

B. LECLERC

B. MONJARDET

Avant-propos

Mathématiques et sciences humaines, tome 98 (1987), p. 5-7

http://www.numdam.org/item?id=MSH_1987__98__5_0

© Centre d'analyse et de mathématiques sociales de l'EHESS, 1987, tous droits réservés.

L'accès aux archives de la revue « Mathématiques et sciences humaines » (<http://msh.revues.org/>) 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/>

AVANT-PROPOS

B. LECLERC* et B. MONJARDET**

Ce numéro de Mathématiques et Sciences humaines ainsi qu'un numéro ultérieur sont consacrés au thème "Combinatoire et Analyse des données". Le terme Combinatoire revêtant des acceptions assez variées est malaisé à définir. Disons que nous l'entendons ici au sens des mathématiques traitant de ces configurations discrètes que sont par exemple, les relations binaires (ou n-aires), les ordres, les graphes, les hypergraphes, les plans en bloc, les géométries d'incidence..., lorsqu'elles sont définies sur des ensembles finis. Quant à l'analyse de données, nous l'entendons bien sûr en un sens large qui contient les nombreuses techniques permettant d'explorer, décrire, résumer, visualiser,... des "données" que leurs nombres et (ou) leur caractère multidimensionnel rendent peu transparentes.

Que Mathématiques et Sciences humaines consacre des numéros spéciaux à ce thème n'a rien d'étonnant. L'analyse des données - lorsqu'on n'en fait pas une panacée à l'absence de réflexion - joue et doit jouer un rôle important dans les techniques mathématiques utilisables en Sciences humaines. D'ailleurs beaucoup de méthodes d'analyse des données et notamment celles utilisant des techniques combinatoires ont été créées à propos de données et de problèmes de ces sciences : c'est par exemple le cas des méthodes de sériation, d'un certain nombre de méthodes typologiques, de l'analyse de similitude, de l'analyse hiérarchique (échelles de Guttman) et booléenne des questionnaires ...

Un article de Guénoche et Monjardet dans le second numéro spécial essaiera de répertorier ces méthodes de la "combinatorial data analysis" (Arabie) en les classant d'après la nature des tableaux de données qu'elles permettent de traiter. En fait c'est bien parce que dans les Sciences humaines les données recueillies ont souvent une structure mathématique "pauvre" (qualitative, relationnelle, ordinale,...) que des méthodes spécifiques, évi-

* C.A.M.S.

** Université Paris 5 et C.A.M.S.

tant les problèmes posés par l'arbitraire d'un codage numérique, sont apparues. La contrepartie est que l'utilisation de telles méthodes pose souvent des problèmes de calcul "difficiles" (en un sens que la théorie récente de la complexité permet de préciser). Toutefois les progrès actuels du carrefour mathématiques - informatique que constitue l'algorithmique, permettent de plus en plus de compléter les études antérieures surtout théoriques de certaines méthodes, par la réalisation de programmes ou logiciels; il s'ensuit une pratique permettant d'éclairer les avantages, les désavantages ou les questions posées par l'emploi de telles méthodes.

C'est ainsi que la pratique de l'analyse booléenne de questionnaires à l'aide d'un logiciel approprié a conduit J. Van Buggenhaut, dans un des articles de ce numéro, à proposer une approche pour un problème apparaissant dans une étape importante de cette méthode : le choix du "seuil de dichotomisation".

La difficulté de certains problèmes de calcul peut aussi conduire à une recherche axée sur les algorithmes les résolvant. C'est ainsi que plusieurs algorithmes ont maintenant été proposés pour approcher au mieux une dissimilarité par une "distance additive d'arbre" (i.e. une distance du plus court chemin dans un arbre dont les sommets correspondent aux objets comparés). Dans son article A. Guénoche fait une présentation détaillée de cinq d'entre eux (certains originaux) et conclut par un premier essai de comparaison de leurs performances.

Il arrive toutefois que les choses se passent "bien". Le problème de l' "arbre minimum" (d'un graphe valué) est un exemple célèbre de problème d'optimisation combinatoire où l'on dispose d'algorithmes efficaces de calcul. Il en est de même du problème plus général de rechercher s'il existe un arbre minimum "commun à plusieurs données", i.e. commun aux préordonnances (préordres sur les paires d'objets) canoniquement associées à ces données. Ceci provient de résultats théoriques énoncés dans l'article de B. Leclerc, article qui illustre aussi la variété des contextes où un tel problème peut apparaître.

Nous terminerons la revue des articles de ce numéro par celui de Giakoumakis et Monjardet qui étudient systématiquement une notion - élémentaire - d'ordre entre divers coefficients mesurant la similarité de préordres totaux afin de les comparer.

En conclusion, ces numéros spéciaux de Mathématiques et Sciences humaines illustrent bien le fait que les méthodes combinatoires en analyse des données sont en plein développement. Dans ce développement, les études théoriques s'accompagnent de plus en plus d'implémentations et de logiciels. De ce fait, les utilisations concrètes de ces méthodes tendent à se multiplier; souhaitons que ces numéros puissent contribuer à cette évolution.