ensemble  $E = [1..n]$ .  $\mu_X$  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-ambiguïté du produit des relations définie ci-après:

Un produit de deux relations  $r$  et  $s$  (matrices booléennes) est non-ambigu si  $(i,j) \in rs$  entraîne 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 monoïde de relations non-ambigu est un monoïde 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 monoïdes de relations transitifs: tout couple  $(i,j)$  est présent dans une des relations du monoïde. Il vient:

Théorème 1 [BePe]: Une partie  $X$  de  $A^*$  est un code maximal fini ssi i) existe une représentation  $\mu_X$  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 monoïde 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 identifiant  $(1,1)$  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  $\mu_X(m) = \{(u,v), (u',v') \in S \times S \mid um \in X^*u' \text{ et } mv' \in vX^*\}$

On vérifie alors que  $\mu_X$  satisfait les conditions i-iv du théorème.

Réciproquement, les conditions i et ii assurent que  $X^*$  est un sous-monoïde 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 monoïde de relations non-ambigu transitif ne contenant pas la relation vide. Par ailleurs, étant donné un code maximal fini  $X$ ,  $\mu_X$  désignera une représentation vérifiant les propriétés du théorème.

## I-BOITES ET MONOÏDE 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 boîte 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 boîte par un tableau  $T$  dont l'élément  $T(i,j)$  est l'intersection du  $i$ ème élément de  $L$  et du  $j$ ème élément de  $C$ .

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

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

Un tableau d'éléments de  $E$  est une boîte s'il vérifie:

i- si un élément est présent en  $(i,j)$  et  $(k,l)$  alors il apparaît 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 boîte 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  $L.C = \{lxc \mid l \in L, c \in C\}$ . 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:

$$\text{Sect}(L.C) = \{r \mid l r \in L, r c \in C\}$$

Preuve: Soit  $r$  dans  $\text{Sect}(L.C)$  et  $l$  dans  $L$ ; pour tout  $c$  de  $C$ , il existe un unique  $(i,j)$  de  $lxc/r$ , 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  $\text{Sect}(L.C)$  est un monoïde. Montrons que le produit est non-ambigu: soient  $r$  et  $s$  dans  $\text{Sect}(L.C)$  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  $\text{Sect}(L.C)$ . Enfin remarquons que les relations  $cxl$  pour  $c$  dans  $C$  et  $l$  dans  $L$  appartiennent à  $\text{Sect}(L.C)$ ; ceci assure la transitivité du monoïde 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 éléments maximaux du treillis  $T$  est en bijection avec l'ensemble des boîtes sur  $E$ .

Éléments de preuve: Remarquons tout d'abord que les relations  $cxl$ , appartenant à  $\text{Sect}(L.C)$ , en constituent l'idéal minimal car elles sont de rang 1. Un tel idéal est maximal pour l'inclusion d'après la maximalité des boîtes. Une étude plus détaillée des m.r.n.a. [BOE 76] montre que tout monoïde 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 boîte L.C et on vérifie que les relations du monoïde appartiennent à  $\text{Sect}(L.C)$ .

## II-CODES ET BOITES

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

exemple: soit la boîte

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  $r_3 = \{1,2\} \times \{1,3\}$ , c'est à dire le produit d'une colonne et d'une ligne de la boîte.

### 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  $(x,y) \in X \times Y$  tel que  $m=xy$ .

Conjecture de la factorisation:

Tout code fini maximal factorise le monoïde 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 monoïde libre ssi il existe une représentation  $\mu_X$  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 monoïde syntaxique de  $X'^*$  car tout mot de X' possède au plus une occurrence de w. Mais il en est de même de w car il possède les mêmes contextes que ww puisque  $PwQCX'$ . La représentation standard du monoïde syntaxique de  $X'^*$  [Boe 76] possède donc une relation 1-triangulaire de rang 1 (l'image de w) et sa restriction  $\mu_X$  à  $A^*$  est 1-triangulable.

Réciproquement, soit  $\mu_X$  une représentation telle que  $M = \mu_X(A^*)$  soit 1-triangulable. Définissons une extension  $\mu_{X'}$  de  $\mu_X$  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  $X'-1$  en  $\underline{P}(A+w-1)\underline{Q}$  et donc  $X-1$  en  $\underline{P}(A-1)\underline{Q}$ .

Ce résultat amène à poser la conjecture suivante: Etant donnée une boîte  $(L,C)$ , 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  $\mu_X$  une représentation. D'après la proposition 2,  $\mu_X(A^*)$  est plongé dans  $\text{Sect}(L.C)$  pour une certaine boîte  $(L,C)$ .  $\mu_X(A^*)$  est donc 1-triangulable et  $X$  factorise  $A^*$  d'après le théorème 2.

## V-BOITES ET ALGORITHMIQUE

On a ramené le problème de la factorisation à un problème combinatoire se prêtant mieux à la programmation pour la recherche de contre-exemple. Le but est de trouver des boîtes  $(L,C)$  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'exécution croissent rapidement avec la taille des ensembles.

Le programme s'organise alors comme suit:

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

2- Rechercher les permutations de  $E$  telles que la boîte 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 boîte. Apparaît 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 même 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 paraître.
- [Boe 76] J.M. Boe: Représentations des monoïdes; 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 paraître.
- [Sch 65] M.P. Schutzenberger: Sur certains sous-monoïdes libres, Bull. Soc. Math. France 93(1965) 209-223.