

ISSN 1516-5388

EDUCAÇÃO
MATEMÁTICA
PESQUISA

v. 4 – n. 2 – 2002

v. 4 – n. 2

EDUCAÇÃO MATEMÁTICA PESQUISA

educ

educ

EDUCAÇÃO MATEMÁTICA PESQUISA

revista do programa de estudos pós-graduados em educação matemática

puc-sp

EDUCAÇÃO MATEMÁTICA PESQUISA

revista do programa de estudos pós-graduados em educação matemática

puc-sp

Editores

Sonia Barbosa Camargo Iglori e Wagner Rodrigues Valente

Conselho Executivo

Ana Paula Jahn, Janete Bolite Frant, Lulu Healy, Maria Cristina S. de A. Maranhão, Saddo Ag Almouloud, Sonia Barbosa Camargo Iglori e Wagner Rodrigues Valente

Conselho Científico

Ana Mesquita (Université Strasbourg, França), Beatriz D'Ambrósio (Indianapolis University, EUA), Celia Hoyles (Institut Education University of London, Inglaterra), Circe da Silva Dynnikov (UFES), Gilda de La Roque Palis (PUC-RJ), Joaquim Gimenez (Universidad de Barcelona, Espanha), Marilena Bittar (UFMS), Michele Artigue (Université Paris VII, França), Mirian Jorge Warde (PUC-SP), Nilson José Machado (FEUSP), Raymond Duval (Université Lille, França), Regina Damm (UFSC), Ricardo Nemirovsky (TERC, EUA), Sérgio Nobre (UNESP-Rio Claro), Terezinha Nunes (Oxford Brookes University, Inglaterra), Vinício Macedo Santos (UNESP – Presidente Prudente)

A Educação Matemática Pesquisa conta com o trabalho de pareceristas *ad hoc*.

Correspondência:

Revista do Programa de Estudos Pós-Graduados em Educação Matemática

Rua Marquês de Paranaguá, 111 – CEP 01303-050 – Consolação – São Paulo – SP

Fone: (11) 3124-7210

Fax: (11) 3159-0189

E-mail: pedmat@exatas.pucsp.br

Expediente: de segunda a sexta-feira das 10h30min às 12h e das 13h30min às 17h30min

EDUCAÇÃO MATEMÁTICA PESQUISA

revista do programa de estudos pós-graduados em educação matemática

puc-sp

ISSN 1516-5388

Educ. Mat. Pesqui., São Paulo, v. 4, n. 2, pp. 1-150, 2002

educ
2002

Educação Matemática Pesquisa : Revista do Programa de Estudos
Pós - Graduados em Educação Matemática / Pontifícia Universidade
Católica de São Paulo - n. 1 (março de 1999) - São Paulo : EDUC, 1999 -
semestral
ISSN 1516-5388

1. Educação Matemática Pesquisa - periódicos. I. Pontifícia Universidade
Católica de São Paulo. Programa de Estudos Pós-Graduados em Educa-
ção Matemática

EDUC – Editora da PUC-SP

Direção

Maria Eliza Mazzilli Pereira
Denize Rosana Rubano

Coordenação Editorial

Sonia Montone

Revisão

Sonia Rangel

Revisão de Inglês

Carolina Muniz Ventura Siqueira

Editoração Eletrônica

Digital Press

Capa

Sara Rosa

educ

Rua Ministro Godói, 1.197
Cep 05015-001 - São Paulo - SP
Fone: (11) 3873-3359 / Fax: (11) 3873-6133
E-mail: educ@pucsp.br
Site: www.pucsp.br/educ



Projeto Editorial

A revista *Educação Matemática Pesquisa*, do Programa de Estudos Pós-Graduados em Educação Matemática da PUC-SP, de regularidade semestral, constitui um espaço de divulgação de pesquisas científicas da área.

O projeto editorial da revista prioriza artigos científicos, inéditos no Brasil, da área de Educação Matemática. Mais particularmente, os relacionados às linhas de pesquisa do Programa: A Matemática na estrutura curricular e formação de professores; História, Epistemologia e Didática da Matemática; e Tecnologias da Informação e Didática da Matemática. A prioridade dada às linhas descritas não é extensiva aos referenciais teóricos, ao contrário, procura-se contemplar a diversidade.

Serão acolhidos, também, artigos que favoreçam o diálogo entre Educação Matemática e áreas afins, como a Matemática, a Epistemologia, a Psicologia Educacional, a Filosofia, a História da Matemática e a História das Disciplinas, entre outras.

A seleção dos artigos faz-se mediante a aprovação de dois pareceristas do conselho editorial ou *ad hoc*. Os pareceres serão enviados aos autores.

Os artigos são apresentados sempre na versão original, com resumos bilíngües (português e inglês ou francês).

O projeto editorial prevê, ainda, que os volumes da revista contendam uma ou mais modalidades, como análises ou relatos de pesquisa, comunicações (ciclo de palestras, conferências), entrevistas, depoimentos ou resenhas científicas.

Em cada número, haverá indicações sucintas das dissertações e teses produzidas no Programa, no semestre de edição.

Conselho Editorial

Editorial Project

The journal Educação Matemática Pesquisa of the Post-Graduate Studies in Mathematics Education Program, PUC/SP, is published each semester with the aim of providing a space for disseminating scientific research in the area.

The policy adopted by the editors is to prioritize scientific articles related to Mathematics Education previously unpublished in Brazil and particularly those addressing the lines of research of the program: Mathematics, curriculum structure and teacher training; History, Epistemology and Didactics of Mathematics; and Information technology and the Didactics of Mathematics. The priority given to the described lines is not restricted to theoretical references, on the contrary, it is hoped that the journal will reflect the diversity that characterizes research in Mathematics Education.

The editors also encourage the submission of articles which open dialogues between Mathematics Education and related areas, such as Mathematics, Epistemology, Educational Psychology, Philosophy, History of Mathematics and its teaching amongst others.

In order to be selected, articles should receive two favorable reviews. Referees will be chosen from the editorial committee or ad hoc reviewers. Authors will receive copies of the reviews.

Articles will be presented always in the original language of the author along with abstracts in Portuguese and English or French.

The editors anticipate that the volumes of the journal will include works of various different types including, for example: research reports, papers based on lectures or conferences, interviews, commentaries on issues pertaining to research, critiques of articles and books, literature reviews and theoretical analyses.

Each issue, will also include the titles of the dissertations and theses produced in the Program during the semester of the edition.

Editorial Committee

Sumário

Editorial	9
L'analyse statistique implicative: ses bases, ses développements <i>(A análise estatística implicativa: suas bases, seu desenvolvimento)</i> <i>(Implicative statistical analysis: bases and development)</i> Régis Gras	11
Une étude diagnostique en vue de la formation des enseignants en géométrie <i>(Um estudo diagnóstico com vistas à formação dos professores em geometria)</i> <i>(A diagnostic study aiming at geometry teachers' education)</i> Saddo Ag Almouloud	49
A implicação estatística usada como ferramenta em um exemplo de análise de dados multidimensionais <i>(Statistical Implicative Analysis as a tool in a multidimensional data case)</i> Régis Gras e Saddo Ag Almouloud	75
Possíveis relações entre o processo de formação e a prática pedagógica: proposta de procedimentos de pesquisa <i>(Possible relationships between the education process and teaching practice: a proposal for research procedures)</i> Ana Lúcia Manrique e Marli E. D. A. André	89
O uso do CHIC na análise de registros textuais em ambiente virtual de formação de professores <i>(The use of CHIC in the analysis of textual registers in a virtual environment for teacher education)</i> Maria Elisabette Brisola Brito Prado	103

Formação de professores para inserção do computador na escola: inter-relações entre percepções evidenciadas pelo uso do software CHIC <i>(Teacher education for the insertion of computers in schools: inter-relationships between perceptions revealed by the use of the software CHIC)</i> Maria Elizabeth Bianconcini de Almeida	125
Dissertações defendidas no segundo semestre de 2002 <i>(Dissertations completed during the second semester of 2002)</i>	145
Normas para a publicação	149

O presente número da revista *Educação Matemática Pesquisa*, segundo do volume 4, referente ao segundo semestre de 2002, contém seis artigos enfocando o método estatístico da análise implicativa desenvolvido pelo pesquisador francês Régis Gras, professor emérito da Universidade de Nantes (França). Os artigos exploram também a utilização do programa CHIC, ferramenta informática que possibilita o uso do método de Gras. O CHIC é um programa em constante atualização, estando hoje em sexta versão. A primeira versão foi desenvolvida por Saddo Ag Almouloud, pesquisador do Programa de Pós-graduação em Educação Matemática da PUC-SP.

No primeiro artigo, Régis Gras, seu autor, destaca a importância do uso de regras na evolução do pensamento calculatório do homem e, em sua teoria, define regras não simétricas necessárias para modelizar as relações do tipo "se a então quase".

No segundo artigo, Saddo Ag Almouloud apresenta uma investigação diagnóstica sobre o ensino e a aprendizagem da geometria em nível do ensino fundamental. Ele utiliza o método estatístico de análise implicativa como instrumental de análise de dados coletados na referida pesquisa.

Os dois pesquisadores, Régis Gras e Saddo Ag Almouloud apresentam, no terceiro artigo, um exemplo de análise realizada pelo método da implicação estatística.

Ana Lúcia Manrique e Marli André, respectivamente doutoranda e professora do Programa de Estudos Pós-Graduados em Psicologia da Educação da PUC-SP, desenvolvem um estudo para identificar possíveis relações entre o processo de formação profissional, proporcionado pelo Proformação, e a prática pedagógica do professor cursista. No artigo, é apresentada uma análise hierárquica de similaridade dos dados coletados, utilizando o programa CHIC.

Os artigos cinco e seis têm como foco o tema da formação do professor e o uso do computador. As autoras, respectivamente Maria Elisabete Brisola Brito Prado e Maria Elizabeth Bianconcini de Almeida, são professoras do Programa de Estudos Pós-Graduados em Educação: Currículo, da PUC-SP. No quinto artigo são abordadas as contribuições do CHIC na ampliação das possibilidades de análise do processo de aprendizagem dos professores-alunos na vivência de um curso de formação que enfatizava a reconstrução da prática pedagógica. No sexto artigo é apresentada a metodologia de análise de dados qualitativos desenvolvida com o uso do programa CHIC evidenciando as inter-relações entre percepções dos sujeitos participantes de uma ação de formação de professores para a inserção do computador na prática pedagógica.

Completam esse periódico os títulos das dissertações de mestrado do Programa de Estudos Pós-Graduados em Educação Matemática da PUC-SP, defendidas no segundo semestre de 2002.

Editores

The current issue of the journal *Educação Matemática Pesquisa* (volume 4, number 2, corresponding to the second semester of 2002) consists of six articles focussing on the method of implicative statistical analysis, developed by the French researcher Régis Gras, professor emeritus of the University of Nantes (France). The articles also explore the use of the software CHIC, a computational tool which implements Gras' method. CHIC has undergone a series of revisions and is today in its 6th version. The first version was developed by Saddo Ag Almouloud, researcher of the Post-Graduation Program in Mathematics Education of PUC-SP.

In the first article, Régis Gras highlights the importance of the use of rules in the evolution of computational thinking and develops his theory involving the definition of asymmetrical rules that are necessary to model relationships of the type « if a then nearly b ».

In the second article, Saddo Ag Almouloud presents a diagnostic investigation related to the teaching and learning of geometry in primary and secondary school. He utilises the method of implicative statistical analysis as the instrument for analysing the data collected in his research.

In the third article, the researchers Régis Gras and Saddo Ag Almouloud present an example of the method of implicative statistical analysis.

Ana Lúcia Manrique and Marli André, respectively doctoral student and lecturer of the Post-Graduation Program in Educational Psychology of PUC-SP, describe in the fourth article a study aiming to identify possible relationships between the process of professional development provided by Proformação (the Brazilian Ministry of Education's teacher-training program) and the participant teachers' pedagogical practice. The article presents a hierarchical analysis of the similarities within the collected data using the software CHIC.

Articles five and six focus on the theme of teacher education and computer use. The authors, Maria Elisabete Brisola Brito Prado and Maria Elizabeth Bianconcini, are lecturers of the Post-Graduation Program in Education and Curriculum of PUC-SP. The fifth article approaches the contribution of CHIC to amplify the possibilities for analysing teacher-students' learning processes during a teacher education course that emphasised the reconstruction of pedagogic practice. The sixth and final article discusses the methodology of qualitative data analysis developed with the use of the software CHIC and reveals the inter-relationships between perceptions of the subjects, who are participants of a teacher education activity related to computer insertion into pedagogic practice.

Finally, the journal presents the titles of the Masters dissertations completed by students of the Post-Graduation Program in Mathematics Education of PUC-SP during the second semester of 2002.

The Editors

L'analyse statistique implicative: ses bases, ses développements

RÉGIS GRAS*

Resumo

Como parte de situações de didática da matemática, o método implicativo se desenvolveu à medida que os problemas foram encontrados e as questões formuladas. Seu objetivo maior visa à estruturação de dados cruzando indivíduos e variáveis, à extração de regras entre variáveis e, a partir da contingência dessas regras, à explicação e, portanto, à previsão em diversos campos: psicologia, sociologia, biologia, etc. É assim que os conceitos de intensidade de implicação, de coesão de classes, de implicação-inclusão, de significância de níveis hierárquicos, de contribuição de variáveis suplementares, etc. foram construídos. Ao tratamento de variáveis binárias acrescentou-se o das variáveis modais, freqüenciais e, recentemente, o tratamento das variáveis-intervalo.

Palavras-chave: implicação estatística; implicação-inclusão; entropia; coesão; grafo implicativo; árvore hierárquica; nível significativo; variável binária; variável modal; variável freqüencial; variável-sobre-intervalo; regra; metaregra; contribuição; tipicidade.

Résumé

Partie de situations de didactique des mathématiques, la méthode implicative se développe au fil des problèmes rencontrés et des questions posées. Son objectif majeur vise la structuration de données croisant individus et variables, l'extraction de règles entre les variables et, à partir de la contingence de ces règles, l'explication et donc la prévision dans divers domaines : psychologie, sociologie, biologie, etc. C'est ainsi que les concepts d'intensité d'implication, de cohésion de classes, d'implication-inclusion, de significativité de niveaux hiérarchiques, de contribution de variables supplémentaires, etc., sont fondés. De la même façon, au traitement de variables binaires s'ajoutent ceux des variables modales, fréquentielles et, récemment, de variables-intervalles.

Mots-clé: implication statistique; implication-inclusion; entropie; cohésion; graphe implicatif; arbre hiérarchique; niveau significatif; variable binaire; variable modale; variable fréquentielle; variable-sur-intervalle; règle; méta-règle; contribution; typicalité.

Abstract

As part of mathematics teaching situations, the implicative method was developed at the same time that problems were being found and questions, formulated. Its major objective aims at: data structuring,

* Professeur Emérite à l'Ecole Polytechnique de l'Université de Nantes, La Chantrerie, BP 50609 44306 Nantes cedex 03. E-mail: regisgra@club-internet.fr.

by crossing individuals and variables; the extraction of rules between variables; and, from the contingency of those rules, the explanation and, therefore, prediction in several fields: psychology, sociology, biology, etc. It was in this way that the concepts of implication intensity, class cohesion, implication-inclusion, significance of hierarchical levels, contribution of supplementary variables, etc., were constructed. In addition to the treatment of binary variables, we now deal with the treatment of modal variables, frequency variables, and interval variables.

Key-words: statistical implication; implication-inclusion; entropy; cohesion; implicative graph; hierarchical tree; significant level; binary variable; modal variable; frequency variable; interval variable; rule; meta-rule; contribution; typicality.

Preambule

Les connaissances opératoires de l'homme sont principalement constituées selon deux composantes: celle des faits et celles des règles entre les faits ou entre règles elles-mêmes. Ce sont ses apprentissages à travers sa culture ou ses expériences personnelles qui lui permettent une élaboration progressive de ces connaissances, même si des régressions, des remises en cause viennent continuellement lui assurer un équilibre fonctionnel. Or les règles se forment de façon stable dès lors que le nombre de succès quant à leur qualité prédictive ou anticipatrice atteint un certain seuil de confiance en deçà duquel elles seront susceptibles d'être prudemment mises en oeuvre. En revanche, ce seuil étant dépassé, l'économie de l'individu le fera résister à son abandon ou à sa critique. En effet, il est coûteux de substituer à la règle initiale une autre règle lors de l'apparition d'un contre-exemple, du fait qu'elle a été confortée par un nombre important de réussites. Un second contre-exemple (ou plus selon la qualité de robustesse du seuil de confiance en la règle) sera peut-être nécessaire à un réajustement, voire à son abandon. Laurent Fleury, dans sa thèse (Université de Nantes, 1997) cite avec pertinence l'exemple de la règle: "toutes les Ferrari sont rouges". Cette règle, très robuste, ne sera pas abandonnée lors de l'observation d'un seul ou de deux contre-exemples. D'autant qu'elle ne manquerait pas d'être rapidement reconfortée.

Ainsi, à l'opposé de ce qui est légitime en mathématiques où toute règle (théorème) ne souffre pas d'exception, où le déterminisme est total, les règles en sciences humaines, plus généralement en sciences dites "molles", sont acceptables et donc opératoires tant que le nombre de contre-exemples restera "supportable" par la fréquence de situations où elles seront efficaces. Le problème, en analyse des données, est alors

d'établir un critère, relativement consensuel, pour définir la notion de seuil de confiance ajustable au niveau d'exigence de l'utilisateur de la règle. Qu'il soit établi sur des bases statistiques a tout lieu de ne pas surprendre. Qu'il possède une propriété de résistance au bruit (faiblesse du ou des premiers contre-exemples) peut également paraître naturel, conforme au sens "économique" évoqué plus haut. Qu'il s'effondre si les contre-exemples se répètent semble aussi devoir guider notre choix dans la modélisation du critère recherché. Ce texte présente le choix épistémologique que nous avons fait. En tant que tel il est donc réfutable mais le nombre de situations où il s'est avéré pertinent et fécond nous conduit à en restituer ici la genèse.

Introduction

Différentes approches théoriques ont été adoptées pour modéliser l'extraction et la représentation de règles d'inférence imprécises (ou partielles) entre variables binaires (ou attributs ou caractères) décrivant une population d'individus (ou sujets ou objets). Mais les situations de départ et la nature des données ne modifient pas la problématique fondamentale. Il s'agit de découvrir des règles non symétriques pour modéliser des relations du type "si *a* alors presque *b*". C'est, par exemple, l'option des réseaux bayésiens (par exemple: Amarger et Pearl, 1988) ou des treillis de Galois (par exemple: Simon, 2000). Mais le plus souvent, la probabilité conditionnelle est le moteur de la définition de l'association, même si l'indice de cette association retenu est de type multivarié (par exemple: Bernard, 1999). De plus et à notre connaissance, d'une part, le plus souvent les développements s'arrêtent à la proposition d'un indice d'implication partielle pour des données binaires, d'autre part, cette notion n'est pas étendue à l'extraction et la représentation hiérarchique de méta-règles, ni à la recherche des sujets et catégories de sujets responsables des associations. Nous proposons ici ces prolongements après avoir rappelé le paradigme fondateur.

L'intensité d'implication dans le cas binaire

Une population E d'objets ou de sujets est croisée avec des variables (caractères, critères, réussites) que l'on interroge de la façon suivante: "dans

quelle mesure peut-on considérer que le fait de relever de la variable a implique celui de relever de la variable b ? Autrement dit, les sujets ont-ils tendance à être b si l'on sait qu'ils sont a ?". Dans les situations naturelles, humaines ou sciences de la vie, où les théorèmes (si a alors b) au sens mathématique du terme ne peuvent être établis du fait des exceptions qui les entachent, il est important pour le chercheur et le praticien de "fouiller dans ses données" afin de dégager cependant des règles assez fiables (des sortes de "théorèmes partiels") pour pouvoir conjecturer une genèse, décrire, structurer une population et faire l'hypothèse d'une certaine stabilité à des fins prédictives. Mais cette fouille exige la mise au point de méthodes pour la guider et pour la dégager du tâtonnement et de l'empirisme.

Pour cela, à l'instar de la méthode de mesure de la similarité de Lerman (1970, 1981), nous définissons (Gras, 1979, 1996) la mesure de la relation implicative $a \Rightarrow b$ à partir de l'in vraisemblance de l'apparition, dans les données, du nombre de cas qui l'infirmant, c'est-à-dire pour lesquels a est vérifié sans que b ne le soit. Cette mesure est relativisée au nombre de données vérifiant respectivement a et non b . Elle quantifie "l'étonnement" de l'expert devant le nombre invraisemblablement petit de contre-exemples eu égard à une indépendance présumée et aux effectifs en jeu.

Précisons. Un ensemble fini V de v variables est donné: a, b, c, \dots Dans la situation paradigmatique classique, il s'agit des performances (réussite-échec) à des items d'un questionnaire. A un ensemble fini E de n sujets x , on associe, par abus d'écriture, les fonctions du type: $x \rightarrow a(x)$ où $a(x) = 1$ (ou $a(x) = \text{vrai}$) si x satisfait ou possède le caractère a et 0 (ou $a(x) = \text{faux}$) sinon. En intelligence artificielle, on dira que x est un exemple ou une instance pour a si $a(x) = 1$ et un contre-exemple dans le cas contraire.

La règle " $a \Rightarrow b$ " est logiquement vraie si pour tout x , $b(x)$ n'est nul que dans le cas où $a(x)$ l'est aussi ; autrement dit si l'ensemble A des x pour lesquels $a(x)=1$ est contenu dans l'ensemble B des x pour lesquels $b(x)=1$. Cependant, cette inclusion stricte n'est qu'exceptionnellement observée dans la réalité. Dans le cas d'un questionnaire de connaissances, on pourrait en effet observer quelques rares élèves réussissant un item a et ne réussissant pas l'item b , sans que soit contestée la *tendance* à avoir b quand on a a . Relativement aux cardinaux de E (soit n), mais aussi de A (soit n_a) et B (soit n_b), c'est donc le «poids» des contre-exemples (soit $n_a \wedge \bar{b}$) qu'il faudra prendre en compte pour accepter statistiquement de conserver ou non la *quasi-implication* ou la *quasi-règle* " $a \Rightarrow b$ ".

Pour mathématiser cette quasi-règle, nous considérons, comme le fait Lerman pour la similarité, deux parties quelconques X et Y de E , choisies aléatoirement et indépendamment (absence de lien a priori entre ces deux parties) et de mêmes cardinaux respectifs que A et B . Soit \bar{Y} et \bar{B} les complémentaires respectifs de Y et de B dans E de même cardinal $n_{\bar{b}} = n - n_b$.

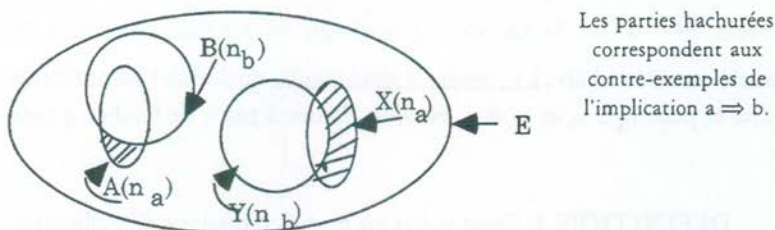


Figure 1

Nous dirons alors:

DEFINITION 1: $a \Rightarrow b$ est admissible au niveau de confiance $1 - \alpha$ si et seulement

$$\text{si } \Pr[\text{card}(X \cap \bar{Y}) \leq \text{card}(A \cap \bar{B})] \leq \alpha$$

Il est établi (Lerman, 1981) que, pour un certain processus de tirage: la variable aléatoire $\text{Card}(X \cap \bar{Y})$ suit la loi de Poisson de paramètre $\frac{n_a n_{\bar{b}}}{n}$.

Dans le cas où $n_{\bar{b}} \neq 0$, nous réduisons et centrons cette variable de Poisson en la variable:

$$Q(a, \bar{b}) = \frac{\text{Card}(X \cap \bar{Y}) - \frac{n_a n_{\bar{b}}}{n}}{\sqrt{\frac{n_a n_{\bar{b}}}{n}}}$$

Dans la réalisation expérimentale, la valeur observée de $Q(a, \bar{b})$ est $q(a, \bar{b})$

$$\text{DEFINITION 2: } q(a, \bar{b}) = \frac{n_{a \wedge \bar{b}} - \frac{n_a n_{\bar{b}}}{n}}{\sqrt{\frac{n_a n_{\bar{b}}}{n}}} \quad \text{est appelé indice}$$

d'implication, nombre que nous retenons comme indicateur de la non-implication de a sur b.

Dans les cas légitimant convenablement l'approximation (par exemple, $\frac{n_a n_{\bar{b}}}{n} \geq 3$), la variable $Q(a, \bar{b})$ suit approximativement la loi normale centrée réduite. L'intensité d'implication, qualité de l'admissibilité de $a \Rightarrow b$, pour $n_a \leq n_b$ et $n_b \neq n$, est alors définie à partir de l'indice $q(a, \bar{b})$ par:

DEFINITION 3: Dans le cas où $n_b \neq n$, l'intensité d'implication de a sur b est:

$$\varphi(a, b) = 1 - \Pr[Q(a, \bar{b}) \leq q(a, \bar{b})] = \frac{1}{\sqrt{2\pi}} \int_{q(a, \bar{b})}^{\infty} e^{-\frac{t^2}{2}} dt$$

Par suite, la définition de l'implication statistique devient:

DEFINITION 4: L'implication $a \Rightarrow b$ est admissible au niveau de confiance $1 - \alpha$ si et seulement si: $\varphi(a, b) \geq 1 - \alpha$.

Rappelons que cette modélisation de la quasi-implication mesure l'étonnement de constater la petitesse des contre-exemples en regard du nombre surprenant des instances de l'implication. Par conséquent, si la règle est triviale, comme dans le cas où B est très grand ou coïncide avec E, cet étonnement devient petit. Nous démontrons (Gras, 1996) d'ailleurs que cette trivialité se traduit par une intensité d'implication très faible, voire nulle:

Si, n_a étant fixé et A étant inclus dans B, n_b tend vers n (B "croît" vers E), alors $\varphi(a, b)$ tend vers 0.

Remarque 1: D'autres modélisations, autres que celle de Poisson, sont possibles. Citons:

* *une modélisation binomiale:* considérant les variables duales $\text{card}(A \cap \bar{Y})$ et $\text{card}(X \cap \bar{B})$, où X et Y sont des parties choisies de façon indépendante dans E et respectant les propriétés cardinales respectives de A et B , tout élément de E , par exemple, a la probabilité $\frac{na}{n} \frac{n\bar{b}}{n}$ d'appartenir à $(A \cap \bar{Y})$. Par suite:

$$\Pr[\text{card}(A \cap \bar{Y}) = k] = C_n^k \left(\frac{na}{n} \frac{n\bar{b}}{n}\right)^k \left(1 - \frac{na}{n} \frac{n\bar{b}}{n}\right)^{n-k} = \Pr[\text{card}(X \cap \bar{B}) = k]$$

* *une modélisation hypergéométrique:* on peut le voir rapidement en considérant encore les variables aléatoires $\text{card}(A \cap \bar{Y})$ et $\text{card}(X \cap \bar{B})$ où X et Y possèdent les mêmes propriétés cardinales respectives que A et B . On a, en effet:

$$\begin{aligned} \Pr[\text{card}(A \cap \bar{Y}) = k] &= \frac{C_{na}^k C_{n-na}^{n-nb-k}}{C_n^{n-nb}} = \frac{n_a! n_{\bar{a}}! n_b! n_{\bar{b}}!}{k! n! (n_a - k)! (n_{\bar{b}} - k)! (n_{\bar{b}} - n_a - k)!} \\ &= \frac{C_{n-n_b}^k C_{n_b}^{n_a - k}}{C_n^{n_a}} = \Pr[\text{card}(X \cap \bar{B}) = k] \end{aligned}$$

Si la modélisation binomiale reste compatible avec la sémantique de l'implication, relation binaire non symétrique, il n'en est plus de même pour la modélisation hypergéométrique. Aussi, nous ne retiendrons que le modèle de Poisson et le modèle binomial.

Remarque 2: La quasi-implication, d'indice $q(a, \bar{b})$ non symétrique, ne coïncide pas avec le coefficient de corrélation $\rho(a, b)$ qui est symétrique et qui rend compte de la liaison entre les variables a et b . En effet, nous

montrons que si $q(a, \bar{b}) \neq 0$ alors $\frac{\rho(a, b)}{q(a, \bar{b})} = -\sqrt{\frac{n}{n_b n_{\bar{a}}}}$

Remarque 3: Nous pouvons définir des conjonctions de variables du type “a et b” ou “(a et b) ou c...” afin de modéliser les phénomènes qui relèvent de concepts comme il est fait en apprentissage ou en intelligence artificielle. Les calculs associés restent compatibles avec la logique des propositions reliées par des connecteurs.

Remarque 4: Contrairement à l'indice de Loevinger (Loevinger, 1947) et à la probabilité conditionnelle ($\Pr[B/A]$) et tous ses dérivés, l'intensité d'implication varie avec la dilatation des ensembles E, A et B, résiste aux bruits, c'est-à-dire en particulier au voisinage de 0 pour $n_{a\bar{b}}$, ce qui ne peut que rendre statistiquement crédible la relation que nous voulons modéliser.

Cas des variables modales et fréquentielles

Dans la suite de nos travaux, nous étendons la notion d'implication statistique à des variables autres que binaires. C'est le cas des variables modales qui sont associées à des phénomènes où les valeurs $a(x)$ sont des nombres de l'intervalle $[0, 1]$ et qui décrivent des degrés d'appartenance ou de satisfaction comme en logique floue, par exemple, «peut-être», «un peu», «quelquefois», etc. C'est aussi le cas des variables fréquentielles qui sont associées à des phénomènes où les valeurs de $a(x)$ sont des réels positifs quelconques.

J.B. Lagrange (Lagrange, 1998) a démontré que, dans le cas modal,
 – si $a(x)$ et $\bar{b}(x)$ sont les valeurs prises en x par les variables modales a et \bar{b} , avec $\bar{b}(x) = 1 - b(x)$
 – si s_a^2 et s_b^2 sont les variances empiriques des variables a et \bar{b}
 alors l'indice d'implication, qu'il dénomme *indice de propension*, devient:

DEFINITION 5: $q(a, \bar{b}) = \frac{\sum_{x \in E} a(x)\bar{b}(x) - n_a n_{\bar{b}}}{n \sqrt{\frac{(n^2 s_a^2 + n_a^2)(n^2 s_b^2 + n_b^2)}{n^3}}}$ est l'indice de propension de variables modales.

Il prouve également que cet indice coïncide avec l'indice défini précédemment dans le cas binaire si le nombre de modalités de a et de b est justement 2, car dans ce cas:

$$n^2 s_a^2 + n_a^2 = nn_a, n^2 s_b^2 + n_b^2 = nn_b \text{ et } \sum_{x \in E} a(x)\bar{b}(x) = n_{a \wedge \bar{b}}.$$

Cette solution apportée au cas modal est aussi applicable au cas des *variables fréquentielles*, voire des *variables numériques positives*, à condition d'avoir normalisé les valeurs observées sur les variables, telles que a et b, la normalisation dans [0,1] étant faite à partir du maximum de la valeur prise respectivement par a et b sur l'ensemble E.

Cas des variables-sur-intervalles et variables-intervalles

Variables-sur-intervalles

Situation fondamentale

Deux variables réelles a et b prennent un certain nombre de valeurs sur 2 intervalles finis $\{a_1, a_2\}$ et $\{b_1, b_2\}$. Soit A (resp. B) l'ensemble des valeurs de a (resp. b) observées sur $\{a_1, a_2\}$ (resp. $\{b_1, b_2\}$). Par exemple, a représente les poids d'un ensemble de n sujets et b les tailles de ces mêmes sujets.

Deux problèmes se posent:

1°) peut-on définir des sous-intervalles adjacents de $\{a_1, a_2\}$ (resp. $\{b_1, b_2\}$.) afin que la partition la plus fine obtenue respecte au mieux la distribution des valeurs observées dans $\{a_1, a_2\}$ (resp. $\{b_1, b_2\}$.)?

2°) peut-on trouver les partitions respectives de $\{a_1, a_2\}$ et $\{b_1, b_2\}$ constituées de réunions des sous-intervalles adjacents précédents, partitions qui maximisent l'intensité d'implication moyenne des sous-intervalles de l'un sur des sous-intervalles sur l'autre appartenant à ces partitions?

Nous allons tenter de répondre à ces deux questions dans le cadre de notre problématique en faisant choix des critères à optimiser pour satisfaire l'optimalité attendue dans chaque cas. A la première question, de nombreuses solutions ont été apportées dans d'autres cadres (par exemple, par Lahanier-Reuter, 1998).

Premier problème

On va s'intéresser à l'intervalle $[a_1, a_2]$ en le supposant muni d'une partition initiale triviale de sous-intervalles de même longueur, mais pas nécessairement de même distribution des fréquences observées sur ces sous-intervalles.

Notons $P_0 = \{A_{01}, A_{02}, \dots, A_{0p}\}$, cette partition en p sous-intervalles. On cherche à obtenir une partition P_q^* de $[a_1, a_2]$ en p sous-intervalles $A_{q1}, A_{q2}, \dots, A_{qp}$ de telle façon qu'au sein de chaque sous-intervalle on ait une bonne homogénéité statistique (faible inertie intra-classe) et que ces sous-intervalles présentent une bonne hétérogénéité mutuelle (forte inertie inter-classe). On sait que si l'un des critères est vérifié l'autre l'est nécessairement (théorème de Koenig-Huyghens).

Pour ce faire, on adoptera une méthode directement inspirée de la méthode des nuées dynamiques conçue par Diday (1972) et adaptée à la situation présente. Pour cela, on cherche à minimiser une certaine fonction W , définie sur l'ensemble G des points réels de $[a_1, a_2]$ et l'ensemble des partitions P de $[a_1, a_2]$ en p sous-intervalles A_i , de la façon suivante:

$$W(G, P) = \sum_{i=1}^p D(G_i, A_i) \text{ avec } D(G_i, A_i) = \sum_{x \in A_i} (G_i - x)^2 \text{ pour tout } i=1, 2, \dots, p.$$

Ainsi, si G est le barycentre des valeurs observées dans A , si G_i est le barycentre des valeurs observées dans A_i , alors $W(G, P)$ est l'inertie intra-classe de A et $D(G_i, A_i)$ est l'inertie de A_i .

1^{ère} étape

On part de la partition $P_0 = \{A_{01}, A_{02}, \dots, A_{0p}\}$. On choisit ce que E. Diday appelle les *noyaux*, en nombre p , dans $[a_1, a_2]$. Ces noyaux sont choisis confondus avec les barycentres respectifs, tels que G_{1j} , des sous-intervalles A_{0j} des valeurs qui y sont observées. Soit G_1 leur barycentre.

On cherche alors la partition $P_1 = \{A_{11}, A_{12}, \dots, A_{1p}\}$ telle que pour tout i :

$$A_{1i} = \{x \in A / \forall j (G_{1i} - x)^2 \leq (G_{1j} - x)^2\}$$

Cela revient à constituer A_{1i} à l'aide des points qui sont les plus proches du barycentre de A_{0i} . Cela revient aussi à reporter dans les sous-intervalles voisins les points qui sont plus proches de leurs propres barycentres respectifs. En cas d'égalité, on affecte x au sous-intervalle de plus petit indice.

On démontre alors que $W(G_1, P_1) \leq W(G_1, P_0)$ (voir plus loin)

2^{ème} étape

On dispose de la partition : $P_1 = \{A_{11}, A_{12}, \dots, A_{1p}\}$

On choisit p noyaux G_{2i} qui minimisent respectivement les quantités $D(y, A_{1i})$, c'est-à-dire: pour tout y , valeur observée dans A ,

$$D(G_{2i}, A_{1i}) \leq D(y, A_{1i}) = \sum_{x \in A_{1i}} (y-x)^2, \text{ qui est l'inertie de } A_{1i} \text{ autour de } y$$

à un coefficient près. Par conséquent, le noyau G_{2i} est le barycentre de A_{1i} .

On obtient donc une nouvelle suite de noyaux : $\{G_{21}, G_{22}, \dots, G_{2p}\}$ dont G_2 est le barycentre et l'on est ramené au procédé de l'étape précédente.

On détermine en effet la partition $P_2 = \{A_{21}, A_{22}, \dots, A_{2p}\}$ telle que:

$$A_{2i} = \{x / \forall j (G_{2j} - x)_2 \leq (G_{2i} - x)_2\}$$

On obtient de même: $W(G_2, P_2) \leq W(G_2, P_1)$

3^{ème} étape

On dispose de la partition: $P_{k-1} = \{A_{(k-1)1}, A_{(k-1)2}, \dots, A_{(k-1)p}\}$

On choisit p noyaux G_{ki} qui minimisent respectivement les quantités $D(y, A_{(k-1)i})$, c'est-à-dire tels que:

$$\text{pour tout } y, \text{ valeur observée dans } A, D(G_{2i}, A_{(k-1)i}) \leq D(y, A_{(k-1)i}) = \sum_{x \in A_{(k-1)i}} (y-x)^2$$

Le noyau G_{ki} est le barycentre de $A_{(k-1)i}$. On obtient ainsi une nouvelle suite de noyaux: $\{G_{k1}, G_{k2}, \dots, G_{kp}\}$, dont G_k est le barycentre.

On détermine la partition $P_k = \{A_{k1}, A_{k2}, \dots, A_{kp}\}$ telle que:

$$A_{ki} = \{x / \forall j (G_{ki} - x)^2 \leq (G_{kj} - x)^2\}$$

On obtient encore: $W(G_k, P_k) \leq W(G_k, P_{k-1})$.

Le processus est fini. En effet, la suite $W(G_k, P_k)$ est non croissante. Si elle devient stationnaire, elle a convergé vers son minimum, les noyaux

et les partitions étant inchangées du fait que la somme $\sum_{i=1}^p D(G_i, A_i)$ est positive et constituée d'éléments non croissants.

Exemple

Supposons que nous ayons observé les poids suivants de 17 individus:

$$A = \{ \underbrace{54, 55, 55, 57, 57, 58}_{A_{01}}, \underbrace{60, 66, 67, 67, 68}_{A_{02}}, \underbrace{70, 71, 74, 78, 79, 79}_{A_{03}} \}$$

1^{ère} étape

On choisit

* la partition $P_0 = \{[50, 60[, [60, 70[, [70, 80]\}$

* et les noyaux: $G_{11} = 56$; $G_{12} = 65,25$; $G_{13} = 75,17$

On calcule les valeurs $(G_{11} - x)^2$ pour tous les x de A , puis de la même façon tous les $(G_{12} - x)^2$ et les $(G_{13} - x)^2$. On associe à G_{11} les valeurs de x qui minimisent les expressions $(G_{11} - x)^2$, soit les nombres: 54, 55, 55, 57, 57, 58, 60.

Donc $A_{11} = \{54, 55, 55, 57, 57, 58, 60\}$.

On obtient de même: $A_{12} = \{66, 67, 67, 68, 70\}$, puis $A_{13} = \{71, 74, 78, 79, 79\}$

2^{ème} étape

On choisit pour noyaux les barycentres respectifs des A_{1i} , pour $i=1,2,3$.

Soit $G_{21} = 56,57$; $G_{22} = 67,75$; $G_{23} = 76,2$

On obtient sans peine: $A_{21} = \{54, 55, 55, 57, 57, 58, 60\}$.

$A_{22} = \{66, 67, 67, 68, 70, 71\}$, puis $A_{23} = \{74, 78, 79, 79\}$

3^{ème} étape

On choisit pour noyaux les barycentres respectifs des A_{2i} , pour $i=1,2,3$.

Soit $G_{31} = 56,57$; $G_{32} = 68,4$; $G_{33} = 77,5$

On obtient: $A_{31} = A_{21}$; $A_{32} = A_{22}$; $A_{33} = A_{23}$

Le processus a donc convergé et l'algorithme s'arrête sur la dernière partition.

Remarque

Afin d'évaluer la qualité de la partition obtenue, on calcule le rapport:

$$\tau = \frac{\text{Inertie inter-classe}}{\text{Inertie totale}}$$

Or l'inertie inter-classe de A pour la partition $P_3 = \{A_{31}, A_{32}, A_{33}\}$ est:

$$\sum_{i=1,2,3} m(G_{3i}) (G_{3i} - G)^2 = 47,433 \text{ où } m(G_{3i}) \text{ est l'effectif des } x \text{ dans } A_{3i}$$

et G est le barycentre de A.

L'inertie totale est $\sum_{x \in A} (G - x)^2 = 49,82$, soit le taux τ excellent de 0,95.

Deuxième problème¹

On suppose maintenant que les intervalles $[a_1, a_2]$ et $[b_1, b_2]$ sont munis de partitions optimales P et Q, respectivement, au sens des nuées dynamiques. Soit p et q les nombres respectifs de sous-intervalles composant P et Q. A partir de ces deux partitions, il est possible d'engendrer 2^{p-1} et 2^{q-1} partitions obtenues par réunions itérées de sous-intervalles adjacents respectivement de P et de Q².

1 D. Lahanier-Reuter en propose une autre approche dans sa thèse (cf infra).

2 Il suffit de considérer l'arborescence dont A_1 est la racine, puis de le réunir ou non à A_2 qui lui-même sera ou non réuni à A_3 , etc. Il y a donc 2^{p-1} branches dans cette arborescence.

On calcule les intensités d'implication respectives de chaque sous-intervalle réuni ou non à un autre de la première partition sur chaque sous-intervalle réuni ou non à un autre de la seconde, puis les valeurs des intensités des implications réciproques.

Il y a donc au total $2 \cdot 2^{p-1} \cdot 2^{q-1}$ familles d'intensités d'implication, chacune d'entre elles nécessitant le calcul de tous les éléments d'une partition de $[a_1, a_2]$ sur tous les éléments d'une des partitions de $[b_1, b_2]$ et réciproquement.

On choisit comme *critère d'optimalité* la moyenne géométrique des intensités d'implication, moyenne associée à chaque couple de partitions d'éléments réunis ou non définies inductivement. On note les deux maxima obtenus (implication directe et sa réciproque) et on retient les deux partitions associées en déclarant que l'implication de la variable-sur-intervalle a sur la variable-sur-intervalle b est optimale lorsque l'intervalle $[a_1, a_2]$ admet la partition correspondant au premier maximum et que l'implication réciproque optimale est satisfaite pour la partition de $[b_1, b_2]$ correspondant au deuxième maximum.

Remarque

1°) Il n'existe pas de relation d'ordre entre $\varphi(a,b)$ et $\varphi(a,b \vee c)$ connaissant $\varphi(a,b)$ et $\varphi(a,c)$.

En effet, on peut avoir $\varphi(a,b) < \varphi(a,b \vee c)$ et $\varphi(a,c) < \varphi(a,b \vee c)$.

exemple 1: $n = 100$; $n_a = 16$; $n_b = 35$; $n_c = 30$; $n_{a \wedge b} = 10$; $n_{a \wedge c} = 8$;

$$n_{a \wedge (b \vee c)} = 2$$

Mais on peut également avoir: $\varphi(a,b) > \varphi(a,b \vee c)$ et $\varphi(a,c) > \varphi(a,b \vee c)$.

exemple 2: $n = 100$; $n_a = 30$; $n_b = 50$; $n_c = 49$; $n_{a \wedge b} = 15$; $n_{a \wedge c} = 16$;

$$n_{a \wedge (b \vee c)} = 1$$

2°) De même, il n'existe pas a priori de relation entre $\varphi(a,c)$, $\varphi(b,c)$ et $\varphi(a \vee b, c)$

En effet, on peut avoir, $\varphi(a,c) < \varphi(a \vee b, c)$ et $\varphi(b,c) < \varphi(a \vee b, c)$.

exemple 3: $n = 100$; $n_a = 20 = n_b$; $n_c = 35$; $n_{a \wedge c} = 16$; $n_{b \wedge c}$;

$$n_{a \vee b, c} = 10$$

Mais on peut avoir également: $\varphi(a,c) > \varphi(a \vee b, c)$ et $\varphi(b,c) > \varphi(a \vee b, c)$.

exemple 4: $n = 100$; $n_a = 20$; n_b ; $n_c = 48$; $n_{a \wedge c} = 11 = n_{b \wedge c}$; $n_{a \vee b, c} = 22$

Ceci montre que l'algorithme de recherche de l'optimum des implications au cours des réunions successives doit fonctionner "jusqu'au bout", c'est-à-dire lorsque toutes les réunions ont été produites et estimées quant à leur puissance implicative.

Décroissance de la fonction W

Quelques notations

Soit A l'ensemble des valeurs observées dans l'intervalle $[a_1, a_2]$ et $L = \{N_1, N_2, \dots, N_p\}$ un ensemble de parties de A . N_i est appelé $i^{\text{ème}}$ noyau de L . Ces noyaux sont choisis de telle façon que $\text{card } N_i$ soit le même pour tout i .

Soit $P = \{A_1, A_2, \dots, A_p\}$ une partition de A en p classes.

$$\text{On pose } W(L, P) = \sum_{i=1}^p D(N_i, A_i) = \sum_{i=1}^p \sum_{x \in A_i, y \in N_i} d(x, y)^3$$

$D(N_i, A_i)$ est une sorte de mesure de dissemblance entre le noyau N_i et la classe A_i .

Le problème des nuées dynamiques vise à minimiser $W(L, P)$ par la construction d'un ensemble convenable de p noyaux dans L^* et d'une partition P^* en p classes.

Algorithme des nuées dynamiques

Les noyaux $L_0 = \{N_{01}, N_{02}, \dots, N_{0p}\}$ sont donnés (ou choisis arbitrairement).

On définit la partition $P_1 = \{A_{11}, A_{12}, \dots, A_{1p}\}$ qui s'en déduit par:

$$A_{1i} = \{x \in A / \forall j \ d(N_{0j}, x) \leq d(N_{0i}, x)\}.$$

3 $d(x, y)$ peut être égal à $(x-y)^2$ comme nous l'avons choisi dans le premier problème.

En cas d'égalité, on affecte x à la classe d'indice plus petit.

On pose $P_1 = f(L_0)$.

On construit alors les noyaux de $L_1 = \{N_{11}, N_{12}, \dots, N_{1p}\}$ par le procédé:

$$N_{1k} = \{x \in A / d(x, A_{1k}) = \inf_j d(x, A_{1j})\}$$

On note $L_1 = g(P_1)$ et on itère l'algorithme.

Proposition: $W(L, P)$ décroît à chaque itération, i.e.:

a) $\forall P, W(L, f(L)) \leq W(L, P)$

b) $\forall L, W(g(P), P) \leq W(L, P)$

Preuve:

a) Soit la partition quelconque au cours de l'algorithme $P = \{A_1, A_2, \dots, A_p\}$ et soit $f(L) = Q = \{B_1, B_2, \dots, B_p\}$ la partition obtenue à l'aide des noyaux $L = \{N_1, N_2, \dots, N_p\}$ définis à partir de P .

$$\text{Alors } W(L, P) = \sum_{i=1}^p D(N_i, A_i) = \sum_{i=1}^p \sum_{x \in A_i} d(N_i, x) = \sum_{j=1}^p \sum_{x \in N_j} d(x, A_j)$$

$$\text{et } W(L, Q) = \sum_{j=1}^p \sum_{y \in B_j} d(N_j, y)$$

Soit $x \in A$. Pour tout i et tout x tel que $x \in A_i \cap B_i$, alors x a la même contribution aux sommes $W(L, P)$ et $W(L, Q)$. Par contre, si $x \in A_i$ et $x \in B_j$, alors $d(N_j, x) \leq d(N_i, x)$

car par construction: $B_j = \{x \in A / \forall i, d(N_j, x) \leq d(N_i, x)\}$

Par suite, pour tout $x \in A$, les contributions de x à $W(L, Q)$ sont inférieures ou égales à celles de $W(L, P)$, soit encore $\forall P, W(L, f(L)) \leq W(L, P)$.

b) Soit $g(P) = \{O_1, O_2, \dots, O_p\}$ un ensemble de p noyaux obtenus à la suite de la partition $P = \{A_1, A_2, \dots, A_p\}$, elle-même dérivée de $L = \{N_1, N_2, \dots, N_p\}$ et définis ainsi:

$$\forall k, O_k = \{x \in A / d(x, A_k) = \inf_j d(x, A_j)\}.$$

$$\text{Par suite, } \forall j, \sum_{x \in O_j} d(x, A_j) \leq \sum_{x \in N_j} d(x, A_j) \text{ d'où } \forall L, W(g(P), P) \leq W(L, P)$$

Variables-Intervalles

Situation fondamentale

On dispose des données fournies par une population de n individus (qui peuvent être chacun ou certains des ensembles d'individus, par ex. une classe d'élèves) selon p variables (par ex. notes sur une année en français, math, physique, ... mais aussi bien: poids, tailles, tour de poitrine, ...). Les valeurs prises par ces variables selon chaque individu sont des intervalles de réels positifs. Par ex., l'individu x donne la valeur $\{12 ; 15.50\}$ à la variable note de math. E. Diday parlerait à ce sujet de p variables symboliques à valeurs intervalles définies sur la population.

On cherche à définir une implication d'intervalles, relatifs à une variable a , constitués eux-mêmes des intervalles observés, vers d'autres intervalles pareillement définis et relatifs à une autre variable b . Ceci permettra de mesurer l'association implicative, donc non symétrique, de certain(s) intervalle(s) de la variable a avec certain(s) intervalle(s) de la variable b , ainsi que l'association réciproque à partir de laquelle on retiendra la meilleure pour chaque couple de sous-intervalles en jeu, comme il vient d'être fait au §1.

Par exemple, on dira que le sous-intervalle $\{2 ; 5,5\}$ de notes de mathématiques implique généralement le sous-intervalle $\{4,25 ; 7,5\}$ de notes de physique, ces deux sous-intervalles appartenant à une partition optimale au sens de la variance expliquée des intervalles respectifs de valeurs $\{1 ; 18\}$ et $\{3 ; 20\}$ prises dans la population. De même, on dira que $\{14,25 ; 17,80\}$ en physique implique le plus souvent $\{16,40 ; 18\}$ en mathématiques.

En suivant la problématique de E. Diday et ses collaborateurs, si les valeurs prises selon les sujets par les variables a et b sont de nature

symbolique, en l'occurrence des intervalles de R^+ , il est possible d'étendre les algorithmes ci-dessus [Gras R. 17 c, 2001, c]. Par exemple, à la variable a sont associés des intervalles de poids et à la variable b des intervalles de tailles, intervalles dus à une imprécision des mesures. Effectuant la réunion des intervalles I_x et J_x décrits par les sujets x de E selon respectivement chacune des variables a et b, on obtient deux intervalles I et J recouvrant toutes les valeurs possibles de a et de b. Sur chacun d'eux on peut définir une partition en un certain nombre d'intervalles respectant comme plus haut un certain critère d'optimalité. Pour cela, les intersections des intervalles tels que I_x et J_x avec ces partitions seront munies d'une distribution prenant en compte les étendues des parties communes. Cette distribution peut être uniforme ou d'un autre type discret ou continu. Mais ainsi, nous sommes ramenés à la recherche de règles entre deux ensembles de variables-sur-intervalles qui prennent comme précédemment leurs valeurs sur $[0, 1]$ et sur lesquelles on pourra chercher les implications optimales.

L'implication-inclusion

Deux raisons nous ont conduits à améliorer le modèle réalisé par l'intensité d'implication:

– lorsque la taille des échantillons traités, et en particulier celui de E, croît (de l'ordre du millier et plus), l'intensité $\varphi(a,b)$ a tendance à ne plus être suffisamment discriminante car ses valeurs peuvent être très voisines de 1, alors que l'inclusion dont elle cherche à modéliser la qualité, est loin d'être satisfaite (phénomène signalé dans (Bodin, 1997) qui traite des grandes populations d'élèves à travers des enquêtes internationales);

– le modèle de la quasi-implication précédent retient essentiellement la mesure de la force de la règle $a \Rightarrow b$. Or, la prise en compte d'une puissance concomitante de non $b \Rightarrow$ non a (contraposée de l'implication) est indispensable pour renforcer l'affirmation d'une bonne qualité de la relation quasi-implicative de a sur b⁴. En même temps, elle pourrait

4 Ce phénomène est signalé par Y. Kodratoff dans son article publié dans les Actes du Colloque "Fouille dans les données par la méthode implicative", IUFM de Caen, juin 2000.

permettre de corriger la difficulté évoquée ci-dessus (si A et B sont petits par rapport à E, leurs complémentaires seront importants et réciproquement).

La solution que nous apportons utilise à la fois l'intensité d'implication et un autre indice qui rend compte de la dissymétrie entre les situations $S_1 = (a \text{ et } b)$ et $S'_1 = (a \text{ et non } b)$ (resp. $S_2 = \text{non } a \text{ et non } b$) et $S'_2 = (a \text{ et non } b)$) en faveur de la première nommée. La faiblesse relative des instances qui contredisent la règle et sa contraposée est ainsi fondamentale. D'ailleurs, le nombre de contre-exemples $n_{a\bar{b}}$ à $a \Rightarrow b$ est celui à la contraposée. C'est donc au concept d'entropie de Shannon que nous faisons référence:

$$H(b/a) = -\frac{n_{a\wedge b}}{n_a} \log_2 \frac{n_{a\wedge b}}{n_a} - \frac{n_{a\wedge \bar{b}}}{n_a} \log_2 \frac{n_{a\wedge \bar{b}}}{n_a},$$

entropie conditionnelle relative aux cases (a et b) et (a et non b) lorsque a est réalisée

$$H(\bar{a}/\bar{b}) = -\frac{n_{\bar{a}\wedge \bar{b}}}{n_{\bar{b}}} \log_2 \frac{n_{\bar{a}\wedge \bar{b}}}{n_{\bar{b}}} - \frac{n_{\bar{a}\wedge b}}{n_{\bar{b}}} \log_2 \frac{n_{\bar{a}\wedge b}}{n_{\bar{b}}}$$

entropie conditionnelle relative aux cases (non a et non b) et (a et non b) lorsque non b est réalisée.

Ces entropies, à valeurs dans $[0,1]$, devraient donc être simultanément faibles et donc les dissymétries entre les situations S_1 et S'_1 (resp. S_2 et S'_2) devraient être simultanément fortes si l'on souhaite disposer d'un bon critère d'inclusion de A dans B. En effet, les entropies représentent l'incertitude moyenne des expériences qui consistent à observer si b est réalisé (resp. si non a est réalisé) lorsque l'on a observé a (resp. non b). Le complément à 1 de cette incertitude représente donc l'information moyenne recueillie par la réalisation de ces expériences. Plus cette information est importante, plus forte est la garantie de la qualité de l'implication et de sa contraposée. Nous devons maintenant adapter ce critère numérique entropique au modèle attendu dans les différentes situations cardinales.

Pour que le modèle ait la signification attendue, il doit satisfaire, selon nous, les contraintes suivantes:

1°) il devra intégrer les valeurs de l'entropie et, pour les contraster, par exemple, intégrer ces valeurs au carré,

2°) comme ce carré varie de 0 à 1, afin de dénoter le déséquilibre et donc l'inclusion, afin de s'opposer à l'entropie, la valeur retenue sera le complément à 1 de son carré tant que le nombre de contre-exemples sera inférieur à la moitié des observations de a (resp. de non b). Au delà de ces valeurs, les implications n'ayant plus de sens inclusif, on affectera au critère la valeur 0,

3°) afin de prendre en compte les deux informations propres à $a \Rightarrow b$ et $\text{non } b \Rightarrow \text{non } a$, le produit rendra compte de la qualité simultanée des valeurs retenues. Le produit a la propriété de s'annuler dès que l'un de ses termes s'annule, i.e. dès que cette qualité s'efface,

4°) enfin, le produit ayant une dimension 4 par rapport à l'entropie, sa racine quatrième sera de la même dimension.

Posons $\alpha = \frac{n_a}{n}$, la fréquence de a et $\bar{\beta} = \frac{n_{\bar{b}}}{n}$, la fréquence de non b. Notons, en fonction de la fréquence $t = \frac{n_{a \wedge \bar{b}}}{n}$ de contre-exemples, les deux termes significatifs des qualités respectives de l'implication et sa contraposée:

$$h_1(t) = H(b/a) = -(1 - \frac{t}{\alpha}) \log_2(1 - \frac{t}{\alpha}) - \frac{t}{\alpha} \log_2 \frac{t}{\alpha} \quad \text{si } t \in [0, \frac{\alpha}{2}] \quad [\text{et } h_1(t) = 1 \text{ si } t \in [\frac{\alpha}{2}, \alpha]$$

$$h_2(t) = H(\bar{a}/\bar{b}) = -(1 - \frac{t}{\bar{\beta}}) \log_2(1 - \frac{t}{\bar{\beta}}) - \frac{t}{\bar{\beta}} \log_2 \frac{t}{\bar{\beta}} \quad \text{si } t \in [0, \frac{\bar{\beta}}{2}] \quad [\text{et } h_2(t) = 1 \text{ si } t \in [\frac{\bar{\beta}}{2}, \bar{\beta}]$$

D'où la définition permettant de déterminer le critère entropique:

DEFINITION 6: L'indice d'inclusion de A, support de a, dans B, support de b, est le nombre:

$$i(a,b) = [(1-h_1^2(t))(1-h_2^2(t))]^{1/4}$$

qui intègre l'information délivrée par la réalisation d'un faible nombre de contre-exemples, d'une part à la règle $a \Rightarrow b$ et, d'autre part, à la règle

non b \Rightarrow non a L'intensité d'implication-inclusion (ou intensité entropique) est le nombre:

$$\psi(a,b) = (i(a,b) \cdot \Psi(a,b))^{1/2}$$

qui intègre à la fois l'étonnement statistique et la qualité inclusive.

La fonction ψ suivant la variable t admet une représentation qui a la forme indiquée par la figure 2, pour n_a et n_b fixés. On remarquera la différence de comportement de la fonction par rapport à la probabilité conditionnelle $P(B/A)$, indice fondamental des autres modélisations de la mesure des règles, par exemple chez Agrawal et son école. Outre son caractère linéaire, donc peu nuancé, cette probabilité conduit à une mesure qui décroît trop vite dès les premiers contre-exemples et résiste ensuite trop longtemps lorsque ceux-ci deviennent importants.

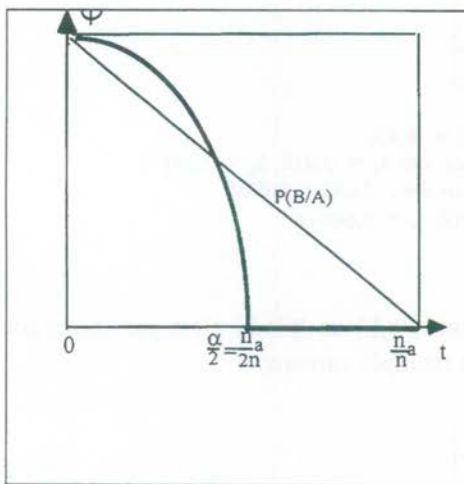


Figure 2

On constate que cette représentation de fonction continue de t traduit les propriétés attendues du critère d'inclusion:

- * "réaction" lente aux premiers contre-exemples (résistance au bruit),
- * "accélération" du rejet de l'inclusion au voisinage de l'équilibre soit $\frac{n_a}{2n}$,
- * rejet au-delà de $\frac{n_a}{2n}$ ce que n'assurait pas l'intensité d'implication.

Exemple 1

	b	\bar{b}	marge
a	200	400	600
\bar{a}	600	2800	3400
marge	800	3200	4000

L'intensité d'implication est 0,9999 ($q(a, \bar{b}) = -3,65$)

Les valeurs entropiques de l'expérience sont $h_1 = 0 = h_2$

La valeur du coefficient modérateur est donc: $i(a, b) = 0$

Par suite $\psi(a, b) = 0$ alors que $P(B/A) = 0,33333$

Ainsi, les fonctions "entropiques" "modèrent" l'intensité d'implication dans ce cas où justement l'inclusion est médiocre.

Exemple 2

	b	\bar{b}	marge
a	400	200	600
\bar{a}	1000	2400	3400
marge	1400	2600	4000

L'intensité d'implication est 1 ($q(a, \bar{b}) = -8,43$)

Les valeurs entropiques de l'expérience sont $h_1 = 0,918$; $h_2 = 0,391$

La valeur du coefficient modérateur est donc: $i(a, b) = 0,6035$

Par suite $\psi(a, b) = 0,777$ alors que $P(B/A) = 0,66666$

Remarque

La correspondance entre $\varphi(a, b)$ et $\psi(a, b)$ n'est pas monotone comme le montre le deuxième exemple suivant:

	b	\bar{b}	marge
a	40	20	60
\bar{a}	60	280	340
marge	100	300	400

L'intensité d'implication est inférieure à la précédente car: ($q(a, \bar{b}) = -6,47$).

Les valeurs entropiques sont: $h_1 = 0,918$; $h_2 = 0,353$

La valeur du coefficient modérateur est: $i(a, b) = 0,608$

Donc $\psi(a, b) = 0,78$ alors que $P(B/A) = 0,66666$

Ainsi, alors que $\varphi(a,b)$ a décru du 1er au 2ème exemples, $i(a,b)$ a crû de même que $\psi(a,b)$. En revanche, la situation contraire est la plus fréquente. Notons que, dans les deux cas, la probabilité conditionnelle ne change pas.

Graphe d'implication

La relation définie par l'implication statistique, si elle est réflexive et non symétrique, n'est pas transitive bien évidemment. Or nous voulons qu'elle modélise la relation d'ordre partiel entre deux variables (les réussites dans notre exemple initial). Par convention, si $a \Rightarrow b$ et si $b \Rightarrow c$, nous accepterons la fermeture transitive $a \Rightarrow c$ seulement si $\psi(a,c) \geq 0,5$, c'est-à-dire si la relation implicative de a sur c est meilleure que la neutralité.

Par exemple, supposons qu'entre les 7 variables a, b, c, d, e et f existent, au seuil supérieur à 0,5, les relations suivantes: $e \Rightarrow c, a, f, b;$ $c \Rightarrow a, f, b \Rightarrow a, f, g \Rightarrow d, f, a \Rightarrow f$.

On pourra alors traduire cet ensemble de relations par le graphe suivant⁵:

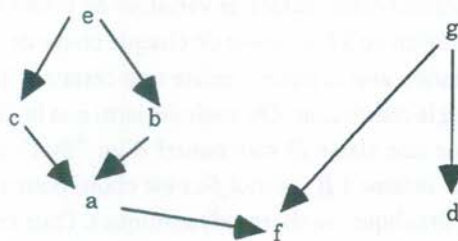


Figure 3

5 Les traitements automatiques des calculs et des graphiques sont exécutés à l'aide du logiciel C.H.I.C. (Classification Hiérarchique Implicative et Cohésitive) disponible sous Windows 95. Ce logiciel, à partir d'une première version établie par R. Gras, révisée sous Pascal par S. Ag Almouloud, [Ag Almouloud, 1992] est maintenant développé par R. Couturier et constamment étendu par lui aux nouveaux concepts et nouveaux algorithmes.

Notons que ce graphe n'est pas un treillis puisque, par exemple la variable a n'implique pas la variable $(a \text{ ou non } a)$ dont le support est E . A fortiori, ce ne peut être un treillis de Galois.

Remarque

A l'instar de la définition 6 d'un indice d'implication-inclusion, A. Bodin (Bodin, 1997) propose un test d'hypothèse afin de valider ou réfuter une hypothèse d'inclusion de l'ensemble A dans l'ensemble B . Ce test est intégré au menu du logiciel C.H.I.C. et conduit au graphe dit d'implication-inclusion établi sur la base d'un seuil donné par l'utilisateur. On peut noter quelques divergences entre ce graphe et le graphe classique d'implication, en particulier, comme l'avait dénoncé A. Bodin, lorsque les populations en jeu deviennent importantes. Par contre, les divergences sont plus faibles lorsque le graphe d'implication est construit à l'aide de l'"intensité entropique" que nous avons construite dans le paragraphe précédent.

Implication entre règles et méta-règles

Une implication entre classes de variables ne prend véritablement son sens qu'à condition qu'à l'intérieur de chaque classe de variables dont on examine la relation avec d'autres, existe une certaine "cohésion" entre les variables qui la constituent. On souhaite ainsi que le "flux" implicatif d'une classe A sur une classe B soit nourri d'un "flux" interne à A et alimente un "flux" interne à B (ce mot *flux* est choisi pour sa connotation métaphorique hydraulique ou thermodynamique). Pour cela, le concept d'entropie H permettant de rendre compte du désordre entre des variables, nous définissons la cohésion entre deux variables par:

DEFINITION 7: La cohésion de la classe (a,b) est le nombre $c(a,b)$ tel que:

$$\begin{aligned} & \text{. si } p = \psi(a,b) \text{ et } H = -p \log_2 p - (1-p) \log_2 (1-p), \text{ alors } \text{coh}(a,b) = \\ & \sqrt{1 - H^2} \end{aligned}$$

$$\text{. si } p = 1, \text{ alors } \text{coh}(a,b) = 1$$

$$\text{. si } p \leq 0,5, \text{ alors } \text{coh}(a,b) = 0$$

Intuitivement, la cohésion mesure le déséquilibre des occurrences des événements $a \wedge b$ et $a \wedge \bar{b}$ en faveur du premier.

DEFINITION 8: La cohésion de la classe de variables $\mathcal{A} = (a_1, \dots, a_r)$ est alors définie par extension:

$$\text{Coh}(\mathcal{A}) = \left[\prod_{\substack{i \in \{1, \dots, r-1\} \\ j \in \{2, \dots, r\}, j > i}} c(a_i, a_j) \right]^{\frac{2}{r(r-1)}}$$

C'est la moyenne géométrique des cohésions de classes à 2 éléments qui, en tant que telle, s'annule dès que l'une des cohésions en jeu s'annule.

Enfin nous pouvons modéliser l'implication statistique d'une classe de variables sur une autre classe en exigeant du modèle qu'il intègre les informations suivantes:

- les cohésions respectives des 2 classes,
- une intensité d'implication extrême des éléments d'une classe sur les éléments de l'autre,
- les cardinaux respectifs des 2 classes.

Chacune de ces informations crédite l'indice que nous retiendrons si:

- l'indice croît avec les cohésions de chaque classe et s'annule lorsque la cohésion de l'une d'entre elles est nulle,
- l'indice croît avec la liaison extrême (minimale si l'on vise un degré d'exigence élevé, maximale si l'on recherche une souplesse réaliste),
- l'indice décroît avec les cardinaux des classes, eu égard à la prise en compte d'une liaison maximale.

Par suite, notant \mathcal{A} et \mathcal{B} des classes de variables d'éléments génériques a_i et b_j , puis $\text{Coh}(\mathcal{A})$ et $\text{Coh}(\mathcal{B})$ leurs cohésions respectives, l'intensité d'implication de \mathcal{A} sur \mathcal{B} est donnée par:

DEFINITION 9: L'intensité d'implication de \mathcal{A} sur \mathcal{B} est:

$$\psi(\mathcal{A}, \mathcal{B}) = \left[\sup_{\substack{i \in \{1, \dots, r\} \\ j \in \{1, \dots, s\}}} \psi(a_i, b_j) \right]^{\text{rs}} [\text{Coh}(\mathcal{A}) \cdot \text{Coh}(\mathcal{B})]^{1/2}$$

On pourra constater que cet indice satisfait les contraintes sémantiques déclarées ci-dessus. Définissant à partir de cet indice une méthode de classification descendante classique par un critère de cohésion décroissante, qui définit une distance ultramétrique, on obtiendra par exemple des arbres comme celui-ci :

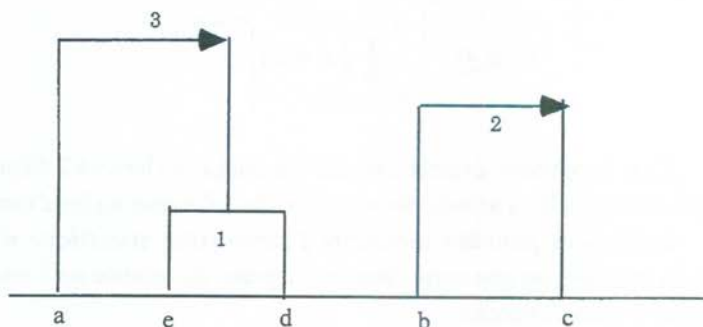


Figure 4

Des interprétations de telles méta-règles sont plus complexes, comme par exemple, illustrée par la figure 4, la méta-règle $a \Rightarrow (e \Rightarrow d)$. Mais le mathématicien pourra faire la comparaison entre de tels assemblages et le fonctionnement interne d'une théorie mathématique: une propriété ou un axiome entraîne un théorème, un théorème entraîne un autre théorème, etc.. De plus, quelques méta-règles sont réductibles à des assemblages aisément interprétables. Par exemple, $a \Rightarrow (e \Rightarrow d)$ se ramène logiquement à $a \wedge e \Rightarrow d$.

Niveaux significatifs d'une hiérarchie cohesive

Etant donné la multiplicité des niveaux de formation des classes, il est indispensable de dégager ceux qui sont les plus pertinents par rapport à l'intention classificatrice du chercheur et eu égard aux critères choisis. Nous procédons alors de façon comparable à celle adoptée primitivement par I.C.Lerman (Lerman, 1981) et relativement à la hiérarchie de similarité, mais en reconditionnant son approche.

Préordre cohesitif

Considérons l'ensemble V des variables $\{a_1, a_2, \dots, a_m\}$ et l'ensemble des couples (a,b) de $V \times V$ tels que $a \neq b$. Il existe $m(m-1)$ tels couples auxquels on a associé leurs cohésions $c(a,b)$ respectives.

DEFINITION 10: On appelle préordre initial et global cohesitif sur $V \times V$ (ou préordonnance), le préordre Ω induit par l'application cohésion c sur $V \times V$.

Soit $G(\Omega)$ son graphe dans $V \times V$. D'après les deux premiers paragraphes, il s'ensuit que:

* d'une part, la classe de préordre correspondant à $c=0$ contient tous les couples tels que $\psi(a,b) \leq 0,5$,

* d'autre part, si $n_a \leq n_b$ alors $c(b,a) \leq c(a,b)$.

Remarquons, par contre, que si $c(a,b) \leq c(c,d)$ on n'a pas nécessairement $c(b,a) \leq c(d,c)$ ou $c(b,a) \geq c(d,c)$.

Determination des niveaux significatifs

Plaçons-nous à un niveau quelconque k de la hiérarchie. A ce niveau, se forme une classe de m_k variables ($2 \leq m_k \leq m$) dont la cohésion est moins bonne que celle des classes antérieurement formées, conformément à l'algorithme retenu, et meilleure que celles des classes à venir.

Soit Π_k la partition sur V définie à ce niveau constituée des classes qui y sont déjà formées et, éventuellement, des singletons non encore associés. Π_k est plus fine que Π_{k+1} .

Soit S_{Π_k} l'ensemble des couples séparés à ce niveau et R_{Π_k} l'ensemble des couples qui y sont réunis pour la première fois, étant entendu que l'on dira que le couple (a,b) est réuni si a et b appartiennent à la même classe du type $(\dots(\dots a, \dots) \dots b) \dots$.

L'ensemble $G(\Omega) \llbracket S_{\Pi_k} \times R_{\Pi_k} \rrbracket$ est constitué des couples de couples qui au niveau k respectent le préordre initial. Par exemple, si l'on a $c(e,f) < c(a,b)$ (donc $((e,f), (a,b)) \in G(\Omega)$) et si au niveau k , e et f sont séparés alors que a et b se réunissent dans la classe qui se forme, le couple $((e,f), (a,b))$ appartient à $G(\Omega) \llbracket S_{\Pi_k} \times R_{\Pi_k} \rrbracket$.

Comme il a été fait dans le premier paragraphe pour le cardinal de $A \ll \bar{B}$, nous associons au cardinal de $G(\Omega) \ll [S_{\Pi_k} \times R_{\Pi_k}]$ l'indice aléatoire $\text{card}[G(\Omega^*) \ll [S_{\Pi_k} \times R_{\Pi_k}]]$ où Ω^* est une préordonnance aléatoire dans l'ensemble, muni d'une probabilité uniforme, de toutes les préordonnances de même type cardinal que Ω . Cet indice a pour espérance $1/2 \text{card}[S_{\Pi_k} \times R_{\Pi_k}]$ et pour variance $\text{card}[S_{\Pi_k} \times R_{\Pi_k}] \text{card}[G(\Omega)]$.

Soit $s(\Omega, k)$ l'indice centré réduit obtenu :

$$\frac{(\text{card}[G(\Omega^*) \ll [S_{\Pi_k} \times R_{\Pi_k}]] - 1/2 \text{card}[S_{\Pi_k} \times R_{\Pi_k}])}{(\text{card}[S_{\Pi_k} \times R_{\Pi_k}] \text{card}[G(\Omega)])^{1/2}}$$

DEFINITION 11: On appelle noeud significatif tout noeud correspondant à un maximum local de $s(\Omega, k)$ au cours de la constitution de la hiérarchie implicative. Nous dirons dans ce cas que la partition Π_k est en résonance partielle avec Ω .

Si, de plus, $G(\Omega) \ll [S_{\Pi_k} \times R_{\Pi_k}] = S_{\Pi_k} \times R_{\Pi_k}$, nous dirons que la partition Π_k est en résonance totale avec Ω .

Le logiciel d'analyse de données C.H.I.C. permet le traitement complet de données quantitatives, ainsi que la sortie du graphe d'implication et de la hiérarchie implicative en mentionnant les noeuds significatifs.

Typicalité et contribution des sujets et des variables supplémentaires

Nous introduisons la notion de variable supplémentaire en analyse implicative à l'instar de la même notion définie en analyse factorielle, c'est-à-dire variable extrinsèque, descripteur par exemple, n'intervenant pas directement dans les liaisons exprimées par la classification entre les variables dites principales de V , donc n'intervenant pas dans la structure de cet ensemble sous la forme graphe ou hiérarchie. Par exemple, une variable supplémentaire pourra représenter une catégorie de sujets (âge, sexe, catégorie socio-professionnelle, etc.).

Il nous paraît très important, au moment du choix des variables sur lesquelles portera l'analyse, de bien distinguer les variables qui vont participer à la mise en évidence de règles ou de méta-règles, variables que nous appelons *principales*, des autres variables, dites *supplémentaires* qui sont descripteurs des sujets et qui de ce fait sont objectives, non liées aux comportements des sujets. Les introduire dans l'analyse risque fort, comme nous l'avons observé, de créer des artefacts dans le graphe ou la hiérarchie, car elles y jouent alors un rôle attracteur trop puissant pour laisser les autres variables subjectives s'organiser entre elles. Comme on le sait, les processus de chaînage inhérents aux méthodes de classification, perturberont les raisons véritables qui sont à la base des relations à extraire. En revanche, identifier les phénomènes objectifs qui jouent un rôle dans les structures apparues est d'une grande richesse informative.

Intéressons-nous pour l'instant à l'analyse cohésitive. A un niveau quelconque de la hiérarchie se forme une classe C de cohésion non nulle. Notre objectif, particulièrement dans le cas d'un noeud significatif, est de définir un critère permettant d'identifier un ou des sujets, puis la catégorie de sujets, ou tout autre variable supplémentaire (âge, sexe, catégorie socio-professionnelle, etc.), contribuant le plus à la constitution de cette classe. Le comportement de ces sujets sera ainsi en harmonie avec le comportement statistique à l'origine de la classe. Une approche comparable est faite conjointement pour étudier la typicalité et la contribution des sujets et des variables supplémentaires à la constitution d'un arc ou d'un chemin du graphe implicatif.

Puissance implicative de classe et de chemin

Plaçons-nous à un niveau k de la hiérarchie où viennent de se réunir, pour former C , deux classes A et B telles que $A \Rightarrow B$ au sens du paragraphe "Implication entre règles et méta-règles"

DEFINITION 12: Le couple (a,b) tel que: $\forall i \in A, \forall j \in B \psi(a,b) \geq \psi(i,j)$ est appelé couple générique de C . C'est ce couple, généralement unique, qui intervient par le sup. dans le calcul de l'implication de A sur B . Le nombre $\psi(a,b)$ est appelé implication générique de C .

Mais, dans chaque sous-classe de C , existe également un couple générique. Précisément, si C est constituée de g ($g \leq k$) sous-classes (C comprise), il y a g couples génériques à l'origine de C et g intensités maximales d'implication $\psi_1, \psi_2, \dots, \psi_g$, qui leur correspondent.

Dans le cas d'un chemin C fermé transitivement (chaque arc de la fermeture admet une intensité d'implication au moins égale à 0.50), composé de g nœuds, C présente $g(g-1)/2$ arcs transitifs. A chacun de ces arcs, par exemple (a,b) , on associe, comme pour une classe, l'intensité d'implication $\psi(a,b)$, que l'on dira encore générique.

DEFINITION 13: Le vecteur $\psi_1, \psi_2, \dots, \psi_g$, élément de $[0,1]^g$, est appelé vecteur puissance implicative de C , traduisant une force implicative interne à C .

Puissance Implicative d'un sujet sur une classe ou un chemin et distance a cette classe ou a ce chemin

Un sujet x quelconque respecte ou non l'implication du couple générique d'une classe ou d'un arc de chemin avec un ordre de qualité comparable. Associant logique formelle et considération sémantique, nous poserons, par exemple et en fonction des valeurs prises par a et b en x :

$\psi_x(a,\bar{b})=1$ si $a=1$ ou 0 et $b=1$; $\psi_x(a,\bar{b})=0$ si $a=1$ et $b=0$; $\psi_x(a,\bar{b})=p$ si $a=b=0$ avec $p \in]0,1]$. Dans nos premières expériences, nous choisissons $p=.5$, valeur neutre. Dans le logiciel CHIC, le calcul des typicalités et des contributions se fait cependant en modulant ces valeurs afin de mieux prendre en compte la sémantique des valeurs attribuées par x à a et à b .

Ainsi, à x , nous pouvons associer n nombres $\psi_{x,1}, \psi_{x,2}, \dots, \psi_{x,g}$ correspondant aux valeurs prises en x par les g implications génériques de la classe ou du chemin C .

DEFINITION 14: Le vecteur $(\psi_{x,1}, \psi_{x,2}, \dots, \psi_{x,g})$, élément de $[0,1]^g$, est appelé vecteur puissance implicative de x . Le sujet x_t , peut-être fictif, dont toutes les composantes du vecteur puissance sont égales à 1, est appelé sujet idéal théorique de C .

Dans ces conditions, on peut munir l'espace des puissances $[0, 1]^g$ d'une métrique du type χ^2 afin d'accentuer les effets de fortes implications génériques.

DEFINITION 15: On appelle distance implicative d'un sujet x à la classe ou au chemin C le nombre:

$$d(x, C) = \left[\frac{1}{g} \sum_{i=1}^{i=g} \frac{[\psi_i - \psi_{x,i}]^2}{1 - \psi_i} \right]^{\frac{1}{2}}$$

Ce nombre n'est autre que la distance dite du χ^2 entre les deux distributions $\{1 - \psi_i\}_i$ et $\{1 - \psi_{x,i}\}_i$ qui expriment les écarts entre les implications génériques empiriques et l'implication selon x. C'est pour cette raison que nous avons choisi le mot **typicalité** que nous allons définir plus bas. Si pour un i, $\psi_i = 1$, nous poserons, par convention, $\psi_{x,i} = 1$. Cette convention ne se fait pas contre nature puisque, dans ce cas, l'implication générique est maximale et significative d'une excellente liaison implicative entre ses deux termes, vérifiée par tous les sujets x de E. Ainsi, si le dénominateur s'annule, il en est de même du numérateur, sauf exception, et l'on pourra de toute façon attribuer la valeur 0 au quotient.

Typicalité et contribution d'un sujet et d'une variable supplémentaire à une classe ou à un chemin

Typicalité

Nous définirons la typicalité à partir de la «distorsion» du sujet considéré par rapport au sujet idéal théorique, tout en remarquant qu'il peut exister des sujets réels dont la distance à C soit inférieure à la distance à cette même classe ou à ce même chemin du sujet idéal théorique. La typicalité d'une catégorie de sujets ou d'une variable supplémentaire G s'en déduira.

DEFINITION 16: La typicalité de x à C est: $\gamma(x,C) = \frac{d(x_i,C)}{d(x,C)}$

et celle de G en est la moyenne dans G: $\gamma(G,C) = \frac{1}{\text{card}G} \sum_{x \in G} \gamma(x,G)$

Ces typicalités peuvent être infinies (pour des configurations contenant des x à distance nulle de C) mais, en particulier, supérieures à 1 pour certains sujets. Afin de donner au chercheur le moyen de savoir ou de vérifier rapidement si telle catégorie de sujets qui l'intéresse est statistiquement déterminante dans la constitution d'une classe implicative ou d'un chemin transitif, un algorithme a été élaboré en s'appuyant sur les deux notions suivantes: groupe optimal et catégorie déterminante.

DEFINITION 17: Soit E la population étudiée. Un groupe optimal d'une classe implicative ou d'un chemin C, groupe noté GO(C), est le sous-ensemble de E qui accorde à C une typicalité plus grande que le complémentaire de GO(C) et qui forme avec celui-ci une partition en deux groupes maximisant la variance inter-classe de la série statistique des typicalités individuelles. Une telle partition est dite *significative*.

L'existence de ce groupe optimal est démontrée dans (Gras R. et Ratsimba-Rajohn H. 1997). Les propriétés utilisées sont aussi celles qui le sont pour établir l'algorithme sur lequel se basent les modules des programmes informatiques qui construisent, automatiquement dans C.H.I.C., chaque sous-groupe optimal.

Considérons une partition $\{G_i\}_i$ de E. Cette partition peut être définie par une variable supplémentaire correspondant par exemple à un descripteur de E. Soit X_i une partie aléatoire de E ayant le même cardinal que G_i , et Z_i la variable aléatoire $\text{Card}(X_i \cap \text{GO}(C))$. Z_i suit une loi binomiale de paramètres: $\text{card } G_i$ et $\text{card } \text{GO}(C) / \text{card } E$.

DEFINITION 18: On appelle catégorie la plus typique de la classe implicative ou du chemin C, la catégorie qui minimise l'ensemble $\{p_i\}_i$ des probabilités p_i telles que:

$$\forall i, p_i = \text{Prob} [\text{card } G_i \cap \text{GO}(C) < Z_i]$$

Une catégorie G_0 est dite déterminante au seuil α si la probabilité associée p_0 est inférieure à α .

Ainsi, la signification d'une classe ou d'un chemin ayant été donnée par l'expert, il lui associera la sous-population la plus porteuse de ce sens. Cette approche est comparable à celle de I.-C. Lerman pour l'analyse des similarités, mais au moyen d'une modélisation et de concepts différents.

D'ailleurs, nous pouvons remarquer que nous pouvons associer au sous-groupe optimal une variable binaire correspondant à la fonction indicatrice de ce sous-ensemble de E . De la même façon, nous pouvons également associer à la catégorie G_i ou bien à la variable supplémentaire correspondante, une variable binaire dont l'indice de similarité

$$s = \frac{n_{a \wedge b} - \frac{n_a n_b}{n}}{\sqrt{\frac{n_a n_b}{n}}}, \text{ au sens de I.C. Lerman, vérifie: } p_i = \Pr\{S \geq s\}, S \text{ étant la}$$

valeur aléatoire dont s est la réalisation. Ainsi, minimiser l'ensemble des probabilités $\{p_i\}_i$ revient à maximiser l'indice de similarité entre les variables binaires, indicatrices de sous-ensembles, associées respectivement l'une au groupe optimal $GO(C)$ et les autres aux différentes catégories $\{G_i\}_i$.

Cette remarque permet d'étendre efficacement la notion de variable supplémentaire la plus typique à des variables numériques, prenant leurs valeurs sur $\{0, 1\}$. Il suffit dans ce cas d'extraire la plus forte des valeurs de similarité entre la variable binaire indicatrice définie par le groupe optimal et les différentes variables numériques placées en supplémentaire, l'indice étant calculé selon le principe retenu en analyse implicative pour les variables numériques (cf. Définition 5). Nous savons que sa restriction au cas binaire coïncide avec sa valeur s dans le cas où les deux variables sont binaires.

Ainsi il est possible de dégager à la fois les individus et les groupes d'individus typiques d'une liaison ou d'un ensemble (classe ou chemin) de liaisons. Ce sont donc ceux qui sont le plus en accord avec la qualité de ces liaisons au sein de la population E considérée. Si par exemple la liaison entre les variables a et b est quantifiée par le nombre $\psi(a, b) = 0,92$, les individus x qui attribuent à cette liaison la valeur $\psi_x(a, b) = 0,90$ sont plus typiques que ceux qui lui attribuent la valeur $0,98$. La nuance entre

cette notion et celle de contribution que nous allons définir prend tout son sens dans l'étude des variables modales ou numériques.

Contribution

Cette notion se distingue de la précédente par l'examen de la responsabilité des individus, puis des variables supplémentaires, qui peuvent en être des descripteurs, à l'existence d'une liaison d'une règle ou d'une méta-règle entre variables principales. Supposons, en effet, que deux variables a et b (resp. plusieurs variables sur un chemin du graphe ou bien deux classes de la hiérarchie) soient réunies par un arc sur un graphe à un certain seuil (resp. en un chemin transitif C du graphe ou bien en une classe C dans une hiérarchie à un certain niveau). Connaissant la valeur $\psi_{x,i}$ attribuée par l'individu à la règle $i: a \Rightarrow b$ (resp. règle i du chemin C ou bien de la classe C constituée de g règles génériques) supposée admissible,

DEFINITION 19: On appelle *distance de x à (a,b) ou à C* :

$$d(x,C) = \left[\frac{1}{g} \sum_{i=1}^{i=g} [1 - \psi_{x,i}]^2 \right] \text{ où } g = 1 \text{ dans le cas de l'arc } (a,b)$$

On appelle *contribution de x à C* le nombre: $\gamma(x,C) = 1 - d(x,C)$

Cette contribution a pour maximum 1 dans le cas où l'individu x a donné la valeur 1 à toutes les règles i . Ceci permet de concilier sémantique et définition formelle. La suite des définitions et des algorithmes de calcul (contribution d'une catégorie ou d'une variable supplémentaire G , groupe optimal d'individus, catégorie ou variable supplémentaire la plus contributive) se transpose immédiatement à partir des principes de la typicalité. Mais dans les situations réelles, nous observons la nuance entre les deux concepts ce qui enrichit l'information exploitable par l'utilisateur. Cependant, le concept de contribution est plus volontiers retenu pour l'interprétation.

Conclusion

Les applications de la méthode ont d'ores et déjà donné des résultats très satisfaisants, et non seulement à la discipline où elle a pris naissance mais aussi dans d'autres domaines de l'éducation ou de recherche scientifique plus générale, comme le montrent les autres communications présentées au cours des journées qui lui étaient consacrées. Le plus souvent, les interprétations s'appuient complémentaiement sur l'analyse de similarités ou/et sur les méthodes factorielles tout en apportant des informations qui lui sont spécifiques en raison de son caractère non symétrique. Les analyses bénéficient efficacement du logiciel C.H.I.C., qui permet, avec une certaine convivialité, tous les traitements algorithmiques et graphiques des méthodes évoquées dans cet article. Son développement suit régulièrement toutes les nouvelles avancées de la théorie de l'implication statistique. Citons, à cet égard, quelques pistes de recherche actuelle: données incomplètes, contraction de l'ensemble des variables, variables à valeurs intervalles, etc.

Bibliographie

- ALMOULOU S. A. (1992). *L'ordinateur, outil d'aide à l'apprentissage de la démonstration et de traitement de données didactiques*. Thèse. Université de Rennes 1.
- AGRAWAL, R.; IMIELINSKY, T. et SWAMI A. (1993). Mining association rules between sets of items in large databases. *Proc. of the ACM SIGMOD'93*.
- AMARGER, S.; DUBOIS, D. et PRADE, H. (1991). "Imprecise quantifiers and conditional probabilities". In: KRUSE, R. et SIEGEL, P. *Symbolic and quantitative approaches to uncertainty*. Berlin Heidelberg, Springer-Verlag.
- BAILLEUL, M. (1994). *Analyse statistique implicative: variables modales et contribution des sujets. Application à la modélisation de l'enseignant dans le système didactique*. Thèse. Université de Rennes 1.
- BAILLEUL, M. et GRAS, R. (1995). L'implication statistique entre variables modales. *Mathématique, Informatique et Sciences Humaines*, n. 128, pp. 41-57.

- BERNARD, J.-M. et POITRENAUD, S. (1999). L'analyse implicative bayésienne d'un questionnaire binaire: quasi-implications et treillis de Galois simplifié. *Mathématiques, Informatique et Sciences Humaines*, n° 147, pp. 25-46.
- BODIN, A. (1997). Modèles sous-jacents à l'analyse implicative et outils complémentaires. *Prépublication IRMAR*, n. 97-32.
- BODIN, A. et GRAS, R. (1999). Analyse du préquestionnaire enseignants avant EVAPM-Terminales. *Bulletin de l'Association des Professeurs de Mathématiques de l'Enseignement Public*, n. 425, pp. 772-786.
- COUTURIER, R. et GRAS, R. (1999). Introduction de variables supplémentaires dans une hiérarchie de classes et application à CHIC. *Actes Des 7^{èmes} Rencontres de la Société Francophone de Classification*, 15-17 septembre.
- DIDAY, E. (1972). *Nouvelles méthodes et nouveaux concepts en classification automatique et reconnaissance des formes*. Thèse d'Etat. Université Paris VI.
- GANASCIA, J. G. (1987). *AGAPE et CHARADE: deux techniques d'apprentissage symbolique appliquées à la construction de bases de connaissances*. Thèse d'Etat. Université Paris Sud.
- GRAS, R. (1979). *Contribution à l'étude expérimentale et à l'analyse de certaines acquisitions cognitives et de certains objectifs didactiques en mathématiques*. Thèse d'Etat. Université de Rennes I.
- GRAS, R. et LARHER, A. (1992). L'implication statistique, une nouvelle méthode d'analyse de données. *Mathématique, Informatique et Sciences Humaines*, n. 120, pp. 5-31.
- GRAS, R. et RATSIMBA-RAJOHN, H. (1996). Analyse non symétrique de données par l'implication statistique. *RAIRO-Recherche Opérationnelle*, 30-3, AFCET, pp. 217-232.
- GRAS, R.; BRIAND, H. et PETER, P. "Structuration sets with implication intensity". In: DIDAY, E.; CHEVALLIER, Y. et OPITZ, O. (eds.) (1996). *Proceedings of the International Conference on Ordinal and Symbolic Data Analysis – OSDA 95*. Paris, Springer.
- GRAS, R. et alii (1996). *L'implication Statistique*. Collection Associée à "Recherches en Didactique des Mathématiques". Grenoble, La Pensée Sauvage.

- GRAS, R.; BRIAND, H.; PETER, P. et PHILIPPE, J. (1997). Implicative statistical analysis. *Proceedings of International Congress I.F.C.S.*, 96. Kobé, Tokyo, Springer-Verlag.
- GRAS, R.; KUNTZ, P.; COUTURIER, R. et GUILLET, F. (2001a). Une version entropique de l'intensité d'implication pour les corpus volumineux. Hermès Science Publication. *Extraction des Connaissances et Apprentissage (ECA)*, v. 1, n. 1-2, pp. 69-80.
- GRAS, R.; KUNTZ, P. et BRIAND, H. (2001b). Les fondements de l'analyse statistique implicative et quelques prolongements pour la fouille de données. *Mathématiques et Sciences Humaines*, n. 154-155, pp. 9-29.
- GRAS, R.; DIDAY, E.; KUNTZ, P. et COUTURIER, R. (2001c). Variables sur intervalles et variables-intervalles en analyse statistique implicative. *Actes du 8^{ème} Congrès de la Société Francophone de Classification*. Université des Antilles-Guyane, Pointe-à-Pitre, 17-21 décembre.
- GRAS, R.; GUILLET, F.; GRAS, R. et PHILIPPE, J. (2001d). Réduction des colonnes d'un tableau de données par quasi-équivalence entre variables. *Extraction des connaissances et apprentissage*. Hermès, v. 1, n. 4, pp. 197-202.
- GRAS, R.; KUNTZ, P. et BRIAND, H. (2003). Hiérarchie orientée de règles généralisées en analyse implicative. *Extraction des Connaissances et apprentissage*, Hermès, pp. 145-15.
- LAGRANGE, J. B. (1998). Analyse implicative d'un ensemble de variables numériques; application au traitement d'un questionnaire à réponses modales ordonnées. *Revue de Statistique Appliquée, Institut Henri Poincaré*, pp. 71-93.
- LAHANIER-REUTER, D. (1998). *Étude de conceptions du hasard: approche épistémologique, didactique et expérimentale en milieu universitaire*. Thèse. Université de Rennes 1.
- LERMAN, I. C. (1970). *Les bases de la classification hiérarchique*. Paris, Gauthier-Villars.
- _____ (1981). *Classification et analyse ordinale des données*. Paris, Dunod.
- LERMAN, I. C.; GRAS, R. et ROSTAM, H. (1981). Elaboration et évaluation d'un indice d'implication pour des données binaires. I et II. *Mathématiques et Sciences Humaines*, n. 74, pp. 5-35 et n. 75, pp. 5-47.

- LOEVINGER J. (1947). A systematic approach to the construction and evaluation of tests of abilities. *Psychological Monographs*, v. 61, n. 4.
- PEARL J. (1988). *Probabilistic Reasoning in intelligent systems*. San Mateo, CA, Morgan Kaufmann.
- SIMON A. (2000). *Outils classificatoires par objets pour l'extraction de connaissances dans des bases de données*. Thèse. Université de Nancy 1.

Recebido em mar./2002 aprovado em abr./2002

Une étude diagnostique en vue de la formation des enseignants en géométrie

SADDO AG ALMOULOU*

Resumo

A artigo apresenta um estudo diagnóstico sobre o ensino e a aprendizagem da geometria no ensino fundamental brasileiro. A análise do sistema educativo brasileiro e do discurso de professores permitiu identificar alguns dos fatores que seriam fontes de dificuldades que os professores enfrentam quanto ao ensino da geometria. Os resultados deste estudo serviram de apoio nas escolhas de nossas hipóteses, a propósito dos conteúdos geométricos e das variáveis didáticas a privilegiar na formação dos professores engajados no projeto de pesquisa.

Palavras-chave: geometria; formação de professores; discurso do professor; sistema educativo brasileiro.

Résumé

L'article présente une étude diagnostique sur l'enseignement et l'apprentissage de la Géométrie au niveau de l'enseignement fondamental brésilien. L'analyse du système éducatif brésilien et du discours des enseignants nous ont permis d'identifier certains des facteurs qui seraient à l'origine des difficultés que les enseignants rencontrent dans l'enseignement de la Géométrie. Les résultats de cette étude ont servi de point d'appui quanto aux choix de nos hypothèses à propos des contenus géométriques et des variables didactiques à prendre en compte dans la formation des enseignants engagés dans le projet de recherche.

Mots-clé: géométrie; formation des enseignants; discours du maître; système éducatif brésilien.

Abstract

The article presents a diagnostic study about the teaching and learning of geometry in the Brazilian secondary school system. The analysis of the Brazilian education system and of teachers' discourse revealed some of the factors that would constitute sources of difficulties faced by teachers in geometry teaching. The results of this study supported the choice of our hypotheses concerning geometry contents and the didactic variables that should be focused in the education of teachers engaged in the research project.

Key-words: geometry; teachers' education; teachers' discourse; Brazilian education system.

* PUC-SP. E-mail: saddoag@pucsp.br.

Introduction

Le travail que nous présentons fait partie d'un projet de recherche (financé par la Fapesp) dont l'objectif est l'étude des facteurs et des stratégies susceptibles d'influencer l'enseignement et l'apprentissage des notions géométriques au niveau des classes de la 5^a à la 8^a de l'enseignement fondamental brésilien (élèves de 11 à 14 ans). Le projet de recherche s'est développé selon les points consécutifs suivants:

- étude diagnostique: tests, interviews individuels et observations;
- élaboration de situations, études de ces situations par le groupe (concepts, constructions géométriques, raisonnement, démonstration);
- activités de formation des enseignants en situation papier-crayon;
- activités de formation avec l'aide de l'environnement informatique (Cabri-géomètre, Logo);
- élaboration et analyse d'activités par les enseignants en formation.

Nous discutons ici, en particulier, le discours des enseignants sur le rôle de la Géométrie dans la formation des élèves, sur son enseignement et son apprentissage. Nous présentons essentiellement les résultats de l'une des phases les plus importantes de l'élaboration du dispositif de formation, à savoir, l'étude diagnostique d'un groupe d'enseignants en vue de définir les caractéristiques de la formation à leur proposer dans le domaine de la Géométrie pour des élèves de 11 à 14 ans.

L'étude diagnostique du groupe des enseignants à former se fait selon deux volets:

- les représentations des enseignants concernant l'enseignement et l'apprentissage de la Géométrie;
- l'observation des compétences des enseignants sur des contenus géométriques.

Les résultats de cette étude vont contribuer aux choix de nos hypothèses de travail quant aux contenus géométriques et aux variables à prendre en compte dans la formation des enseignants, mais aussi dans le choix des situations d'enseignement/apprentissage de la Géométrie.

Système éducatif brésilien et la formation des enseignants

L'analyse du système éducatif brésilien et de l'enjeu de la Géométrie nous a permis d'identifier certains facteurs qui seraient à l'origine des difficultés que les enseignants rencontrent pour l'enseignement et l'apprentissage des savoirs et des connaissances géométriques. À l'origine de ces problèmes, nous identifions les faits suivants:

À propos de la formation des enseignants:

- les enseignants ont eu une formation de base très précaire en Géométrie;
- les cours de formation initiale n'intègrent pas suffisamment une réflexion profonde sur l'enseignement de la Géométrie;
- les modalités de formation continue, n'ont pas encore atteint leur objectif par rapport à la Géométrie.

À propos des situations d'enseignement:

De façon générale, les situations proposées dans les manuels scolaires et par la majorité des enseignants sont caractérisées par les faits suivants:

- non-coordination des registres de représentation sémiotique (Duval, 1995);
- non-perception du rôle important de la figure dans la visualisation et les phases d'exploration;
- les problèmes proposés par la majeure partie des livres scolaires brésiliens sont de type "algébrique" (peu de travail sur le raisonnement déductif et sur la démonstration);
- le système éducatif brésilien n'impose pas aux divers niveaux d'enseignement un programme officiel et obligatoire de Mathématiques. Il définit la politique générale de l'éducation, des recommandations et orientations générales sur les méthodes, les savoirs et savoir-faire. Chaque école définit les contenus qu'elle juge importants pour la formation des élèves, et la Géométrie est très souvent laissée de côté;
- le passage de la Géométrie empirique à la Géométrie déductive est quasi-inexistant;

- peu de travail sur la lecture et l'interprétation des textes mathématiques.

Les paramètres curriculaires nationaux de 1998 mettent l'accent sur:

- l'importance de la Géométrie au niveau du quatrième cycle (7^a et 8^a séries, soit des élèves de 13-14 ans);
- l'importance de l'élaboration de situations-problèmes qui favorisent le raisonnement déductif et l'introduction de la démonstration;
- le rôle important de la figure et les principales fonctions d'un dessin: visualisation, résumé d'informations, aide à la preuve et aux conjectures.

Mais la majorité des enseignants de l'enseignement fondamental et de lycée n'est pas préparée pour mettre en œuvre les recommandations et les orientations didactiques et pédagogiques des paramètres curriculaires nationaux.

Méthodologie utilisée pour réaliser l'étude diagnostique

Nous avons fait passer à ces enseignants un questionnaire dont la structure est la suivante:

- informations sur les enseignants (formation, âge, sexe...);
- accès à l'information (TV, journaux, revues, etc.);
- méthodologies utilisées pour l'enseignement et l'apprentissage de la Géométrie;
- leurs difficultés et celles de leurs élèves en Géométrie, origine de ses difficultés et les stratégies envisagées pour les résoudre;
- rôle de la résolution de problèmes dans l'apprentissage de la Géométrie;
- analyse didactique des erreurs des élèves.

Nous avons également fait passer aux élèves de ces enseignants un questionnaire dont l'objectif est d'identifier les problèmes que ces derniers rencontrent dans l'acquisition des savoirs et des connaissances (au sens de Brousseau) géométriques. Mais cette partie de la recherche ne fera pas l'objet d'étude dans ce texte.

Les Instruments d'analyse des données multidimensionnelles

Nous présentons essentiellement les résultats de l'analyse de similarité et de la hiérarchie implicative. Au moyen de ces analyses, nous recherchons à synthétiser et structurer les réponses des enseignants afin d'obtenir une typologie de comportements.

Nous utilisons le logiciel CHIC (Classification Hiérarchique, Implicative et Cohésitive) pour le traitement des données statistiques multidimensionnelles. Ce logiciel développé sous la direction de Régis Gras, traite essentiellement la Classification Hiérarchique (Lerman), l'implication statistique et la hiérarchie implicative (Gras et son équipe).

L'analyse hiérarchique de similarité permet de constituer sur l'ensemble des variables statistiques étudiées des partitions de moins en moins fines, construites de façon ascendante en arbre à l'aide d'un critère de similarité. Elle permet d'étudier et d'interpréter en termes de typologie et de ressemblance (et de dissemblance) des classes de variables, constituées significativement à certains niveaux et s'opposant à d'autres à ces mêmes niveaux.

Au moyen de l'analyse statistique implicative des données, nous cherchons à dégager des structures implicatives au sens suivant: telle attitude *a* s'accompagne, de façon conséquente ou non, de telle attitude *b*. Cette expression s'apparente à l'implication $a \rightarrow b$ ou à l'inclusion de l'ensemble de ceux qui ont *a* dans l'ensemble de ceux qui ont *b*. En fait, cette implication ou cette inclusion stricte étant rarement observée, l'analyse implicative devient statistique lorsqu'une mesure estime l'"étonnement" de l'écart entre la relation stricte et la relation observée.

La hiérarchie implicative de classe permet une analyse de relations intra-classes et inter-classes de réponses. Groupant des réponses dans la mesure où une relation implicative les lie, nous obtenons des classes formées par l'implication.

Pour l'analyse du discours des enseignants, nous avons sélectionné certaines des questions qui nous paraissaient les plus pertinentes. Les variables statistiques (voir les annexes) retenues sont des variables binaires. Elles ont été sélectionnées en nous appuyant sur une analyse qualitative et une analyse statistique descriptive.

Analyse du discours des enseignants

Nous présentons les résultats partiels de l'analyse du discours des enseignants participant au projet. Nous visons essentiellement:

- l'analyse du discours des enseignants par rapport au rôle de la Géométrie dans la formation des élèves, à son enseignement et à son apprentissage;
- l'obtention d'une caractérisation de ces enseignants qui participent à l'analyse;
- l'identification des contenus de Géométrie qu'ils disent enseigner, ainsi que ceux qui posent problème quant à leur enseignement et leur apprentissage.

Caractéristiques des enseignants participant à la recherche

Le tableau suivant nous donne un aperçu général sur les principales caractéristiques des enseignants participant au projet de recherche.

Les enseignants	
Nombre	24
Âge	Entre 31 et 40 ans
Expérience professionnelle	Enseignent depuis 2 à 5 ans
Formation	Ils ont eu une formation en Géométrie
Méthode utilisée pour travailler la Géométrie en classe	
Cours magistral	15 enseignants
Mettent les élèves en situation de recherche	4 enseignants
Travail en groupe	13 enseignants
Résolution de problèmes	13 enseignants
Utilisent les jeux comme moyen didactique	8 enseignants
Activités expérimentales	7 enseignants
À propos des documents officiels sur l'enseignement des Mathématiques	
Programmes officiels de l'État de São Paulo: 12 enseignants les connaissent et 6 s'en ont inspiré	
Expériences Mathématiques (livres scolaires édités par le gouvernement de l'État de São Paulo): 7 enseignants affirment les connaître et 4 disent les avoir utilisés.	
PCN (Paramètres curriculaires Nationaux): 5 enseignants les connaissent et les utilisent	

Les contenus géométriques qu'ils disent avoir l'habitude d'enseigner	
Quadrilatères	11 enseignants
Similitude des triangles	14 enseignants
Circonférence et Cercle	12 enseignants
Théorème de Pythagore	13 enseignants
Aires et Périmètre	15 enseignants
Relations métriques dans un triangle rectangle	13 enseignants
Théorème de Thalès	11 enseignants
Triangles	15 enseignants
Égalité de triangles	9 enseignants
Transformations géométriques	4 enseignants

Analyses multidimensionnelles du discours des enseignants: analyse de similarité

Rappelons que notre population est constituée de deux groupes: le groupe de vendredi est constitué d'enseignants qui ont déjà reçu, au cours de leurs formations continues passées, des informations sur les tendances actuelles sur l'enseignement et l'apprentissage des mathématiques. Nos rencontres avec ces enseignants ont lieu tous les vendredis.

Le groupe de jeudi est composé d'enseignants qui affirment n'avoir jamais reçu une formation continue tant du point de vue des contenus mathématiques ni du point de vue didactique et pédagogique. Ils travaillent tous dans la même école. Nos rencontres ont lieu les jeudis.

Le traitement des données par CHIC nous donne l'arbre de similarité (voir annexe), dont nous analyserons les principaux blocs.

Premier bloc est constitué des variables: 1, 30, 44, 54, 60, 3, 35, 21, 38, 39, 51, 23, 36, 26 et 48.

Il est constitué de trois sous-classes:

Groupe de vendredi: Ce groupe met en évidence l'opinion selon laquelle il est important d'accorder aux élèves une grande autonomie dans la construction de leur connaissance, en opposition avec une attitude plus directive de l'enseignant. Une partie des enseignants semble mettre

l'accent sur l'importance des vérifications empiriques des propriétés et des relations géométriques, l'initiation au raisonnement et à la démonstration. Ils semblent être d'accord que la démonstration est d'une extrême importance pour la formation intellectuelle des élèves. À cet effet, elle ne doit pas être abandonnée, mais ils trouvent qu'il est prématuré de l'introduire dans l'enseignement fondamental et qu'elle doit être initiée en début de lycée où les élèves auront plus de maturité pour l'affronter. Nous pensons que cette conception sur l'initiation à la démonstration semble due à certains facteurs parmi lesquels nous citons:

- la Géométrie, en particulier l'initiation au raisonnement et à la démonstration, n'occupe pas une place de choix dans l'enseignement, l'accent est plutôt mis sur les aspects calculatoires de la Géométrie;
- de façon générale, les enseignants de l'enseignement fondamental pointent la Géométrie comme la matière la plus problématique quant à son enseignement et à son apprentissage;
- comme nous l'avons déjà signalé, la plus grande partie des enseignants actuels a eu une formation initiale très précaire en Géométrie;
- les cours de formation initiale, tant des enseignants de l'enseignement fondamental, que de ceux de lycée, n'ont pas encore atteint les résultats escomptés, il en est de même pour les cours de formation continue.

A la deuxième sous-classe contribuent principalement les enseignantes. Elle met en évidence l'opinion selon laquelle le programme de mathématiques doit être exécuté en respectant le rythme des élèves. L'analyse de ce groupe de variables statistiques révèle également que l'usage de la règle et du compas ne paraît pas fondamental pour l'apprentissage de la Géométrie. Mais une partie des enseignants semble admettre que la Géométrie permet à l'élève de réaliser des recherches, de résoudre des problèmes, d'imaginer différentes stratégies de résolution et de les justifier. Ils pensent, de plus, que la Géométrie prépare l'élève à utiliser différents processus et méthodes de résolution d'un problème en analysant leurs différences et leurs points communs.

La troisième sous-classe s'oppose aux deux premières sous-classes que nous venons de traiter. Cette sous-classe met en évidence l'importance

de la règle et du compas dans la construction des aptitudes en Géométrie. De plus, il révèle la conception selon laquelle la participation active des élèves à la construction de leur connaissance ne provoque pas nécessairement une perte de contrôle de la classe de la part de l'enseignant. Leur opinion confirme aussi le peu d'intérêt à accorder à l'enseignement/apprentissage de la Géométrie par rapport à d'autres thèmes mathématiques.

Le deuxième bloc met en jeu les variables statistiques: 2, 32, 27, 12, 19, 20, 17, 4, 42, 5, 6, 10 et 9. Elle met en évidence trois sous-classes de variables:

Le groupe de jeudi: Ce groupe est caractérisé par les opinions qui ne semblent pas accepter d'accorder une grande autonomie à l'élève dans la construction de ses connaissances. Mais en même temps révèle qu'il est important de l'encourager à la recherche de solution et à la construction de ses propres connaissances. Répondre ainsi indiquerait que les enseignants en question considèrent que accorder une autonomie aux élèves serait un obstacle à la mise en place de certaines connaissances et savoirs. Selon certains enseignants, accorder une certaine autonomie aux élèves serait leur accorder une certaine liberté qui, dans la majorité des cas vécus, provoque un problème lié à la gestion de la classe.

Cette classe met aussi en évidence les contenus géométriques que les enseignants disent enseigner: isométrie de triangles, transformations géométriques, triangles semblables et théorème de Thalès. Mais une analyse de leurs pratiques en classe montre qu'une bonne partie de ces contenus (les transformations géométriques, triangles semblables, par exemple) n'est pas abordée.

Le sexe masculin: Cette sous-classe semble révéler l'opinion selon laquelle la démonstration en Géométrie n'a aucune utilité pratique pour les élèves de l'enseignement fondamental, et qu'elle doit être supprimée des programmes scolaires. Il semble que, pour une partie des enseignants de sexe masculin, la démonstration n'a aucune utilité pour la vie pratique des enfants jusqu'à l'adolescence, et que par conséquent, elle ne doit pas faire l'objet d'enseignement au niveau du cycle fondamental. Son étude doit être initiée en début de lycée. Pour l'enseignement de la Géométrie, il est important de faire travailler les élèves sous forme de recherche, d'activités expérimentales, en utilisant des jeux, mais aussi en faisant des cours magistraux.

Troisième bloc mettant en jeu les variables 7, 8, 45, 57, 11, 13, 14, 16, 15 et 18.

Ce bloc met en évidence l'importance de la résolution de problèmes dans l'apprentissage de la Géométrie, mais cette classe met aussi en évidence l'opinion selon laquelle la Géométrie ne prépare pas les élèves à une attitude réflexive. Elle met aussi en évidence les contenus géométriques qu'ils disent enseigner: quadrilatères, circonférence et cercle, théorème de Pythagore, relations métriques dans un triangle, aires et périmètre. Ces contenus sont caractérisés par leur aspect numérique.

En croisant les réponses des enseignants, nous observons que certains en même temps: "Enseignent en utilisant la résolution de problèmes (variable 8)", sont d'accord que "le travail en Géométrie développe chez l'enfant l'aptitude à rechercher des exemples et contre-exemples, à formuler des hypothèses et à prouver expérimentalement (variable 45)", ne sont pas d'accord avec l'opinion "La Géométrie offre à l'élève la possibilité de réaliser des investigations, résoudre des problèmes, créer des stratégies, les justifier et avoir des arguments sur ses stratégies (variable 39)". Ces discours indiqueraient que les enseignants en questions considèrent comme disjointes les activités d'exploration expérimentale de celles qui entraînent les élèves aux raisonnements et les initient à la démonstration. L'existence de cette idée se confirme aussi pour le cas des enseignants qui sont d'accord avec l'opinion 45, mais ne sont pas d'accord avec l'opinion "On ne doit pas abandonner, au cours de toutes les phases de l'enseignement fondamental, les vérifications empiriques de propriétés et de relation, mais qu'on doit aussi favoriser un travail sur des démonstrations simples".

Nous faisons l'hypothèse que ces opinions sont en étroites relations avec la formation des enseignants, les situations d'enseignement proposées par les manuels scolaires.

Analyses multidimensionnelles du discours des enseignants: hiérarchie implicative

L'arbre de la hiérarchie implicative (voir les annexes) met en évidence 14 classes de variables statistiques. Nous analysons celles qui nous semblent les plus pertinentes.

- La classe (16 \Rightarrow 15) \Rightarrow 30 (classe de cohésion la plus élevée.) met en relation les contenus que les enseignants disent enseigner: les relations métriques dans un triangle et le calcul d'aires. Ils se caractérisent par leur aspect algébrique. Les figures géométriques semblent (l'analyse de Manuels scolaires le montre) servir uniquement de support visuel.

L'opinion des enseignants sur l'enseignement de ces contenus est en relation implicative avec leur opinion "être d'accord que la Géométrie permet à l'élève de développer l'esprit de recherche".

- La classe 10 \Rightarrow (13 \Rightarrow 18) met en évidence l'enseignement de la "circonférence, du cercle et des triangles". L'enseignement de ces contenus semble partir d'activités au cours desquelles les élèves participent à la construction de leurs connaissances à partir de situations expérimentales.

- La classe implicative qui regroupe les variables [(20 \Rightarrow 17) \Rightarrow (7 \Leftrightarrow 14)] \Rightarrow [(8 \Rightarrow 37) \Rightarrow 33] est caractérisée par les contenus de Géométrie que certains enseignants disent avoir l'habitude de travailler en classe, ainsi que les méthodes qu'ils disent utiliser pour leur enseignement/apprentissage. Leur discours met l'accent sur l'importance des méthodes prenant en compte l'élève, méthodes pour lesquelles l'élève est acteur de la construction de ses connaissances. C'est un discours qui semble prendre en compte les tendances des recherches actuelles en Didactique des Mathématiques, mais nos observations révèlent que ce discours est différent de la réalité en classe.

Cette dernière observation rejoint celle d'Aline Robert (Robert et alii, 1989) selon laquelle les pratiques des enseignants sont intimement liées à leurs conceptions sur les Mathématiques et l'enseignement reçu au cours de leur formation.

Ces conceptions sont probablement liées à leurs expériences personnelles, à l'environnement socioculturel présent et passé, à la période de leurs études et à des caractéristiques encore plus personnelles.

La stabilité des conceptions d'un individu présente quelquefois des résistances au changement, à la fois parce qu'un équilibre personnel est à maintenir, mais aussi, parce qu'une partie des conceptions correspond à des convictions (éventuellement implicites, non perçues comme des réponses à des questions, mais admises, de préférence, sans qu'il ait conscience du phénomène ou sans pouvoir argumenter à leur propos).

- Implication: $8 \Rightarrow 37$: la résolution de problèmes liés aux calculs d'aires est associée à la composition et la décomposition de figures. Une partie des enseignants semble mettre l'accent sur les fonctions de la figure (visualiser, faire voir, résumer les informations, aide à la preuve et aux conjectures) pour l'acquisition des connaissances liées au concept d'aire. Nous rapprochons ses idées à celles de Raymond Duval (1995).

L'auteur montre que la résolution de problèmes de Géométrie et l'entrée dans la forme de raisonnement que cette résolution exige, dépendent de la prise de conscience de la distinction des différentes formes d'appréhension de la figure (appréhensions séquentielle, perceptive, discursive et opératoire). L'appréhension opératoire des figures, selon l'auteur, dépend de la prise de conscience des différentes modifications possibles d'une figure (modification méréologique, optique et positionnelle).

Ces modifications sont réalisées psychiquement, graphiquement et mentalement. L'intérêt de fractionner une figure ou de son examen à partir des parties élémentaires est lié à l'opération de configuration intermédiaire. En effet, les parties élémentaires peuvent être regroupées en sous-figures, toutes dans la figure de départ. Cette opération permet alors d'enchaîner immédiatement des traitements tels que les mesures d'aires à partir des sommes des parties élémentaires ou mettre en évidence l'équivalence de deux regroupements intermédiaires.

- La classe implicative ($22 \Leftrightarrow 28$) $\Rightarrow 38$ semble révéler l'importance de l'utilisation d'instruments, comme la règle et le compas, mais aussi la nécessité pour l'enseignant de prendre en compte les intérêts des élèves.

L'ampliation et la réduction des figures, et les instruments de constructions de figures géométriques, semblent constituer des appuis importants pour l'acquisition de notions géométriques.

- La classe implicative ($2 \Leftrightarrow 26$) semble mettre en évidence le désaccord de certains enseignants sur l'importance de laisser une certaine autonomie à leurs élèves.

L'analyse des relations implicatives paraît révéler, pour une partie des enseignants, qu'il est prématuré de travailler la Géométrie déductive

et que cette dernière doit être initiée seulement en début de lycée. Or, comme nous le savons bien, plusieurs recherches montrent le contraire.

Analysant les causes d'échec des élèves dans une tâche de démonstration en Géométrie, Duval (1995) dit qu'elle met en jeu une activité cognitive spécifique et que son apprentissage n'est pas lié à une situation d'interaction sociale, ni subordonnée à un jeu de pressions internes d'un objet. Elle est un type de processus cognitif autonome avec des caractéristiques spécifiques par rapport à d'autres formes de fonctionnement du raisonnement, comme l'induction, l'argumentation, l'interprétation.

D'un côté, elle articule les énoncés en fonction de leur statut et non en fonction de leur signification, de l'autre, elle se fait en progression par substitution d'énoncés et par enchaînement. L'apprentissage de la démonstration, consiste, pour Duval, prioritairement, à la conscientisation qu'il s'agit d'un discours différent de celui qui est pratiqué dans la pensée naturelle. La compréhension opératoire des définitions et des théorèmes suppose que ceux-ci soient vus comme des règles de substitution. Pour l'auteur, la prise de conscience de ce qui est une démonstration est faite seulement à partir de l'articulation de deux registres, parmi lesquels l'utilisation du langage naturel par l'élève. Cette interaction va surgir de l'interaction entre la représentation non discursive produite et le discours écrit ou oral.

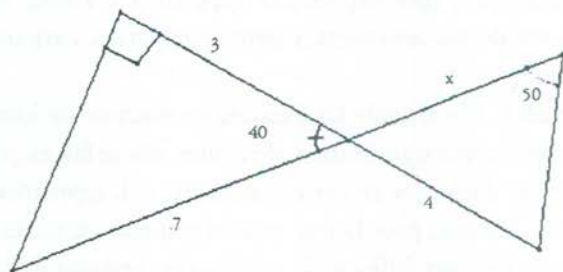
L'apprentissage de la démonstration par les élèves est un processus long, et doit être initié dès l'enseignement fondamental. Pour amener une majorité d'élèves (de collège) à la prise de conscience de la structure profonde de la démonstration, nous pensons que l'enseignement doit prendre en compte plus d'activités de résolution de problèmes, parmi lesquelles nous citons l'exploration guidée ou libre des propriétés d'une figure liée à un énoncé d'un problème et l'aide à la démonstration, permettant non seulement de développer les capacités de raisonnement de l'enfant, mais aussi de comprendre le statut de la démonstration géométrique.

Analyse des questions relatives aux contenus

Analyse de la situation 1

Situation 1

Un enseignant propose à ses élèves le problème suivant. Observe la figure ci-dessous:



Quelle est la valeur de x ?

- Penses-tu que tes élèves vont répondre que $x = \dots$?
- Justifie ta réponse?
- Cite trois difficultés au moins que, à ton avis, les élèves commettent fréquemment face à ce type de situation.
- Quelles sont les raisons de ces difficultés?
- Quelle est l'importance de la figure pour la résolution du problème proposé aux élèves?
- Explique comment tu corrigerais cet exercice à tes élèves?
- Considères-tu l'énoncé du problème: () Bien formulé
() mal formulé? Propose une autre formulation du problème.

Quelques résultats

Dix-neuf des 24 enseignants ont répondu que leurs élèves ne sauront pas répondre ou ne sauront pas résoudre la question posée. Ils justifient cet état de fait, par celui que la Géométrie est reléguée en dernière position et par la méconnaissance de cette discipline par la majorité des enseignants. L'une des raisons des difficultés que les enseignants soulèvent

est liée à l'interprétation de la figure, interprétation qui généralement n'est pas prise en compte dans l'enseignement de la Géométrie.

Les difficultés sont aussi dues, selon certains enseignants, aux facteurs suivants:

- manque de contenu géométrique tels que: congruence des triangles, triangles semblables, proportionnalités, théorèmes de Thalès;
- manque d'habiletés et de certaines attitudes comme: réfléchir, être capable de l'abstraction, raisonner, savoir calculer, savoir interpréter;
- la peur des mathématiques, le manque d'intérêt et d'assurance face aux problèmes de Géométrie.

La figure a été considérée très importante dans la résolution du problème, parce qu'elle permet de visualiser les données et les sous-figures pertinentes. Certaines recherches associent la visualisation à l'apprentissage de la Géométrie (Ponte et alii, 1998). Ces recherches suggèrent que les compétences du domaine spatial soient travaillées avant celles normalement attendues quant à la Géométrie plane.

L'analyse des résultats de cette situation montre que les difficultés attribuées aux élèves sont en réalité, aussi celles des enseignants à propos de la Géométrie. À la question qui leur demandait de résoudre le problème pour leurs élèves, seuls huit l'ont résolu correctement, et un seulement avait indiqué de façon adéquate les connaissances mobilisables pour la résolution du problème.

Ces résultats confirment celles de Belchior (1994, cité par Ponte et alii, 1998) selon lequel certains enseignants de l'enseignement fondamental et de lycée avaient, dans une tâche de résolution de problèmes de Géométrie, des résultats proches de ceux de leurs élèves.

Par rapport à l'énoncé du problème, le tableau suivant met en évidence les opinions des enseignants sur sa formulation:

	Bonne formulation	Mal formulée	Sans opinion
Nombre d'enseignants	11	5	8

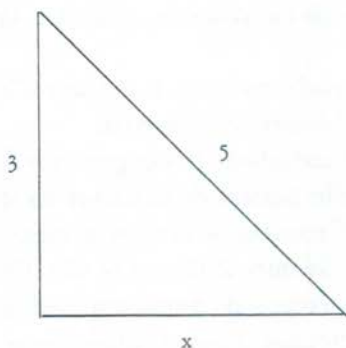
Les enseignants qui pensent que le problème est mal formulé suggèrent qu'il y ait plus d'informations afin que les élèves puissent utiliser le théorème de Pythagore dans la résolution du problème.

Analyse de la situation 2

Situation 2

Le problème suivant a été proposé aux élèves:

Que peut-on dire de la mesure x du côté du triangle, sachant qu'on s'est servi de la même unité de longueur pour toutes les mesures?



Un de tes élèves a répondu: $x = 4$.

La réponse de ton élève est-elle juste?

Quelle serait la stratégie que ton élève aurait utilisée pour résoudre ce problème?

Se tu considères que la réponse de ton élève est fautive, quelle solution pourrais-tu lui proposer? Justifie ta réponse.

Quelques résultats

Sur les 24 enseignants qui ont résolu cette situation, 16 trouvent que la solution proposée par l'élève est correcte. Six enseignants disent qu'ils ne peuvent la considérer correcte car les informations contenues dans la figure ne permettent pas de considérer le triangle donné comme un triangle rectangle. Dix-sept des enseignants pensent que l'élève a utilisé le théorème de Pythagore pour la résolution du problème. Deux enseignants ont affirmé que l'élève avait utilisé la moyenne arithmétique

des mesures pour trouver la valeur de x : $\frac{3+5}{2} = \frac{8}{2} = 4$, trois enseignants

disent que leurs élèves ne pourront pas résoudre le problème, raison invoquée: l'angle droit n'est pas fourni.

La résolution de ce type de situation doit prendre en compte les trois formes de processus cognitifs identifiés par Duval (1995) et qui ont des fonctions épistémologiques spécifiques:

- la visualisation qui permet l'exploration heuristique d'une situation complexe;
- la construction géométrique au cours de laquelle les actions représentées et les résultats observés sont liés aux objets mathématiques représentés;
- le raisonnement. Suivant l'auteur, l'heuristique des problèmes de Géométrie se réfère à un registre spatial, qui donne lieu à des formes d'interprétation autonomes mettant en jeu les différentes appréhensions de la figure: appréhension séquentielle, perceptive, discursive et opératoire.

Analyse de la situation 3

Situation 3

On sait d'un quadrilatère qu'il a trois côtés de même longueur.

Henrique dit: Il suffit que ses diagonales se coupent en leur milieu pour qu'il soit un losange.

Carlos Magalhães argumente: Il suffit qu'il ait un angle droit pour qu'il soit un carré.

Pedro Malan dit: Il suffit que le quatrième côté soit le double de l'un des trois côtés pour qu'il soit un trapèze.

- Qui a raison?
- Qui a tort?
- Justifie ta réponse?

Quelques résultats

Le tableau ci-dessous résume les réponses des enseignants:

	A raison	N'a pas raison
Henrique	3	21
Magalhães	14	10
Pedro Malan	11	13

Les justifications de leurs réponses ne s'appuie pas sur des arguments scientifiques, aucune démonstration ou contre-exemple n'est proposé pour valider ou réfuter les réponses des trois élèves. De façon générale, ces enseignants ont eu beaucoup de difficultés à résoudre le problème. L'absence de support visuel (la figure, par exemple) serait l'une des causes des difficultés des élèves à résoudre le problème proposé (selon leurs propres commentaires).

Conclusions et perspectives

L'étude des manuels scolaires, des PCN et les informations obtenues à partir de ce questionnaire révèlent une certaine réalité de l'enseignement de la Géométrie, ainsi que la nécessité d'une formation continue de ces enseignants. Ces résultats nous ont guidé dans le choix de nos hypothèses de travail quant aux contenus géométriques et aux variables à prendre en compte dans la formation des enseignants, mais aussi dans le choix des situations d'enseignement/apprentissage de la Géométrie.

Le travail de formation des enseignants que nous avons entrepris ensuite prend en compte trois aspects qui nous paraissent importants:

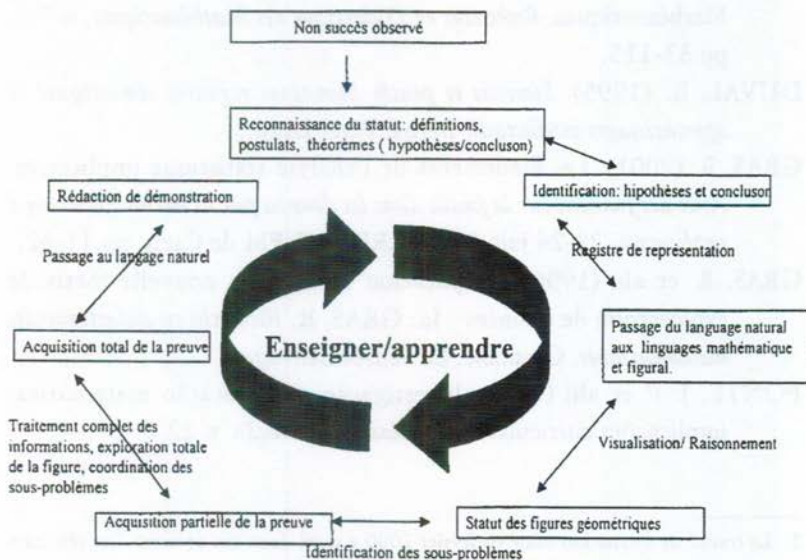
- faire un travail sur les savoirs et les savoir-faire en Géométrie, ayant pour objectif la formation des enseignants participant au projet de recherche; nous faisons l'hypothèse que cette formation leur permettra, tout au moins en partie, de s'approprier de certains savoirs et connaissances géométriques en favorisant un contrôle significatif de ceux-ci au moment de leur enseignement/apprentissage;
- faire un travail de formation intégrant certains résultats de la Didactique des Mathématiques et ayant pour objectif la construction d'instruments d'analyse des situations didactiques que ces enseignants sont amenés à développer en classe. Nos observations et celles de divers chercheurs montrent que, de façon générale, les enseignants ont un discours prenant en compte certains des résultats des recherches sur l'enseignement et l'apprentissage des mathématiques. Mais, ils semblent

éprouver de grandes difficultés à prendre ce discours en compte dans la construction et l'expérimentation de situations de classe;

- une étude des pratiques enseignantes par l'équipe de recherche et une analyse réflexive et constructive par les enseignants de leurs pratiques en classe. Cette analyse se fait (suivant l'idée de Robert, 2001) du point de vue de la construction des connaissances géométriques proposée aux élèves, et en ne tenant compte que des aspects épistémologiques et cognitifs (id., ibid., p. 65). Nous utilisons les expressions "pratiques enseignantes" et "pratiques en classe" suivant le sens d'Aline Robert (id., ibid.):

Nous réservons l'expression pratiques enseignantes à l'ensemble des activités de l'enseignant qui aboutissent à ce qu'il met en œuvre en classe.

Les pratiques en classe désignent tout ce que dit et fait l'enseignant en classe, en tenant compte de sa préparation, de ces conceptions et connaissances en mathématiques et de ses décisions instantanées, si elles sont conscientes. (Id., ibid., p. 66)



Le travail de formation¹ que nous avons mené ensuite s'appuie essentiellement sur les niveaux de compréhension de la Géométrie de Van Hiele (niveau de visualisation, niveau d'analyse, niveau de déduction formelle, niveau de la rigueur) et la théorie des registres de représentation sémiotique de Duval (1995), mettant en jeu l'importance de la coordination de différents registres de représentation sémiotique, du rôle de la figure dans la résolution des problèmes de Géométrie, de la lecture et l'interprétation des textes mathématiques, de la constitution d'un réseau sémantique des objets mathématiques et des théorèmes (et/ou définitions) qui peuvent être utilisées dans une démonstration. Le schéma ci-dessus résume le processus de construction des savoirs et des connaissances géométriques que nous avons mis en place chez les enseignants.

Bibliographie

- BRASIL. SECRETARIA DE EDUCAÇÃO (1998). *Parâmetros curriculares nacionais: Ensino Fundamental – Matemática*. Brasília, MEC, SEF.
- BROUSSEAU, G. (1986). Fondements et méthodes de la Didactique des Mathématiques. *Recherches en Didactique des Mathématiques*, v. 7.2, pp.33-115.
- DUVAL, R. (1995). *Sémiosis et pensée humaine: registres sémiotiques et apprentissages intellectuels*. Berna, Peter Lang.
- GRAS, R. (2001). Les fondements de l'analyse statistique implicative. *Actes des Journées sur la fouille dans les données par la méthode d'analyse implicative*, 23-24 juin 2000, ARDM-IUFM de Caen, pp.11-32.
- GRAS, R. et alii (1996). "Implication statistique: nouvelle méthode exploratoire de données". In: GRAS, R. *Recherche en didactique des mathématiques*. Grenoble, La Pensée Sauvage.
- PONTE, J. P. et alii (1998). Investigaç o em educaç o matem tica: implicaç es curriculares. *Ci ncias da Educaç o*, v. 22.

1 Le travail de formation initi  en f vrier 2000 a dur  deux ans et demi. Les r sultats scientifiques de ce projet feront l'objet d'autres publications.

- ROBERT, A. (2001). Les recherches sur les pratiques des enseignants et les contraintes de l'exercice du métier d'enseignant. *Recherches en Didactique des Mathématiques*, 21/1.2, pp. 57-80.
- ROBERT, A. e ROBINET, J. (1989). Représentation des enseignants de mathématiques sur les mathématiques et leur enseignement. *Cahiers de DIDIREM*. IREM de Paris VII, v. 1.
- SAEB (1995). *Sistema Nacional de Avaliação da Educação Básica*. Brasília, Secretaria de Desenvolvimento, Inovação e Avaliação Educacional, Instituto Nacional de Avaliação de Estudos e Pesquisas Educacionais.
- SECRETARIA DA EDUCAÇÃO DO ESTADO DE SÃO PAULO (1996). *Experiências Matemáticas: 7ª série*. São Paulo, SE/Cenp (versão preliminar).
- VAN HIELE, P. (1980). *Levels of Thinking: How to meet them, How to avoid them*. 58º ENCONTRO ANUAL DO NCTM, Seattle (paper).
- _____ (1986). *Structure and insight: to Theory of mathematics Education*. London, Academic Press.

Recebido em mar./2002; aprovado em abr./2002

Annexes

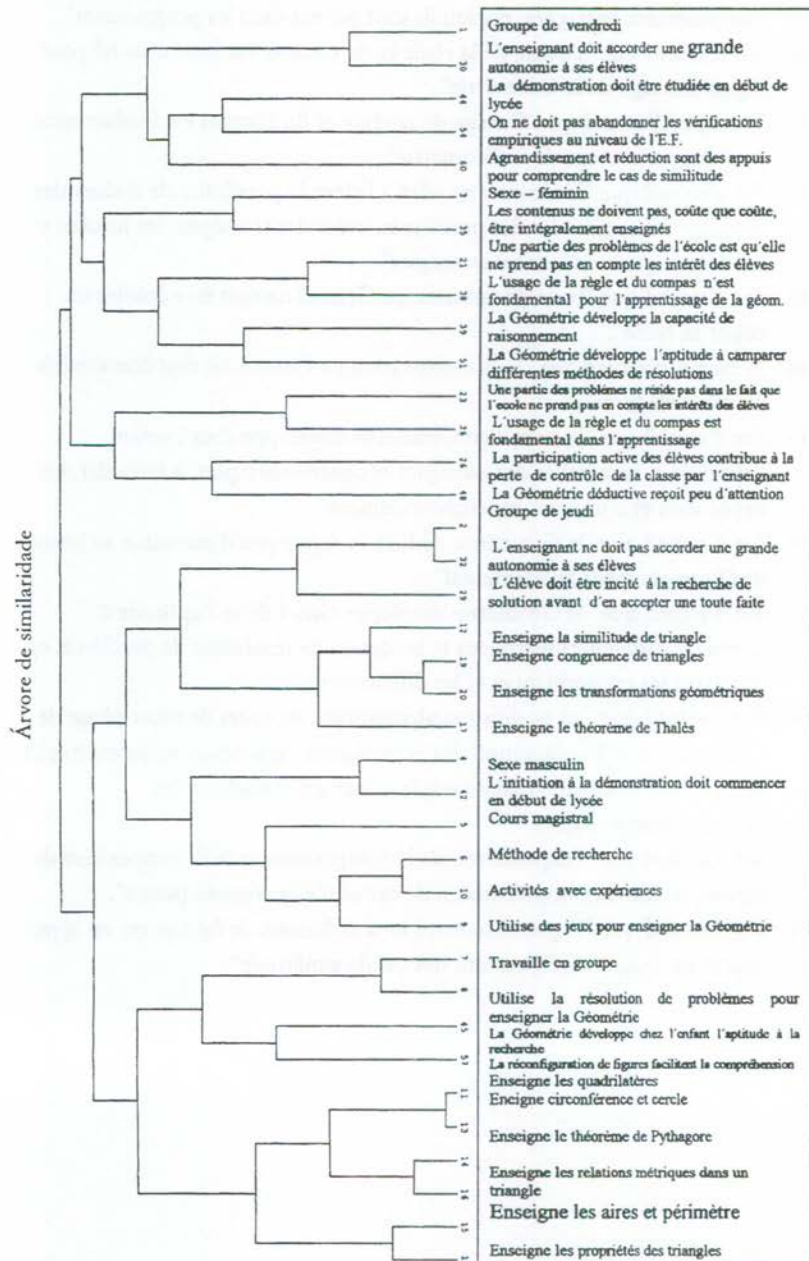
Annexe 1

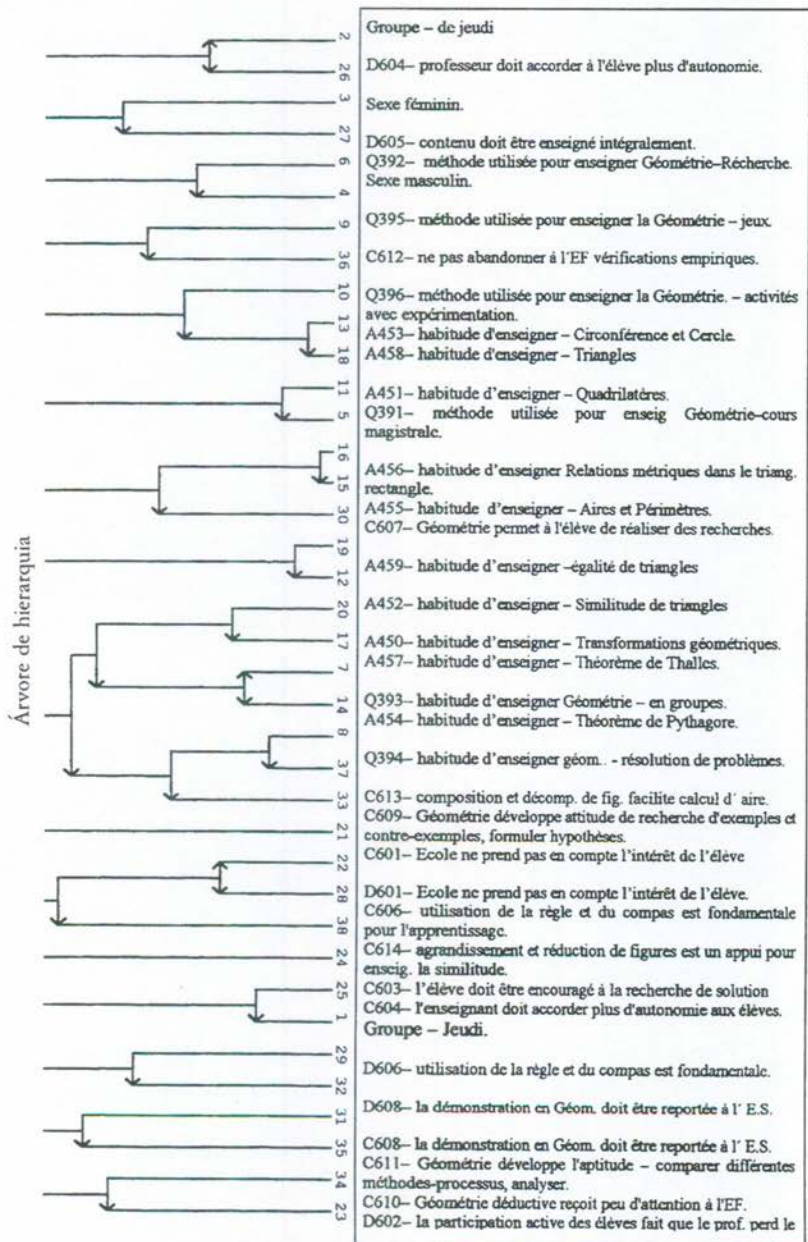
Variables statistiques de l'arbre de similarité de la hiérarchie implicativa

1. Groupe de vendredi: enseignants avec lesquels nous travaillons le vendredi
 2. Groupe de jeudi: enseignants avec lesquels nous travaillons le jeudi
 3. Sexe féminin
 4. Sexe masculin
 5. Cours magistral
 6. Méthode de recherche
 7. Travail en groupes
 8. Résolution de problèmes
 9. Jeux
 10. Activités avec expériences
 11. Quadrilatères
 12. Similitudes de triangles
 13. Circonférence et Cercle
 14. Théorème de Pythagore
 15. Aires et périmètres
 16. Relations métriques – triangle
 17. Théorème de Thalès
 18. Triangles
 19. Congruence de triangles
 20. Transformations géométriques
 21. Est d'accord que "une partie des problèmes de l'école est qu'elle ne prend pas en compte les intérêts des élèves".
 23. N'est pas d'accord que "une partie des problèmes de l'école est qu'elle ne prend pas en compte les intérêts des élèves".
 26. Est d'accord que "la participation active des élèves contribue à la perte de contrôle de la classe par l'enseignant".
 29. Est d'accord que "en classe, l'élève doit être incité à la recherche de solution d'un problème avant d'en accepter une toute faite".
 30. Est d'accord que "l'enseignant doit laisser à ses élèves une grande autonomie dans la construction de ses connaissances".
 32. N'est pas d'accord que "l'enseignant doit laisser à ses élèves une grande autonomie dans la construction de ses connaissances".
- Sous quelle forme enseignent-ils?
- Quels contenus enseignent-ils ?

35. N'est pas d'accord que "les contenus doivent être, coûte que coûte, intégralement enseignés, puisqu'ils sont prévus dans les programmes".
36. Est d'accord que "l'usage de la règle et du compas est fondamental pour l'apprentissage de la Géométrie".
38. N'est pas d'accord que "l'usage de la règle et du compas est fondamental pour l'apprentissage de la Géométrie".
39. Est d'accord que "la Géométrie offre à l'élève la possibilité de réaliser des investigations, résoudre des problèmes, créer des stratégies, les justifier et avoir des arguments sur ses stratégies".
42. Est d'accord que "la démonstration en Géométrie doit être étudiée en début de lycée".
44. N'est pas d'accord que "la démonstration en Géométrie doit être étudiée en début de lycée".
45. Est d'accord que "le travail en Géométrie développe chez l'enfant l'aptitude à rechercher des exemples et contre-exemples, à formuler des hypothèses et à prouver expérimentalement".
48. Est d'accord que "la Géométrie déductive reçoit peu d'attention au niveau de l'Enseignement Fondamental".
51. Est d'accord que "la Géométrie développe chez l'élève l'aptitude à comparer différentes méthodes et processus de résolution de problème, en analysant les ressemblances et les différences".
54. Est d'accord que "on ne doit pas abandonner, au cours de toute phase de l'Enseignement Fondamental, les vérifications empiriques de propriétés et de relations, mais qu'on doit aussi favoriser un travail sur des démonstrations simples".
57. Est d'accord que "l'exploitation de la composition et la décomposition de figures, facilite la compréhension de calcul d'aires figures planes".
60. Est d'accord que "l'agrandissement et la réduction de figures est un appui important pour l'enseignement des cas de similitude".

Annexe 2





A implicação estatística usada como ferramenta em um exemplo de análise de dados multidimensionais

RÉGIS GRAS* e SADDO AG ALMOULOU**

Resumo

O artigo apresenta os resultados de um questionário, cujo objetivo é analisar as opiniões dos participantes de um colóquio organizado pela PUC-SP sobre a implicação estatística de análise de dados multidimensionais. O colóquio foi realizado no intuito de estudar as potencialidades, a pertinência e a importância da realização de uma análise implicativa nas investigações das Ciências da Educação. A implicação estatística foi a principal ferramenta para analisar as informações envolvidas no questionário.

Palavras-chave: implicação estatística; metodologia de análise de dados; hierarquia de similaridade; hierarquia implicativa; avaliação.

Resumé

L'article présente les résultats d'un questionnaire dont l'objectif est l'étude des opinions des participants d'un colloque organisé à la Pontificia Universidade Católica de São Paulo, sur l'analyse implicative de données statistiques multidimensionnelles. Le colloque a pour objectif d'étudier les potentialités, la pertinence et l'importance de l'analyse implicative pour les recherches en sciences de l'éducation. L'implication statistique a été l'outil principal de traitement et d'analyse des données issues de ce questionnaire.

Mots-clés: implication statistique; méthodologie d'analyse de données; hiérarchie de similarité; hiérarchie implicative; évaluation.

Abstract

The article presents the results of a questionnaire whose purpose is to analyse the opinions of the participants in a seminar organised by PUC-SP about the statistical implication of multidimensional data analysis. The seminar aimed at investigating the potentialities, the pertinence and importance of carrying out implicative analysis in the investigations in the area of Education Sciences. Statistical implication was the main tool to analyse the information contained in the questionnaire.

Key-words: statistical implication; data analysis methodology; similarity hierarchy; implicative hierarchy; evaluation.

* Professeur Emérite à l'École Polytechnique de l'Université de Nantes, La Chantrerie, BP 50609 44306 Nantes cedex 03. E-mail: regisgra@club-internet.fr

** Professor do Programa de Estudos Pós-Graduados em Educação Matemática - PUC/SP. E-mail: saddoag@pucsp.br.

Introdução

Este artigo discute os resultados de um questionário, cujo objetivo é avaliar o colóquio intitulado "O método estatístico implicativo utilizado em estudos qualitativos de regras de associação: contribuição à pesquisa em educação", realizado em julho de 2003 no Programa de Estudos Pós-Graduados em Educação Matemática da PUC-SP.

O colóquio tem por objetivo principal realizar um estudo de refinamento sobre as pontencialidades, a pertinência, bem como a importância da realização de análises estatísticas de dados multidimensionais (análise hierárquica de similaridade, análise implicativa) nas investigações da Educação Matemática e em investigações em Educação de um modo mais abrangente. Foram previstos, para o desenvolvimento do trabalho, três momentos:

1. Realização de estudo das fases fundamentais de uma análise de dados multidimensionais (instrumentos de coleta de dados, organização e exploração, instrumentos de tratamentos, interpretação, levando em conta a questão e os objetivos da pesquisa);

2. Realização de um breve estudo de caráter teórico e intuitivo sobre os diferentes métodos de análise de dados multidimensionais e oferta de oficinas sobre os métodos;

3. Realização de estudo dos exemplos de dados encontrados em pesquisas desenvolvidas no Programa de Pós-graduação em Educação Matemática e em Educação: Currículo da PUC-SP;

4. Organização de um Colóquio sobre os métodos, aberto a distintos Programas de Pós-Graduação da PUC-SP, bem como de outras Universidades.

O colóquio contou com a participação ativa do professor emérito da Escola Politécnica da Universidade de Nantes: Régis Gras. O referido autor e seu grupo de pesquisa, desde 1979, procuraram, entre outros assuntos, colocar à disposição dos pesquisadores (em matemática, em psicologia, em biologia, em educação, etc.) ferramentas estatísticas (a análise implicativa, a hierarquia implicativa) que permitem evidenciar a dinâmica dos comportamentos de sujeitos (alunos, por exemplo) em situação de resolução de problemas, no caso da Educação Matemática. A análise implicativa (cf. o artigo de Gras nesta revista), como todos os

métodos de análise estatística de dados multidimensionais, permite visualizar, organizar, construir modelos e explicar fenômenos associados aos dados.

Estrutura do questionário

Para a constituição do corpo de informações a ser analisado foi aplicado um questionário-avaliação aos participantes do colóquio. O questionário (cf. Anexo 1), além de informações pessoais, solicitava que os participantes respondessem a quatro questões subjetivas, que objetivavam fornecer uma visão de como eles vivenciaram a programação e os conteúdos dos trabalhos no evento.

As questões subjetivas foram as seguintes:

1) *Suas expectativas foram satisfeitas?*

Sim Parcialmente Não

2) *Quais pontos da programação despertaram mais seu interesse?*

3) *Quais aspectos você desejaria aprofundar mais?*

4) *Na lista abaixo, escolha a(s) expressão(ões) que melhor expressa(m) seu sentimento, a propósito do colóquio.*

muito longo muito curto cansativo estimulante

curiosidade muito bom inútil abertura

difícil gostaria fazer mais

As respostas foram codificadas (cf. Anexo 1) e as variáveis estatísticas que dizem respeito aos dados pessoais foram consideradas como variáveis suplementares (o código dessas variáveis é do tipo X s).

Analizamos as duas questões abertas (“Quais os pontos da programação despertaram mais seu interesse?” e “Quais aspectos você desejaria aprofundar mais?”), destacando as palavras mais significativas das respostas e essas palavras foram agrupadas por sinonímia. Obtivemos, assim, 6 variáveis para a primeira questão e 4, para a segunda, retomando, às vezes, as alegações que eram comparáveis de uma questão a outra.

Mesmo sabendo que o número de questionários esteja relativamente fraco, algumas estruturas interessantes são interpretáveis, sem que seu caráter explicativo seja de uma fidedignidade absoluta; por consequência, precisa ser confirmada por outro estudo. Nós nos contentaremos, então, em destacar as tendências e os pontos mais assegurados.

A tabela do Anexo 2 apresenta as ocorrências, as médias e os desvios-padrão das variáveis estatísticas estudadas. Observa-se que as variáveis que tiveram maiores ocorrências são: XSS (expectativas satisfeitas), EST (o trabalho foi estimulante) e IMA (manipulação de CHIC). Essas três variáveis expressam o grau de satisfação dos participantes e a pertinência da metodologia de trabalho adotada.

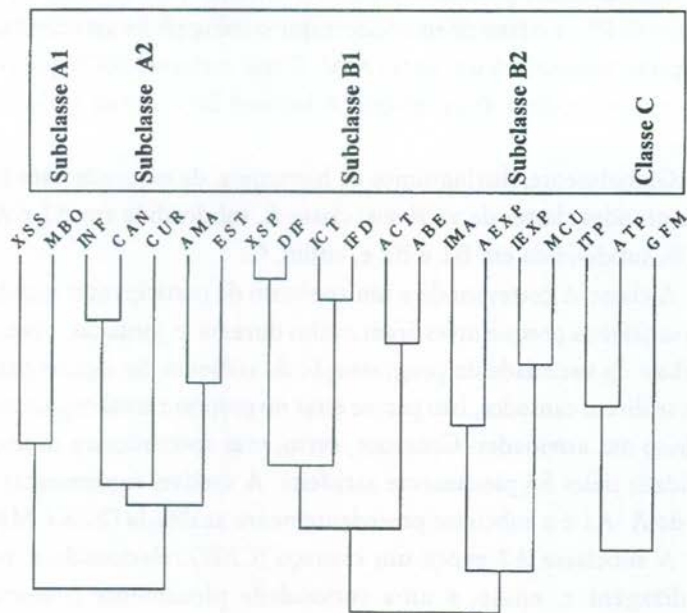
Análise de similaridades segundo I. C. Lerman

Como em todos os métodos de classificação, procuramos constituir, em um conjunto V de variáveis, partições de V cada vez menos finas, construídas de maneira ascendente. Essas partições encaixadas são representadas por uma árvore construída usando um critério de similaridade ou de semelhança estatística entre variáveis. A similaridade se define a partir do cruzamento do conjunto V das variáveis com um conjunto E de sujeitos (ou de objetos). Esse tipo de análise permite ao usuário estudar e interpretar, em termos de tipologia e de semelhança (e não semelhança) decrescente, classes de variáveis constituídas significativamente a certos níveis da árvore e se opondo a outras nesses mesmos níveis.

O *critério de similaridade* se exprime da maneira seguinte, nos casos das variáveis binárias (presença – ausência, verdadeiro – falso, sim – não, etc.): 2 variáveis a e b , satisfeitas respectivamente por dois subconjuntos A e B de E , são muito semelhantes quando o número k dos sujeitos de $A \cap B$ é importante de um lado, pelo que teria sido no caso da ausência de ligação entre a e b , e por outro lado, com relação aos cardinais de E , A e B . Medimos essa semelhança pela probabilidade de que k seja superior ao número aleatório esperado nesta situação.

A modelagem probabilista da variável aleatória, cujo k é a realização, pode ser uma distribuição *binomial* ou de *Poisson*, à escolha do usuário. A segunda distribuição supõe que E seja uma amostra de uma população grande, o que a primeira não supõe. Se E não tem nenhuma razão estatística, *a priori*, de ser representativo, é preferível usar o modelo binomial que analisa a estrutura de E enquanto tal. Quando os parâmetros o permitem, uma aproximação gaussiana dessas duas leis é efetuada.

O índice de similaridade entre variáveis é usado para definir um índice de similaridade entre duas classes de variáveis, segundo o princípio de comparação entre a observação e o que seria dado pelo acaso. Um índice, dito de coesão, permite reagrupar as classes de variáveis.



Árvore de similaridade

Assim, para construir uma árvore de similaridade (cf. árvore acima), reunimos em uma classe de primeiro nível, primeiramente, as 2 variáveis que são mais similares, no sentido do índice de similaridade, depois 2 outras variáveis ou uma variável e a classe já formada, no sentido do índice da classe e, depois, outras variáveis ou classes de variáveis.

Interpretação sucinta

No primeiro nível aparece uma relação de similaridade entre as únicas duas variáveis, mais ou menos restritivas: "expectativas parcial-

mente satisfeitas” e “difícil”. O ritmo do colóquio pôde explicar o cansaço, tanto mais que a carga teórica era muito importante. Identificamos aqui, talvez, o desequilíbrio entre aportes teóricos e ilustrações. Mas a representação desse fenômeno é muito minoritária.

Sublinhamos, por outro lado e ao contrário, que a relação entre satisfação (84%) e o fato de reconhecer que o colóquio foi satisfatório, de modo geral, aparece em um outro nível. É esse caráter (MBO) que consideramos como indício de satisfação. A variável EF é a mais típica dessa relação.

Globalmente, distinguimos na hierarquia, da esquerda para direita, três grandes classes de variáveis: classe A, subdividida em A1 e A2, a classe B, subdividida em B1 e B2 e, enfim, C.

A classe A corresponde a um conjunto de participantes que estão muito satisfeitos porque investiram muito durante as jornadas, apesar da extensão e da variedade da programação do colóquio. Se alguns participantes se dizem cansados, isto parece estar no próprio e total engajamento do sujeito nas atividades. Cansados, certo, mas apreenderam muito e a curiosidade deles foi plenamente satisfeita. A variável suplementar G é típica de A. A1 é a subclasse precedentemente analisada (XSS e MBO).

A subclasse A2 evoca um cansaço (CAN) relacionado a novas aprendizagens e, então, a uma curiosidade plenamente preenchida (tipicalidade G).

A classe B contém as nuances explicando a percepção da dificuldade e, então, da satisfação parcial:

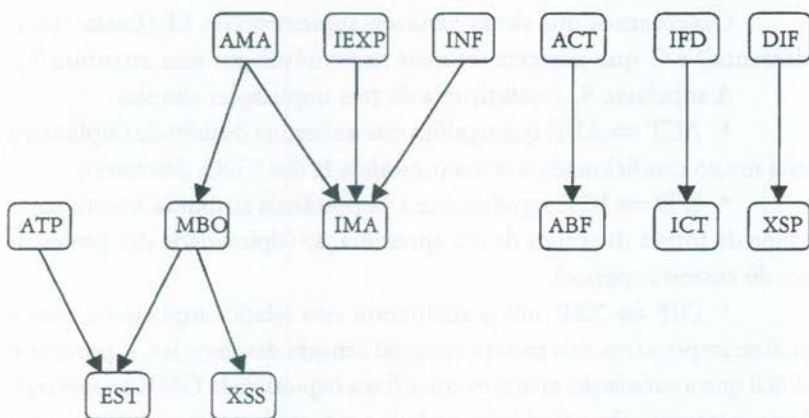
- a subclasse B1 é justamente a já analisada (XSP e DIF). GME é a variável a mais típica desta subclasse e, globalmente, de B.
- a subclasse B2 explica que a origem da restrição vem da complexidade teórica, o que não impede que haja demanda posterior (ACT) de mais atividades sobre o método, em razão das possibilidades (ABE) que a teoria permite. A variável M é a mais típica.

A classe C revela a satisfação e uma nova demanda (GFM), de uma programação na qual a teoria e prática (ITP e ATP) seriam intimamente associadas. GMA é a variável a mais típica.

Análise Implicativa

Análise do grafo implicativo obtido

Consideramos a construção do grafo implicativo no valor mínimo de 0,70 que permite apresentar uma estrutura interessante e, ainda, significativa do ponto de vista estatístico. É possível diminuir esse valor, a fim de obter mais relações implicativas, relações essas se tornando ainda mais fracas. Mas pode-se igualmente aumentar esse número para conservar as implicações mais fortes.



Grafo implicativo

Interpretação sucinta

1ª Hierarquia coesitiva observada

Duas subestruturas aparecem, sem relação entre si.

A subclasse S1 corresponde à satisfação total (XSS), a subclasse S2 à satisfação parcial (XSP). Essas duas variáveis são colocadas nas extremidades dos caminhos implicativos, significando, portanto, que elas são mais as conseqüências dos sentimentos expressos.

A subclasse S1 reúne dois caminhos:

O primeiro corresponderia, antes de tudo, a futuras expectativas; o segundo corresponderia à satisfações de expectativas *a priori*.

A satisfação dessas últimas provém da manipulação de CHIC (IMA) e, particularmente, da descoberta de suas novas funcionalidades (INF) (EM variável típica), mas, igualmente, das aplicações e experiências associadas (IEXP). São essas relações que deram sentido à manipulação (IEXP \Rightarrow IMA) (G é a variável mais típica).

Notamos, portanto, que é a variável AMA que articula os dois caminhos. Pode-se interpretá-la de modo contraposto, assim: “se não estivesse satisfeito com as manipulações, então não desejaria ter outras, na ocasião de um próximo colóquio”.

O segundo mostra que a satisfação total (MBO) está muito ligada ao caráter estimulante das atividades (EST).

Observamos que são as variáveis suplementares EF (Ensino Fundamental) e G que parecem ser mais responsáveis por essa estrutura S_1 .

A subclasse S_2 é constituída de três implicações simples:

- ACT \Rightarrow ABF: que significa que a abertura do método implicativo está muito condicionada à teoria (tipicalidade dos GDO, doutores);

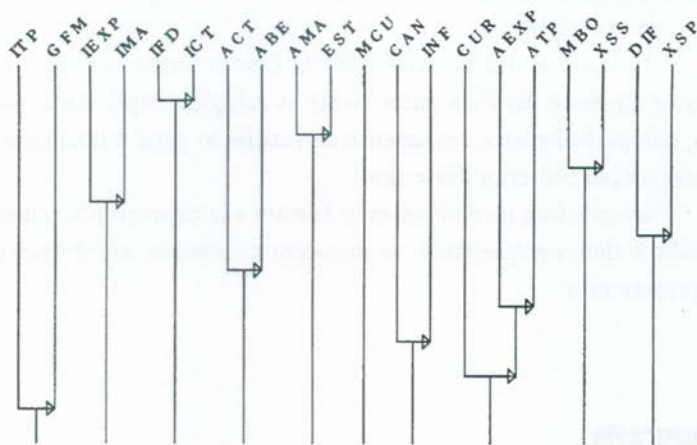
- IFD \Rightarrow ICT: significa que a importância atribuída à teoria teria vindo da forma dinâmica de sua apresentação (tipicalidade dos professores do ensino superior).

- DIF \Rightarrow XSP: nós já analisamos esta relação implicativa, mas a análise implicativa nos mostra em qual sentido devemos ler. É por que é difícil que a satisfação apareceu mais fraca (tipicalidade GME ou mestre). Uma interpretação psicológica poderia ser tentada aqui: os participantes que começaram a se engajar na pesquisa parecem ter dificuldade no caminho a percorrer, pois eles devem perceber os aportes teóricos como uma fronteira a ultrapassar para acessar ao estatuto de pesquisador. Este não é o caso dos mais jovens, não ainda implicados no uso do método, nem os mais avançados ou já iniciados.

2ª Hierarquia coesitiva observada

Observa-se que poucas grandes classes se formaram. Todas as classes, exceto uma, são constituídas de dois elementos. Esse fato é devido muito ao número relativamente importante de variáveis em relação ao número de sujeitos. Analisaremos, então, brevemente, cada uma dessas classes, respeitando a sua ordem de qualidade de coesão decrescente.

• IFD \Rightarrow ICT: a forma dinâmica e eficiente é fonte de interesse no que tange ao aporte teórico. Esta observação enfatiza a prudência com a qual um curso magistral deve ser pensado. Um curso não acompanhado de exemplo, nos quais os sujeitos encontram um domínio já conhecido e manipulado se torna desestimulante e sem grande efeito sobre a aprendizagem (tipicalidade de ES).



Árvore coesita

• AMA \Rightarrow EST: as futuras manipulações de CHIC poderiam ser estimulantes (tipicalidade de EM).

• MBO \Rightarrow XSS: a significação vem sobretudo do caráter total, completo (teoria – prática), adequado do colóquio (tipicalidade de EF).

• IEXP \Rightarrow IMA: as manipulações se justificam por que elas foram acompanhadas da experimentação não formal (tipicalidade de G).

• DIF \Rightarrow XSP: já foi analisada.

• ACT \Rightarrow ABE: já foi analisada.

• CAN \Rightarrow INF: lemos esta implicação de modo contraposto, pois ela é mais significativa nesse sentido: “se CHIC não tivesse apresentado novas funcionalidades, então não teríamos terminado tão cansados”. Sabe-se bem que são os desequilíbrios dos conhecimentos que são dolorosamente ressentidos (tipicalidade de GDA).

• $CUR \Rightarrow (AEXP \Rightarrow ATP)$: esta relação é estatisticamente significativa. Assim, a curiosidade será tão satisfeita quanto a regra: “as experiências permitem esclarecer a relação teórico-prática” também será. O sujeito é curioso de ver como a teoria é utilizada como ferramenta para a prática (tipicalidade de GDA).

• $ITP \Rightarrow GFM$: as demandas posteriores são principalmente induzidas pelo interesse em trabalhar mais a relação teórico-prática (tipicalidade de EM).

Observamos que a variável MCU (tempo muito curto) é neutra como se ela fosse dividida entre todas as relações implicativas. Isto é, claro, compatível com o sentimento de satisfação geral e uma demanda de intervenção posterior desse tipo.

Para concluir, incentivamos os leitores a interpretar por si mesmos os gráficos dados esquecendo, se julgaram necessário, as interpretações que propusemos.

Bibliografia

- ALMOULOU, S. A. (1992). *L'ordinateur, outil d'aide à l'apprentissage de la démonstration et de traitement de données didactiques*. Thèse. Université de Rennes 1.
- BODIN, A. (1997). Modèles sous-jacents à l'analyse implicative et outils complémentaires. *Prépublication IRMAR*, n. 97-32.
- BODIN, A. et GRAS, R. (1999). Analyse du préquestionnaire enseignants avant EVAPM-Terminales. *Bulletin de l'Association des Professeurs de Mathématiques de l'Enseignement Public*, n. 425, pp. 772-786.
- COUTURIER, R. et GRAS, R. (1999). Introduction de variables supplémentaires dans une hiérarchie de classes et application à CHIC. *Actes des 7èmes Rencontres de la Société Francophone de Classification*, Nancy, 15-17 septembre.
- GRAS, R. (1979). *Contribution à l'étude expérimentale et à l'analyse de certaines acquisitions cognitives et de certains objectifs didactiques en mathématiques*. Thèse d'Etat. Université de Rennes I.

- GRAS, R. et LARHER, A. (1992). L'implication statistique, une nouvelle méthode d'analyse de données. *Mathématique, Informatique et Sciences Humaines*, n. 20, pp. 5-31.
- GRAS, R. et RATSIMBA-RAJOHN, H. (1996). Analyse non symétrique de données par l'implication statistique. *RAIRO-Recherche Opérationnelle*, 30-3, AFCE, pp. 217-232.
- GRAS, R.; BRIAND, H. et PETER, P. (1996). Structuration sets with implication intensity. In: DIDAY, E.; CHEVALLIER, Y. et OPITZ, O. (eds.). *Proceedings of the International Conference on Ordinal and Symbolic Data Analysis – OSDA 95*, Paris, Springer, pp. 147-156.
- GRAS R. et alii (1996). *L'implication Statistique*. Collection Associée à "Recherches en Didactique des Mathématiques". Grenoble, La Pensée Sauvage.
- GRAS R.; BRIAND H.; PETER P. et PHILIPPE J. (1997). Implicative statistical analysis. *Proceedings of International Congress I.F.C.S.*, 96. Kobé, Tokyo, Springer-Verlag, pp. 412-419.
- GRAS R.; KUNTZ P.; COUTURIER R. et GUILLET F. (2001). Une version entropique de l'intensité d'implication pour les corpus volumineux. *Extraction des Connaissances et Apprentissage (ECA)*, v. 1, n. 1-2, pp. 69-80.
- GRAS, R.; KUNTZ, P. et BRIAND, H. (2001). Les fondements de l'analyse statistique implicative et quelques prolongements pour la fouille de données. *Mathématiques et Sciences Humaines*, n. 154-155, pp. 9-29.
- GRAS, R.; DIDAY, E.; KUNTZ, P. et COUTURIER, R. (2001). Variables sur intervalles et variables-intervalles en analyse statistique implicative. *Actes du 8^{ème} Congrès de la Société Francophone de Classification*. Université des Antilles-Guyane, Pointe-à-Pitre, 17-21 décembre, pp. 166-173.
- GRAS, R.; GUILLET, F.; GRAS, R. et PHILIPPE, J. (2002). Réduction des colonnes d'un tableau de données par quasi-équivalence entre variables. *Extraction des connaissances et apprentissage, Hermès*, v. 1, n. 4, pp. 197-202.
- GRAS, R.; KUNTZ, P. et BRIAND, H. (2003). Hiérarchie orientée de règles généralisées en analyse implicative. *Extraction des Connaissances et apprentissage, Hermès*, pp. 145-157.

LERMAN, I. C.; GRAS, R. et ROSTAM, H. (1981). Elaboration et évaluation d'un indice d'implication pour des données binaires, I et II. *Mathématiques et sciences Humaines*, n. 75, pp. 5-47.

Recebido em mar./2002; aprovado em abr./2002

Anexos

Anexo 1

1ª Parte: Identificação (variáveis suplementares)

Data: _____

- 1) Sexo: Masculino M s Feminino F s
- 2) Em que grau(s) de ensino leciona neste ano?
 Ensino Fundamental EF s
 Ensino Médio EM s
 Ensino Superior ES s
 Pós-Graduação PG s
- 3) Qual a sua formação acadêmica?
 Graduado G s
 Pós-Graduado – Mestrando GMA s
 Pós-Graduado – Mestre GME s
 Pós-Graduado – Doutorando GDA s
 Pós-Graduado – Doutor GDO s

2ª Parte: Opinião (variáveis principais)

- 4) Suas expectativas foram satisfeitas:
 Sim XSS Parcialmente XSP Não XSN
- 5) Quais os pontos da programação despertaram mais seu interesse?
 teoria, conceitual ICT novas funcionalidades de CHIC INF
 manipulação de CHIC IMA relação teoria-prática ITP
 aplicações, experiências IEXP forma dinâmica eficaz IFD
- 6) Quais aspectos você desejaria aprofundar mais?
 teoria, conceitual ACT manipulação de CHIC AMA
 relação teoria-prática ATP aplicações, experiências AEXP
- 7) Na lista abaixo, escolha a(s) expressão(ões) que melhor expressa(m) seus sentimentos a propósito do colóquio.
 muito longo MLO muito curto MCU cansativo CAN
 estimulante EST curiosidade CUR muito bom MBO
 inútil INU abertura ABE difícil DIF
 gostaria fazer mais GFM

Anexo 2

	Ocorrência	Média	Desvio padrão
XSS	21.00	0.84	0.37
XSP	4.00	0.16	0.37
ICT	10.00	0.40	0.49
INF	5.00	0.20	0.40
IMA	17.00	0.68	0.47
ITP	3.00	0.12	0.32
IEXP	11.00	0.44	0.50
IFD	4.00	0.16	0.37
ACT	9.00	0.36	0.48
AMA	8.00	0.32	0.47
ATP	16.00	0.64	0.48
AEXP	3.00	0.12	0.32
MLO	0.00	0.00	0.00
MCU	4.00	0.16	0.37
CAN	1.00	0.04	0.20
EST	18.00	0.72	0.45
CUR	11.00	0.44	0.50
MBO	13.00	0.52	0.50
INU	0.00	0.00	0.00
ABE	10.00	0.40	0.49
DIF	2.00	0.08	0.27
GFM	19.00	0.76	0.43
M s	6.00	0.24	0.43
F s	19.00	0.76	0.43
EF s	4.00	0.16	0.37
EM s	10.00	0.40	0.49
ES s	15.00	0.60	0.49
PG s	4.00	0.16	0.37
G s	4.00	0.16	0.37
GMA s	9.00	0.36	0.48
GME s	4.00	0.16	0.37
GDA s	8.00	0.32	0.47
GDO s	1.00	0.04	0.20

Possíveis relações entre o processo de formação e a prática pedagógica: proposta de procedimentos de pesquisa

ANA LÚCIA MANRIQUE* e MARLI E. D. A. ANDRÉ**

Resumo

Esse estudo busca identificar possíveis relações entre o processo de formação profissional, proporcionado pelo Proformação e a prática pedagógica do professor cursista. Algumas das relações foram estudadas a partir de informações obtidas pela análise do desempenho dos professores em um teste de simulação didática – instrumento organizado na forma de questões que simulam situações didático-pedagógicas do cotidiano escolar, referente às áreas: ensino de português, ensino de matemática e didática. O teste de simulação didática foi aplicado em professores cursistas de cinco Agências Formadoras (AGF) do Proformação – Guandu e Feira de Santana, no estado da Bahia, Rio Branco, no Acre, Quixadá e Massapê, no Ceará – em dois momentos do processo de formação. Inicialmente, apresentamos um breve perfil dos professores cursistas, que responderam ao teste no primeiro momento de aplicação, em relação a sexo, idade e série em que lecionam. Em seguida, realizamos uma análise hierárquica de similaridade com as informações obtidas na aplicação de um programa de tratamento de dados estatísticos multidimensionais – CHIC, nos resultados do teste de simulação didática, apenas do primeiro momento. Concluímos que o uso desse instrumento – CHIC – enriqueceu a análise das informações, oferecendo o recurso de visualizar relações existentes entre o processo de formação e a prática pedagógica de professores.

Palavras-chave: formação de professores; análise de similaridade; Proformação.

Abstract

The objective of this study is to identify possible relationships between the professional training provided by the Brazilian Ministry of Education's teacher-training program Proformação and the practice of teachers who are being trained. Certain relationships were studied using information obtained in the analysis of teacher performance by means of a simulated teaching test – a questionnaire that simulates everyday classroom situations, dealing with: the teaching of Portuguese, the teaching of Mathematics and pedagogical approach. Teachers attending five Proformação Teaching Agencies

* Professora da PUC-SP e da USJT. E-mail: manrique@pucsp.br.

** Professora do Pós-Graduação em Educação: Psicologia da Educação – PUC-SP.
E-mail: marliandre@pucsp.br.

(AGF) – Guandu and Feira de Santana, in the state of Bahia; Rio Branco, in the state of Acre; Quixadá and Massapê, in the state of Ceará – underwent the simulated teaching test at two different stages in their professional training. First, we elaborated a brief profile of the teachers, who answered the test at the time of its application, regarding sex, age and the grade they teach. Then, we conducted an initial hierarchical similarity analysis between the information obtained and the results of the simulated teaching test using a multidimensional statistical data analysis software (CHIC). We believe that the use of the software enriched the analysis, enabling us to visualize relationships between professional training and teaching practice.

Key-words: teacher training; similarity analysis; Proformação.

Introdução

O Projeto Proformação – Programa de Formação de Professores em Exercício – foi desenvolvido pelo MEC (Fundescola-SEED/MEC), em parceria com estados e municípios das regiões Norte, Nordeste e Centro-Oeste do Brasil. É um curso de nível médio, com habilitação em magistério e tem por metas:

- a) titular em nível médio, com habilitação para o magistério;
- b) oferecer condições técnicas para a melhoria da prática pedagógica;
- c) contribuir para a melhoria do processo ensino-aprendizagem e a elevação do desempenho escolar do aluno;
- d) valorizar o magistério pelo resgate da melhoria da qualidade do ensino.

O curso desenvolveu-se na modalidade de educação à distância, com dois anos de duração, incluindo fases presenciais ao final de cada semestre e reuniões quinzenais aos sábados.

Uma avaliação externa, financiada pelo Programa das Nações Unidas para o Desenvolvimento (PNUD), buscou identificar possíveis relações entre o processo de qualificação profissional, veiculado pelo Proformação, e a prática pedagógica do professor cursista.

Para essa avaliação foram realizados estudos de caso que utilizaram, entre outros instrumentos, testes de simulação didática para avaliar o desempenho dos cursistas. Esse teste é composto de questões que simulam situações didático-pedagógicas do cotidiano escolar, referente ensino de português, ensino de matemática e didática. Foi aplicado em cursistas

de cinco AGFs (Agências Formadoras) do Proformação – Guandu e Feira de Santana, no estado da Bahia, Rio Branco, no Acre, Quixadá e Massapé, no Ceará – em dois momentos da implantação do Programa: no início e no final de 2001.

Primeira análise dos resultados dos testes de simulação

Um total de 135 professores cursistas respondeu o teste de simulação didática na primeira aplicação – momento 1 – e 120 professores, na segunda aplicação – momento 2. Cada um dos dois testes é composto por 10 questões referentes ao ensino de português, 10 ao ensino de matemática e 5 à didática, perfazendo um total de 25 questões.

O Quadro 1 apresenta o total de professores cursistas, distribuídos por AGF, que responderam às questões do referido teste.

Quadro 1 – Distribuição dos professores cursistas por estado e AGF

Estado	AGF	Total de professores	
		Momento 1	Momento 2
Bahia	Guandu	21	20
	Feira de Santana	20	20
Acre	Rio Branco	27	24
Ceará	Quixadá	38	31
	Massapé	29	25
Total		135	120

Verificamos que apenas em uma das AGFs da Bahia, Feira de Santana, manteve-se o mesmo número de professores (20) que responderam ao teste nos dois momentos propostos. Essa diferença decorre dos diferentes momentos de aplicação do teste. Percebemos, também, que em Quixadá e Massapé – AGFs pertencentes ao estado do Ceará – concentra-se a maior parte dos sujeitos avaliados, mesmo havendo uma diminuição de professores no momento 2.

Vale ressaltar que, na AGF de Quixadá, apenas dez professores que realizaram o teste no primeiro momento estiveram presentes na segunda aplicação. Assim, alguns professores foram substituídos por outros, durante o processo.

O Quadro 2 mostra a média de acertos das questões dos testes de simulação didática por AGF e por momento de aplicação.

Quadro 2 – Média de acertos das questões dos testes de simulação didática por AGF e por momento de aplicação

AGF	Acertos (total de questões: 25)				
	Momento 1		Momento 2		Variação da média
	Professores	Média	Professores	Média	
Guandu	21	10,2	20	17,3	7,1
Feira de Santana	20	13,9	20	18,4	4,5
Rio Branco	27	12,5	24	16,1	3,6
Quixadá	38	11,6	31	17,6	6,0
Massapé	29	11,8	25	16,8	5,0
Total	135	11,9	120	17,2	5,3

Verificamos que houve 47,6% de acertos (11,9 de 25 questões) no momento 1 da aplicação do teste e um aumento significativo desse índice no momento 2 (68,8% – 17,2 de 25 questões). Ao visualizar essas informações, notamos que a AGF de Feira de Santana (Bahia) obteve o maior índice de acertos entre todas as AGFs nos dois momentos de aplicação do teste; o momento 1 apresentou uma média de 13,9 de questões certas em um total de 25 e o momento 2, uma média de 18,4 de questões certas. As AGFs de Guandu, Massapé e Quixadá obtêm índices de acerto abaixo da média no primeiro momento da aplicação e as AGFs de Rio Branco e Massapé, no segundo momento.

Além disso, é possível perceber que a variação média mais significativa foi na AGF de Guandu, ou seja, os professores cursistas de Guandu foram os que conseguiram obter o maior acréscimo no índice de acertos das questões do momento 2 de aplicação do teste em relação ao do momento 1. E os professores de Rio Branco obtiveram a variação média menos significativa.

Os professores da AGF de Massapé obtiveram índices de acertos abaixo da média total, tanto no primeiro momento de aplicação do teste quanto no segundo momento. Entretanto, conseguiram obter a variação média próxima da variação média geral, o que mostra uma recuperação no desempenho, com um aumento do número de acertos.

Porém, apenas essas informações não são suficientes para percebermos se as metas do Proformação foram alcançadas. Para tentarmos identificar possíveis relações entre o processo de formação profissional, proporcionado pelo Proformação, e a prática pedagógica do professor cursista, delineamos, a seguir, um breve perfil dos professores cursistas que responderam ao teste no primeiro momento, em relação a sexo, idade e série em que lecionam. Depois, analisamos os resultados obtidos com a aplicação do software CHIC nos dados dos testes de simulação dos dois momentos.

O Quadro III apresenta a distribuição dos professores cursistas no momento 1, por AGF e por sexo.

Quadro 3 – Distribuição dos professores cursistas por AGF e por sexo

AGF	Sexo		Total
	Feminino	Masculino	
Guandu	21	0	21
Feira de Santana	18	2	20
Rio Branco	19	8	27
Quixadá	37	1	38
Massapé	26	3	29
Total	121	14	135

Os dados do Quadro 3 mostram que a grande maioria (89,6% – 121 professores) dos professores cursistas é do sexo feminino, sendo muito pequena a participação masculina (10,4% – 14 professores). Esses dados referendam a questão de gênero na profissão docente, havendo uma predominância do sexo feminino. Nota-se que 57,1% dos homens (8 de 14 professores) são da AGF de Rio Branco e que na AGF de Guandu todo o grupo é formado por mulheres.

O próximo quadro – Quadro 4 – apresenta dados sobre a faixa etária a que pertencem os professores que responderam ao teste de simulação no momento 1.

Quadro 4 – Distribuição dos professores cursistas por AGF e faixa etária

AGF	Faixa Etária						Total
	Não respondeu	Até 19 anos	20 a 29 anos	30 a 39 anos	40 a 49 anos	Acima de 50 anos	
Guandu	1	0	11	8	1	0	21
Feira de Santana	4	0	2	8	5	1	20
Rio Branco	2	0	5	10	4	6	27
Quixadá	1	0	1	15	16	5	38
Massapé	1	3	15	7	3	0	29
Total	9	3	34	48	29	12	135
Porcentagem	6,7%	2,2%	25,2%	35,6%	21,2%	8,9%	100%

No Quadro 4 observamos que a maioria dos professores cursistas tem idade entre 20 e 49 anos (82%), encontrando-se mais de um terço do total dos professores na faixa dos 30 aos 39 anos (35,6%). Notamos, também, que são poucos os professores com idade inferior a 20 e acima dos 50 anos.

A seguir, apresentamos o Quadro 5, que nos mostra para qual ciclo do ensino fundamental os professores cursistas estão lecionando. O 1º ciclo corresponde às 1ª e 2ª séries; o 2º ciclo, às 3ª e 4ª séries; o multisseriado nos informa que o professor possui em sua sala de aula alunos de diversas séries. Além disso, quando aparecerem os 1º e 2º ciclos juntos, isso significa que o professor leciona em dois períodos, sendo em um período para o 1º ciclo e no outro, para o 2º ciclo.

Quadro 5 – Distribuição dos professores cursistas por AGF e ciclo em que lecionam

AGF	Não respondeu	1º ciclo	2º ciclo	Multis-seriado	1º e 2º ciclos	Total
Guandu	6	7	3	4	1	21
Feira de Santana	0	7	5	7	1	20
Rio Branco	13	5	4	5	0	27
Quixadá	2	27	4	5	0	38
Massapé	0	23	6	0	0	29
Total	21	69	22	21	2	135
Porcentagem	15,6%	51,1%	16,3%	15,6%	1,5%	100%

Notamos, no Quadro 5, que um pouco mais da metade (51,1%) dos professores cursistas estão lecionando para as duas primeiras séries do ensino fundamental, ou seja, estão alfabetizando crianças. Observamos, também, que nas AGFs de Quixadá e Massapé, a grande maioria dos professores – 50 de 67 professores – está trabalhando com a alfabetização. Esclarecemos, ainda, que alguns dos professores classificados como não respondentes, na verdade, colocaram demitidos em sua resposta quanto ao ciclo que lecionavam.

Assim, por meio da análise dos dados organizados nos três quadros anteriores concluímos que o perfil dos professores cursistas é, na maioria, de pessoas do sexo feminino, com idade entre 30 e 39 anos e lecionando nas duas primeiras séries do ensino fundamental.

Nova análise dos dados

Apresentamos, a seguir, uma análise hierárquica de similaridade, por permitir realizar outras interpretações das informações obtidas com o primeiro teste de simulação didática. Para realizarmos a análise hierárquica de similaridade, utilizamos um software de tratamento de dados estatísticos multidimensionais desenvolvido no Instituto de Recherche Mathématique de Rennes (Irmarm) da Universidade de Rennes 1, denominado CHIC.

Para aplicar esse software, selecionamos apenas os professores que responderam aos dois testes aplicados nos dois momentos propostos. Dessa maneira, a análise realizada corresponde às respostas fornecidas por 93 professores das cinco AGFs e o Quadro 6 apresenta a distribuição dos professores por estado e AGF.

Quadro 6 – Distribuição dos professores cursistas por estado e AGF

Estado	AGF	Total de professores
Bahia	Guandu	20
	Feira de Santana	20
Acre	Rio Branco	21
Ceará	Quixadá	10
	Massapé	22
Total		93

Observamos que apenas a AGF de Quixadá apresentou uma diminuição significativa de professores cursistas, isto é, dos 38 professores que responderam o teste no momento 1, apenas 10 (26,3%) estavam presente no momento 2.

A análise hierárquica de similaridade foi realizada apenas para os dados obtidos com a aplicação do teste de simulação didática no momento 1, por fornecerem informações significativas de possíveis posturas dos professores em sala de aula. Além disso, o primeiro teste foi aplicado em professores que já tinham participado do Proformação por um semestre, ou seja, apresentavam influências do curso realizado.

Com a aplicação do software CHIC, as semelhanças e as diferenças entre as variáveis, ou classes de variáveis, são traduzidas graficamente por uma árvore, denominada árvore de similaridade. Nela podemos perceber seis blocos de variáveis.

O primeiro bloco é referente aos acertos das questões 3 de didática, 2 e 4 de ensino de matemática e 2 de ensino de português. Os acertos dessas questões fornecem uma caracterização dos professores cursistas da AGF de Guandu (Bahia). Observando as situações propostas nessas questões, podemos perceber uma certa postura dos professores:

- discriminam algumas das atividades de ensino de matemática que não são adequadas para nenhum dos dois grupos, como, por exemplo, planejar o procedimento de como passar aos alunos as regras para efetuar a multiplicação e tomar tabuada (questões 2 e 4 de matemática);
- percebem a adequação de uma atividade de ensino de português para os dois grupos de alunos: explorar a capa dos livros (questão 2 de português).

Notam que expondo aos alunos os limites que asseguram a boa convivência social podem estar contribuindo para que os próprios alunos definam as regras de convivência baseadas no respeito pelo outro (questão 3 de didática).

O *segundo bloco* refere-se ao acerto das questões 3, 4, 7, 8 e 9 de ensino de português, 6 e 8 de ensino de matemática e 4 de didática que caracterizaram os professores cursistas de Feira de Santana (Bahia). Esse bloco apresenta os níveis mais fortes de similaridade, além dos mais significativos. Esse grupo ficou caracterizado pelo acerto de algumas das questões não adequadas a nenhum dos ciclos e de duas questões adequadas ao segundo ciclo (3ª e 4ª séries), mostrando um certo discernimento ante situações didático-pedagógicas propostas pelo teste. Além disso, o grupo acertou a única questão de ensino de português considerada a mais adequada para o 1º ciclo (questão 8).

- propõem a produção oral de texto por todos os alunos do primeiro ciclo (questão 8 de português);
- discriminam algumas das atividades que não são adequadas para nenhum dos dois grupos: apenas o professor faz leitura de texto, mostrando que os alunos devem prestar atenção em suas atitudes (questões 3 e 4 de português) e ensinar regras e dar muitos exercícios para que os alunos exercitem o processo, sem dar muita atenção às finalidades de tal técnica (questão 6 de matemática);
- percebem a adequação de algumas das atividades propostas para o segundo ciclo: reflexão sobre o tipo de linguagem empregada – formal ou informal – (questão 7 e português), elaboração de texto escrito pelos próprios alunos (questão 9 de português) e uso de material didático para construção de algoritmos de algumas situações matemáticas (questão 8 de matemática);

- conseguem perceber a necessidade de ensinar a partir da realidade dos alunos, para despertar o interesse deles (questão 4 de didática).

O bom desempenho dos professores cursistas de Feira de Santana (Bahia) talvez possa ser explicado pelo fato de terem experiências de ensino nos dois ciclos.

O *terceiro bloco* complementa as caracterizações das AGFs de Guandu e Feira de Santana. O acerto das questões 3, 9 e 10 de matemática e a questão 1 de didática. Essas questões nos permitem falar das seguintes características:

- discriminam algumas das atividades de matemática para o primeiro ciclo: apresentação de cartazes com figuras simples para falar de multiplicação (questão 3) e, após esse tipo de atividade, a introdução da representação matemática adequada (questão 9);

- percebem a necessidade de apresentar diversas formas de fazer multiplicações para os alunos do segundo ciclo (questão 10);

- propõem atividades relacionadas às experiências culturais que vivenciam para tentar resolver algumas das questões propostas pelos alunos (questão 1 de didática).

As questões 1 de ensino de português, 5 e 7 de ensino de matemática e 2 de didática caracterizam os professores cursistas de Rio Branco, formando o *quarto bloco*. Essas questões estão relacionadas ao ensino de português e matemática e são adequadas aos dois grupos de alunos – primeiro e segundo ciclos.

- percebem a maneira de introduzir a leitura de um livro para que os alunos se interessem por ele (questão 1 de português) e de dividir os alunos em grupos para que resolvam situações do cotidiano (questões 5 e 7 de matemática);

- têm consciência da necessidade do envolvimento de todos – pais, alunos e comunidade – para conseguir recursos de órgãos governamentais que garantam melhores condições materiais da escola (questão 2 de didática).

As AGFs de Quixadá e Massapé não apresentaram níveis fortes de similaridade, formando o *quinto bloco*. A questão 1 de matemática, adequada para os dois ciclos, caracteriza os professores cursistas de Quixadá da seguinte forma:

- iniciam o trabalho com situações da vivência dos alunos que estejam relacionadas à operação matemática que desejam ensinar (questão 1).

As questões que caracterizam os professores da AGF de Massapé foram as 5 e 10 de ensino de português.

- explicam o significado de palavras que aparecem nos textos que são lidos e permitem que os alunos falem a respeito, ou seja, que se comuniquem verbalmente.

Mais duas questões complementam a caracterização dos professores cursistas de Quixadá e Massapé, identificando o *sexto bloco*. A questão 6 de ensino de português é adequada para alunos do segundo ciclo e procura explorar o papel do narrador do texto lido. A questão 5 de didática está relacionada à avaliação de alunos e fala da importância de considerar todo avanço relativo à aprendizagem do aluno, garantindo a apropriação de conhecimentos escolares.

Breves considerações

Ao considerarmos a análise realizada nas informações obtidas com os testes de simulação didática, não podemos esquecer do esforço e da dedicação que os professores cursistas desempenham em suas atividades profissionais diárias. São docentes sem titulação, não tiveram acesso à formação que lhes garantisse qualidade no exercício do magistério, não têm desenvolvido o hábito de estudar de forma autônoma ou, simplesmente, estão afastados do estudo sistemático há muitos anos. Embora todos esses fatos interfiram nas metas propostas pelo Proformação, a análise realizada indica que mudanças foram desencadeadas em concepções e, possivelmente, na prática desses professores.

As relações apresentadas pelo software CHIC nos mostram que os professores parecem valorizar as experiências culturais, participar da vida da comunidade e estimular as lutas por melhores condições de vida, bem como perceber a adequação ou não de uma atividade para os alunos.

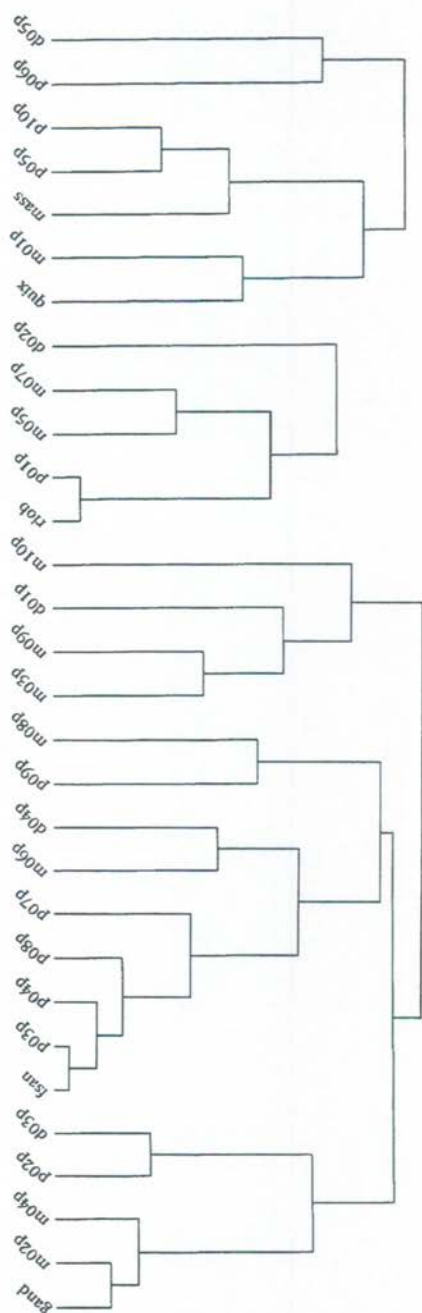
Podemos também considerar que para conhecer os reais benefícios do Proformação, ou de qualquer processo de formação de professores, é necessário buscar informações em diferentes procedimentos e instrumentos de investigação.

A implicação estatística entre variáveis ou classes de variáveis fornece uma ferramenta de estudo de caracteres de natureza variada, que dispõe dados ao pesquisador, cujas ligações podem ser estruturadas em árvores. Essas ligações permitem obter respostas para questionamentos relacionados à concepções ou ao comportamento, como, por exemplo: podemos definir concepções homogêneas e resistentes relativamente a um saber, a partir de classes ordenadas de procedimentos? Com as respostas de um questionário respondido por professores, podemos obter representações de seu ensino? O pesquisador que procura respostas para essas perguntas não pode se privar de outras abordagens metodológicas, pois cada uma permite uma recomposição sintética e, inevitavelmente, reduzida da informação inicial (Gras, 1996).

Referência

GRAS, R. et alii. (1996). *L'implication statistique. Nouvelle méthode exploratoire de données. Applications à la didactique*. Paris, La Pensée Sauvage.

Recebido em mar./2002; aprovado em maio/2002



Árvore de similaridades

Notação:

p01p : questão de ensino de português número 1, primeiro teste

m01p : questão de ensino de matemática número 1, primeiro teste

d01p : questão de didática número 1, primeiro teste

gand: Guandu

fsan: Feira de Santana

riob: Rio Branco

quix: Quixadá

mass: Massapé

O uso do CHIC na análise de registros textuais em ambiente virtual de formação de Professores

MARIA ELISABETTE BRISOLA BRITO PRADO*

Resumo

Este artigo aborda o percurso metodológico de uma pesquisa qualitativa desenvolvida na área de educação, que utiliza o método estatístico multidimensional, por meio do software CHIC – Classificação Hierárquica Implicativa e Coercitiva.

Tal pesquisa teve como cenário um curso de formação de professores voltado para o uso integrado da tecnologia da Informação e Comunicação no contexto da escola, desenvolvido na modalidade a distância. A constituição dos dados desta pesquisa foi se dando a partir da análise interpretativa dos registros textuais dos professores-alunos expressos nos Fóruns de Discussão do ambiente virtual do curso.

Esta análise resultou na identificação de categorias-emergentes, as quais foram submetidas ao referido método estatístico, que permite, por meio do CHIC, a visualização das categorias-emergentes mapeadas em classes e níveis nas árvores de similaridade. A visualização permitiu analisar as relações entre as categorias-emergentes, desvelando novas compreensões sobre o aprendizado dos professores-alunos em função das características do curso de formação desenvolvido no ambiente virtual.

Palavras-chave: formação de professores; educação a distância; percurso metodológico; software CHIC; aprendizagem.

Abstract

This article approaches the methodological course of a qualitative research study carried out in the field of education. We used the multidimensional statistical method through the application of a software called CHIC – Implicative and Coherisitive (Cobesive?) Hierarchical Classification.

We examined a distance teacher-training course on Integrated Use of Information and Communication Technology in the School Context. Data were collected through interpretative analysis of text records written by teacher-students in Discussion Forums that took place in the virtual environment of the course. The analysis resulted in the identification of emerging categories, which were submitted to the referred statistical method. This method allowed us, through application of the CHIC software, to visualize the emerging categories mapped in classes and levels of a similarity tree. That visualization

* Pesquisadora-colaboradora do Núcleo de Informática Aplicada à Educação Nied-Unicamp e doutoranda do Programa de Pós-Graduação em Educação: Currículo, da PUC-SP. E-mail: beprado@terra.com.br.

enabled us to analyze the relationships among the emerging categories and therefore to acquire a new understanding of the teacher-students' learning process, in view of the characteristics of the training course developed in the virtual environment.

Key-words: teacher education; distance education; methodological course; CHIC software; learning.

Introdução

Meu interesse pelo método estatístico multidimensional utilizado em estudos de regras de associação, desenvolvido por meio do CHIC – Classificação Hierárquica Implicativa e Coercitiva, surgiu quando conheci algumas pesquisas na área de Educação e Psicologia que priorizavam a análise qualitativa na busca do entendimento das inter-relações dos dados. Mais particularmente, a pesquisa de Almeida (2000), que tratava da análise de um programa de formação de professores para implantar o computador na prática pedagógica, uma vez que venho atuando e estudando na área de formação de educadores para o uso integrando da informática no contexto de sala de aula.

Assim que comecei a desenvolver a pesquisa de doutorado no Programa de Pós-Graduação em Educação: Currículo da PUC-SP, tendo como cenário o Curso de Especialização em Desenvolvimento de projetos pedagógicos realizados com as novas tecnologias de Informação e Comunicação, o qual foi desenvolvido no mesmo programa da PUC-SP, envolvendo ações de formação na modalidade a distância por meio dos recursos das redes computacionais, vi a possibilidade de utilizar o CHIC.

Para tanto, foi necessário construir um *percurso metodológico*, uma vez que a constituição dos dados dessa pesquisa foi se dando a partir da análise interpretativa dos registros textuais dos professores-alunos expressos nos espaços do ambiente virtual do curso. Assim, este artigo¹ aborda a trajetória metodológica da pesquisa e as contribuições do CHIC, que ampliaram as possibilidades de análise sobre as reflexões expressas durante o processo de aprendizagem dos professores-alunos, vivenciando as ações de formação desenvolvidas no ambiente virtual do curso.

1 Este artigo é baseado na tese de doutorado intitulada "Educação a distância e formação do professor: redimensionando concepções de aprendizagem" da autora, sob a orientação do Prof. Dr. José Armando Valente, defendida em 2003.

Cenário da pesquisa

O Curso de Especialização em Desenvolvimento de projetos pedagógicos realizados com as novas tecnologias de Informação e Comunicação foi desenvolvido no Programa de Pós-Graduação em Educação: Currículo da PUC-SP a partir de uma demanda do Programa Nacional de Informática na Educação (ProInfo), da Secretaria de Educação a Distância (SEED), do Ministério da Educação (MEC), durante o período de agosto de 2000 a maio de 2001, perfazendo uma carga horária² total de 360 horas de curso e mais 60 horas para o desenvolvimento da monografia, finalizando com uma avaliação final presencial.

Nesse curso, participaram 44 (quarenta e quatro) alunos, os quais eram professores da rede pública do ensino de diversas localidades, abrangendo cidades de vários estados do Brasil, inclusive tendo a participação de 1 (uma) aluna da Argentina.

A estrutura do curso foi organizada em três módulos. O módulo I foi realizado presencialmente, com carga horária de 30 horas, sob a responsabilidade dos Núcleos de Tecnologia Educacional, no local de trabalho de cada um dos participantes. Esse módulo tinha o objetivo de oferecer aos participantes noções básicas sobre o manuseio do computador. Os módulos II e III foram realizados totalmente a distância, sob a responsabilidade do Programa de Pós-Graduação em Educação: Currículo da PUC-SP, utilizando o ambiente de suporte para curso a distância TelEduc. Nesse ambiente, os alunos foram divididos em duas turmas, sendo que cada uma era acompanhada sistematicamente por uma coordenadora.

O módulo II, com carga horária de 300 horas, foi organizado em 9 (nove) disciplinas: Realização e Projetos baseados em texto; Realização e Projetos baseados em sistema de autoria programáveis – Logo; Realização e Projetos baseados em Internet; Exploração de diferentes softwares educacionais; Organização, Políticas Educacionais e Currículo; Educação e Tecnologia, Realização e Projetos combinando diversos aplicativos/software;

2 Além dessa carga horária desenvolvida pelo referido Programa, vale esclarecer que houve mais 30 horas, que foram desenvolvidas nos Núcleos de Tecnologia Educacional – NTEs em vários locais do país.

Metodologia para elaboração de monografia e Tópicos em Informática na Educação. Tais disciplinas formam desenvolvidas por docentes e pesquisadores, doutores e mestres da PUC-SP e de instituições credenciadas pelo Ministério de Educação das áreas do conhecimento concernentes.

Esse módulo foi organizado por disciplinas envolvendo a participação de vários docentes, que interagiram com os alunos em diferentes momentos do curso, ou seja, durante a realização de sua disciplina. Isso demandou, da coordenação de turmas, a função de articular as disciplinas, de modo a evitar a fragmentação entre os conteúdos abordados. Além disso, a constante presença das coordenadoras de turmas favoreceu a construção de uma relação em termos de referência humana, a qual permitiu aproximar as pessoas e criar vínculos afetivos, o que, no contexto virtual, assume uma dimensão mais complexa, em relação às questões subjetivas, devido ao meio de expressão e de comunicação ocorrer via escrita (sem os elementos da presencialidade).

Por último, o módulo III, com carga horária de 60 horas, visava à elaboração de uma monografia, a qual deveria expressar uma reflexão-teórico-prática sobre as atividades desenvolvidas ao longo do curso. A orientação das monografias foi feita pelas coordenadoras de turmas³ e pelo coordenador geral⁴, isso porque, nessa estrutura de curso (envolvendo vários e disciplinas), o acompanhamento sistemático de tais profissionais é que favoreceu a orientação dos alunos em suas produções nesse módulo.

A proposta⁵ do curso baseada na abordagem de *formação contextualizada* visava a propiciar ao professor da escola pública que atua ou atuará como multiplicador no Núcleo de Tecnologia Educacional (NTE), o aprendizado dos recursos computacionais e de suas implicações pedagógicas, tanto no âmbito teórico, como no prático. As ações do curso eram voltadas para a vivência do professor-aluno em situações de uso dos recursos

3 Foram coordenadoras de turmas a Profa. Dra. Maria Elizabeth B. de Almeida e a Profa. Ms. Maria Elisabette B. B. Prado.

4 A coordenação geral do curso foi do Prof. Dr. José Armando Valente e a Profa. Dra. Maria Elizabeth B. de Almeida também assumiu a função de coordenadora executiva.

5 A proposta desse curso foi elaborada pelo Prof. Dr. José Armando Valente (docente da Unicamp e da PUC-SP) e pela Profa. Dra. Maria Elizabeth B. de Almeida (docente da PUC-SP).

computacionais, ora como aprendiz, ora como professor, atuando com os alunos no cotidiano de sua prática pedagógica escolar. O contexto da prática foi o foco do curso, propiciando ao professor-aluno refletir sobre as potencialidades e as limitações dos recursos computacionais, para integrá-los nas atividades de sala de aula.

Esse curso, portanto, contemplou o contexto, a realidade de atuação do professor, tendo como base o princípio da reflexão de Schön (1983, 1992), acrescidas dos novos elementos do universo virtual, os quais foram tratados numa perspectiva de educação a distância, denominada por Valente (2000) o *estar junto virtual*, que enfatiza a interação e o processo de reconstrução do conhecimento. Esse cenário do curso me instigou a analisar as reflexões dos professores-alunos expressas nos espaços do ambiente virtual, em relação aos 4 (quatro) eixos temáticos fundamentados teoricamente: (1) Uso pedagógico dos recursos computacionais, (2) Processo de aprendizagem, (3) Prática Pedagógica e (4) Conceituação da aprendizagem e da prática pedagógica.

Construindo um percurso metodológico

Como a viabilização do curso ocorreu por meio de um ambiente (plataforma) computacional, neste caso o TelEduc, todos os documentos que foram sendo produzidos ao longo do curso ficaram armazenados nesse ambiente, gerando um grande volume de dados de diversas características. Esses documentos correspondem aos materiais utilizados e/ou produzidos pelos docentes (textos de leituras, informativos, dinâmicas de encaminhamento de atividades, entre outros), às produções dos alunos (elaboração de atividades: programas computacionais e textos) e às interações que acontecem nos diversos espaços de comunicação do ambiente virtual (correio eletrônico, fórum de discussão, chat e portfólio).

Dentre esses documentos, optei pelos registros textuais do fórum de discussão, pelo fato de ter ocorrido sistematicamente em todas as disciplinas do curso, mas, essencialmente, por dois fatores. Primeiro, porque queria manter o mesmo tipo de registro textual, considerando que,

em cada espaço virtual de comunicação, o uso da linguagem tem uma característica própria. Segundo, porque o fórum de discussão, sendo um espaço de interação assíncrona, permitia aos professores-alunos explicitar suas ações-pensamentos de forma mais elaborada, revelando com isso suas reflexões sobre determinados momentos de aprendizagem que estavam vivenciando no curso.

A partir dessa escolha, comecei a “garimpar” o dado bruto, procurando identificar um fórum representativo de cada uma das disciplinas. Por sua vez, cada disciplina, dependendo da dinâmica, criava mais de um fórum de discussão. Nesse caso, utilizei o mesmo critério de escolha, ou seja, a sistematicidade dos assuntos, por exemplo, em uma disciplina que envolvia o aprendizado de *softwares/aplicativos*, era comum ter um fórum para discutir questões operacionais e outro, voltado para um tema que reportava tanto à aprendizagem como à prática pedagógica relacionada ao uso da tecnologia. Focalizei a seleção dos fóruns por assuntos que retratavam as reflexões dos professores-alunos voltadas para as implicações dos recursos computacionais no processo de ensino e aprendizagem, uma vez que este era o foco da pesquisa.

Assim, após essa seleção dos fóruns, fui em busca do dado achado, ou seja, daquele dado que foi organizado no contexto de interação, mais especificamente, quando o professor-aluno explicitava suas reflexões acerca de uma temática pertinente aos propósitos das disciplinas. Esse momento de nova garimpagem exigiu muitas (re)leituras dos registros textuais dos professores-alunos, para a compreensão da escrita-oralizada de suas reflexões. Nessas (re)leituras procurei selecionar apenas os registros textuais que expressavam as reflexões voltadas para algumas questões relacionadas aos 4 (quatro) eixos temáticos: (1) Uso pedagógico dos recursos computacionais, (2) Processo de aprendizagem, (3) Prática Pedagógica e (4) Conceituação da aprendizagem e da prática pedagógica.

Desta forma, com o universo dos dados mais delineado, o próximo procedimento envolveu a análise-interpretativa de 185 (cento e oitenta e cinco) registros textuais. Nesse momento, com a matéria-prima constituída, a finalidade da análise-interpretativa era de identificar as “categorias-emergentes” relacionadas aos 4 (quatro) eixos temáticos em cada um dos

registros textuais. Esse termo “categoria-emergente”, utilizado na pesquisa, refere-se a um conjunto de idéias expressas nos registros textuais analisados, portanto, não se trata de uma categoria conceitual teórica. Dito de outra maneira, as categorias-emergentes não foram definidas *a priori* e sim identificadas a partir da análise interpretativa dos registros textuais dos professores-alunos expressos no fórum de discussão do ambiente TelEduc.

A interpretação desses registros textuais foi feita com base nos pressupostos teóricos norteadores da minha trajetória profissional e acadêmica na área de formação de professores para o uso da informática na escola. O cerne desse construto teórico se constituiu dos princípios *construcionistas* de Papert (1983, 1990, 1994) e Valente (1993, 1999, 2002) que enfatizam a natureza da aprendizagem e o papel da tecnologia aliado ao *conceito da reflexão* de Schön (1983, 1992) na formação do profissional prático reflexivo. Além disso, essa interpretação estava atrelada ao conhecimento vivencial do contexto, devido à minha constante participação ao longo do curso. De fato, a vivência como coordenadora de turma e docente foi fundamental, pois o fato de conhecer o contexto do curso sob diferentes perspectivas, bem como o seu movimento interno, permitiu-me interpretar as explicitações dos professores-alunos de forma contextualizada em relação às ações de formação, considerando as especificidades das disciplinas.

A partir do primeiro resultado da análise-interpretativa, submeti os dados – os 185 (cento e oitenta e cinco) registros textuais dos professores-alunos, com as 27 (vinte e sete) categorias-emergentes identificadas com seus respectivos indicadores e níveis de reflexão – para uma outra pesquisadora fazer uma análise com vistas a chegarmos a um consenso interpretativo. Esse procedimento possibilita imprimir um caráter de confiabilidade na identificação das categorias-emergentes, bem como na classificação dos três níveis de reflexão, conforme os critérios descritos a seguir:

- NÍVEL 1 reflexão empírica sobre fatos/situações observáveis do cotidiano do professor-aluno.
- NÍVEL 2 reflexão com estabelecimento de relações entre os aspectos envolvidos no contexto do curso e/ou da prática profissional.
- NÍVEL 3 reflexão que evidencia sistematização do conhecimento, potencializando mudanças de concepções e/ou práticas.

Assim, as categorias-emergentes, em seus respectivos níveis relacionadas aos 4 (quatro) eixos temáticos, constituíram um novo cenário de análise, tal como segue na sua organização abaixo:

Tema 01 – Uso pedagógico dos Recursos Computacionais (RC)

Código	Categorias-emergentes
01RC	Incentivo para aprender – NÍVEL 1
02RC	Concepção integradora dos recursos computacionais – NÍVEL 2
03RC	Relação entre os recursos computacionais e os aspectos da aprendizagem – NÍVEL 2
04RC	Indício de mudança na concepção de uso dos recursos computacionais – NÍVEL 3

Tema 02 – PROCESSO DE APRENDIZAGEM (PA)

Código	Categorias-emergentes
05PA	Aspectos Significativos – NÍVEL 1
06PA	Aspectos Cognitivos – NÍVEL 1
07PA	Aspectos Emocionais – NÍVEL 1
08PA	Interação Compartilhamento – NÍVEL 1
09PA	Interação Comprometida – NÍVEL 2
10PA	Contextualização – NÍVEL 2
11PA	Reflexão sobre a própria aprendizagem – NÍVEL 2
12PA	Reflexão sobre aprendizagem do outro – NÍVEL 2
13PA	Reflexão articulada entre a própria aprendizagem e do outro – NÍVEL 3
14PA	Espiral de Aprendizagem – NÍVEL 3

Tema 03 – PRÁTICA PEDAGÓGICA (PP)

Código	Categorias-emergentes
15PA	Análise da concepção tradicional de educação – NÍVEL 1
16PA	Postura do professor diante da tecnologia – NÍVEL 1
17PA	Intencionalidade Pedagógica – NÍVEL 1
18PA	Flexibilidade e abertura – NÍVEL 2
19PA	Reflexão sobre a prática pedagógica – NÍVEL 2
20PA	Estratégias Pedagógicas – NÍVEL 2
21PA	Articulação entre a prática pedagógica e aprendizagem – NÍVEL 2
22PA	Reflexão sobre mudança na escola – NÍVEL 2
23PA	Articulação teoria e prática pedagógica – NÍVEL 3
24PA	Indício de mudança na prática pedagógica – NÍVEL 3

Tema 04 – (CA) CONCEITUAÇÃO DE APRENDIZAGEM E DA PRÁTICA PEDAGÓGICA (CA)

Código	Categorias-emergentes
25CA	Conhecimento como construção – NÍVEL 3
26CA	Recontextualização do aprendizado – NÍVEL 3
27CA	Ressignificação de teoria e/ou prática – NÍVEL 3

A partir dessa constituição dos dados, busquei o apoio de um método estatístico multidimensional utilizado em estudos qualitativos de regras de associação. Esse método é viabilizado por um *software* que faz a análise hierárquica de similaridade, permitindo visualização de semelhanças e classes de variáveis mapeadas em níveis de uma árvore hierárquica. Essa metodologia consiste em uma organização e análise de dados segundo seu agrupamento e interseção que se desenvolve por meio do *software* CHIC – Classificação Hierárquica, Implicativa e Coercitiva (Almouloud, 1997).

Dessa maneira, os dados resultantes da análise interpretativa dos registros textuais (categorias-emergentes), representativos de cada disciplina do curso, foram compilados em uma planilha do Excel. Foram elaboradas três planilhas constituídas pelo agrupamento das disciplinas que tinham características comuns não de conteúdo, mas sim de situações de aprendizagem. Dessa forma, a tabela referente ao agrupamento 1 foi formada pelos registros textuais das disciplinas que envolveram atividades

práticas, usando os *softwares*/aplicativos e atuando com alunos ou professores no contexto da escola. O agrupamento 2 constitui-se dos registros textuais das disciplinas em que os professores-alunos, afastados das atividades práticas, participavam de discussões, refletindo sobre as questões relacionadas ao universo da prática pedagógica. E, o agrupamento 3 foi constituído pelos registros textuais das disciplinas em que os professores-alunos, afastados das atividades práticas, participavam de discussões, refletindo sobre as teorias relacionadas ao processo de ensino e aprendizagem.

Assim, foi feita uma tabela binária para cada agrupamento e para isso codifiquei as categorias-emergentes como variáveis (por exemplo: 01RC, 06PA, 22PP, etc.), as quais foram colocadas na primeira linha da tabela e os registros textuais codificados (por exemplo: R101, R102, etc.), foram colocados na primeira coluna. Os valores (0 ou 1) foram lançados na tabela de dupla entrada, indicando a presença ou ausência da categoria-emergente correspondente ao registro textual. O exemplo abaixo mostra uma tabela parcial.

	01RC	02RC	03RC	04RC	05PA	06PA	07PA
R101	1	0	0	1	1	0	1
R102	1	1	0	0	1	1	0
R103	1	0	1	0	0	1	1
R104	1	1	1	1	0	0	0
R105	1	0	0	0	0	0	0
R106	0	0	0	1	0	0	1

Assim, as tabelas feitas em Excel (com extensão do arquivo em csv) correspondentes a cada um dos agrupamentos foram carregadas pelo software CHIC, gerando três gráficos da árvore de similaridade. O agrupamento 1 gerou uma árvore de similaridade, permitindo fazer a análise relacional das categorias-emergentes na situação APRENDENDO NA AÇÃO.

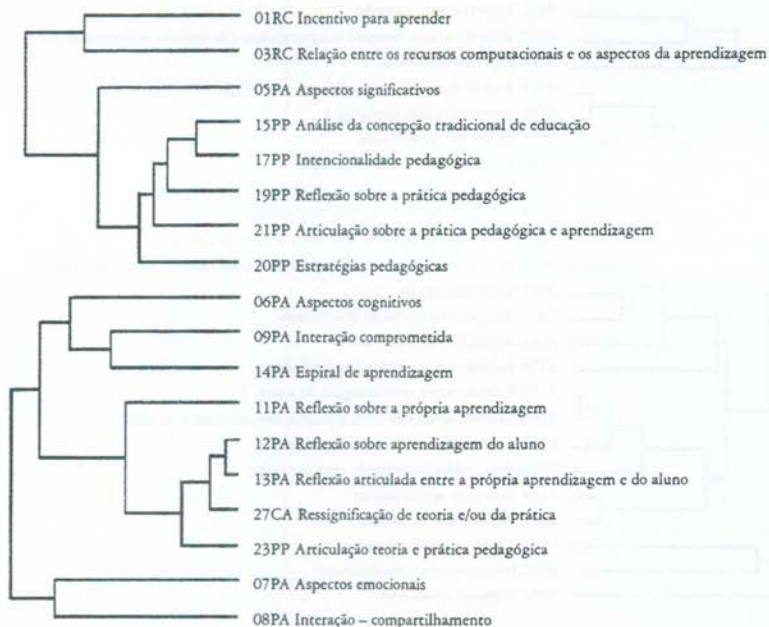


Figura 1 – Gráfico da Árvore de similaridade da situação APRENDENDO NA AÇÃO (Agrupamento 1)

Em síntese, a visualização dessa árvore mostra que na situação APRENDENDO NA AÇÃO aparecem 18 categorias-emergentes de níveis 1, 2 e 3 de reflexão, embora o nível 3 tenha se destacado nas categorias-emergentes relacionadas ao eixo temático PROCESSO DE APRENDIZAGEM.

No entanto, para poder analisar o movimento e as transformações das categorias-emergentes entre os agrupamentos, foi gerada uma árvore de similaridade integrando os agrupamentos 1 e 2, permitindo, com isso fazer a análise relacional na situação TEORIZANDO NA AÇÃO.

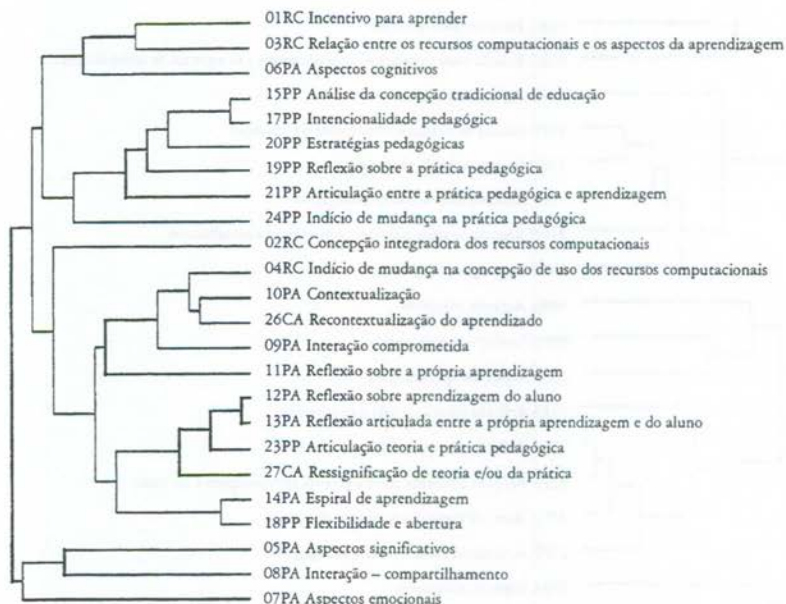


Figura 2 – Gráfico da Árvore de similaridade da situação TEORIZANDO NA AÇÃO (Agrupamentos 1 e 2)

Na visualização dessa árvore de similaridade, observam-se as 24 categorias-emergentes, com predominância dos níveis 2 e 3 de reflexão. Nessa situação, TEORIZANDO NA AÇÃO, aparecem seis novas categorias-emergentes: CONCEPÇÃO INTEGRADORA DOS RECURSOS COMPUTACIONAIS [02RC, NÍVEL 2], INDÍCIO DE MUDANÇA NA CONCEPÇÃO DE USO DOS RECURSOS COMPUTACIONAIS [04RC, NÍVEL 3], CONTEXTUALIZAÇÃO [10PA, NÍVEL 2], FLEXIBILIDADE E ABERTURA [18PP, NÍVEL 2], INDÍCIO DE MUDANÇA NA PRÁTICA PEDAGÓGICA [24PP, NÍVEL 3] e RECONTEXTUALIZAÇÃO DO APRENDIZADO [26CA, NÍVEL 3], representativas dos quatro eixos temáticos.

E, por último, com a mesma intenção de analisar o movimento e as transformações das categorias-emergentes entre os agrupamentos, foi gerada uma árvore de similaridade integrando os agrupamentos 1, 2 e 3, permitindo a realização da análise relacional na situação INTEGRANDO A TEORIA NA AÇÃO.

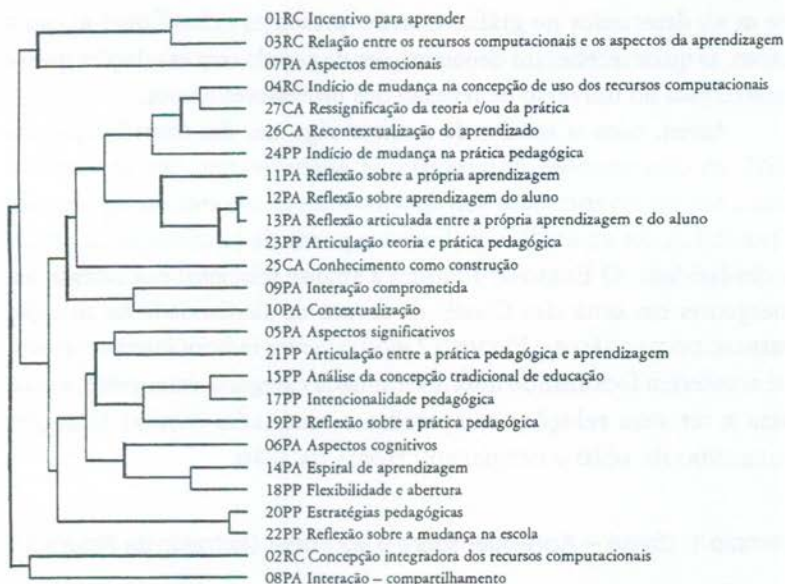


Figura 3 – Gráfico da Árvore de similaridade da situação INTEGRANDO TEORIA NA AÇÃO (Agrupamentos 1, 2 e 3)

Nessa árvore de similaridade, observam-se as 26 categorias-emergentes com predominância dos níveis 2 e 3 de reflexão. Nessa situação, INTEGRANDO TEORIA NA AÇÃO, aparecem duas novas categorias-emergentes: REFLEXÃO SOBRE A MUDANÇA NA ESCOLA [22PP, NÍVEL 2] e CONHECIMENTO COMO CONTRUÇÃO [25CA, NÍVEL 3], representativas dos temas PRÁTICA PEDAGÓGICA e CONCEITUAÇÃO DA PRÁTICA PEDAGÓGICA E A APRENDIZAGEM.

Essa organização dos dados propiciou analisar as relações entre as categorias-emergentes *na* própria situação de aprendizagem e *entre* as situações e aprendizagem (aprendendo na ação, teorizando na ação e integrando a teoria na ação), revelando aspectos importantes do processo de aprendizagem dos professores-alunos vivenciados no curso.

Análise Interpretativa baseada no CHIC

Para analisar as convergências entre as categorias-emergentes, extraí, das árvores de similaridade fornecida pelo CHIC, as associações em

que os *nós* detectados no gráfico estavam presentes e classifiquei-as como CLASSES, as quais receberam denominações de acordo com as relações que se sintetizavam no universo de reflexão dos professores-alunos.

Assim, com o intuito de mostrar algumas das contribuições do CHIC na compreensão do processo de aprendizagem dos professores-alunos, este tópico apresenta dois tipos distintos de recortes, exemplificando as possibilidades da análise qualitativa a partir da visualização da árvore de similaridade. O EXEMPLO 1 mostra a análise relacional das categorias-emergentes em uma das Classes da árvore de similaridade *na* situação APRENDENDO NA AÇÃO e o EXEMPLO 2 mostra novos indícios interpretativos, que acontecem focalizando uma determinada categoria-emergente, a qual passa a ter suas relações comparadas e analisadas *entre* as situações TEORIZANDO NA AÇÃO e INTEGRANDO TEORIA NA AÇÃO.

Exemplo 1: Classe – Aprender sobre o aprender (extraído da Figura 1)

Essa CLASSE apresenta-se constituída pelas categorias-emergentes REFLEXÃO SOBRE A PRÓPRIA APRENDIZAGEM [11PA, NÍVEL 2], REFLEXÃO SOBRE APRENDIZAGEM DO OUTRO [12PA, NÍVEL 2], REFLEXÃO ARTICULADA ENTRE A PRÓPRIA APRENDIZAGEM E A DO OUTRO [13PA, NÍVEL 3], RESSIGNIFICAÇÃO DE TEORIA E/OU DA PRÁTICA [27CA, NÍVEL 3] e ARTICULAÇÃO TEORIA E PRÁTICA PEDAGÓGICA [23PP, NÍVEL 3], como mostra a figura abaixo:

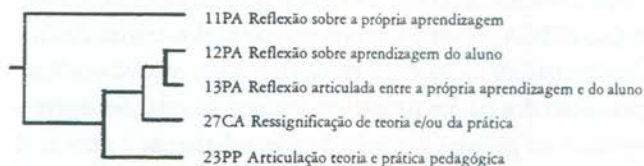


Figura 4 – CLASSE – APRENDER SOBRE APRENDER

O *nó* de similaridade mais forte dessa CLASSE, formado pelas categorias-emergentes, REFLEXÃO SOBRE APRENDIZAGEM DO OUTRO [12PA, NÍVEL 2] e REFLEXÃO ARTICULADA ENTRE A PRÓPRIA APRENDIZAGEM E A DO OUTRO [13PA, NÍVEL 3] mostra que existe uma grande probabilidade de que as atividades práticas desenvolvidas nas disciplinas do curso – particular-

mente aquelas que desafiam o professor-aluno a colocar em ação o pensamento lógico, criativo e estético, tal como aconteceu nas disciplinas Realização de projetos baseados em sistemas de autoria programáveis – Logo e Realização de projetos baseado na Internet – favoreçam a REFLEXÃO SOBRE A APRENDIZAGEM. O extrato do registro textual abaixo exemplifica este tipo de reflexão:

(...) quanto pode ser interessante e expressivo observar como os alunos constroem os links numa atividade no Powerpoint. Nesse tipo de atividade enuncia-se muito da forma de pensar do sujeito. Isso me ocorreu há pouco vendo as atividades dos colegas. Enquanto... [os colegas] construíam os organogramas claros, objetivos, enunciando o pensamento de quem construiu as apresentações, eu ia me vendo, me observando, num espelho mágico e inusitado. Comparando as apresentações destes colegas com as minhas me enxerguei, naquilo que muitas vezes se constitui até um problema de relacionamento para mim... Compreender a forma de pensar, os caminhos percorridos pelo raciocínio pelo sujeito, que é um caminho que denuncia sua história de vida e seus interlocutores, é um passo importante para o professor encaminhar um processo de aprendizagem que seja interessante, bonito, firme, com significado para os participantes daquele fazer. [Registro Textual de M.C.B.O.M., 11/2000].

Nessa situação, o professor-aluno envolvido no processo de observar e comparar a trajetória cognitiva do colega que se expressa na atividade de produção de um hipertexto (como parte da construção de página), reflete sobre a aprendizagem do outro e a utiliza como espelho para analisar e conhecer seu próprio processo de aprender. Essa articulação entre a aprendizagem do outro e a própria aprendizagem pode ser vista como um meio que favorece ao professor-aluno atribuir novos significados e, portanto, compreender o processo de aprendizagem.

Esse *nó*, por sua vez, propiciou a formação encadeada de outros *nós*, representados pelo NÍVEL 3 DE REFLEXÃO, particularmente pelas categorias-emergentes RESSIGNIFICAÇÃO DE TEORIA E/OU DA PRÁTICA [27CA, NÍVEL 3] e ARTICULAÇÃO TEORIA E PRÁTICA PEDAGÓGICA [23PP, NÍVEL 3]. Essas relações são indicativas de que, possivelmente, houve para alguns professores-alunos uma compreensão mais sistematizada do conhecimento sobre o processo de aprendizagem. O extrato a seguir exemplifica essas relações:

Tentarei fazer um paralelo entre a experiência e o texto analisado.... A [aluna] foi muito espontânea ao organizar seus pensamentos para construir o que estava querendo. No momento em que descobriu que poderia comandar a Tat, sentiu-se segura, pois, iria poder alcançar os seus objetivos... Esta experiência com a [aluna] se enquadra no processo construcionista... Achei tão interessante este momento: eu já tinha analisado o texto, mas, relendo novamente na tentativa de enquadrar com a minha experiência, a leitura foi diferente, em certos trechos, percebi que eu tinha acabado de vivenciar com a [aluna]. Parece-me que as autoras escrevem o texto baseado no meu relato... agora, é tudo tão real: o Logo, a produção pelo aluno... [Registro Textual de T.K.A.N., 10/2000]

Nesse texto, a reflexão do professor-aluno mostra a relação entre aquilo que observa e analisa na prática atuando com a aluna e a leitura feita sobre os princípios teóricos que fundamentam o uso da Linguagem de programação Logo no processo de ensino e aprendizagem. Nessa situação, a atividade prática propicia ao professor-aluno fazer novas interpretações teóricas, atribuindo novos significados à teoria e à prática.

Durante a realização das atividades práticas, os professores-alunos tiveram acesso aos textos específicos sobre os fundamentos educacionais que norteiam o uso computador no contexto da escola. As leituras desses textos eram indicadas para subsidiar o desenvolvimento do projeto/atividade do professor-aluno na sua atuação com os alunos, usando os *softwares*/ aplicativos e, intencionalmente, para propiciar a articulação entre a teoria e a prática. Isso nos mostra que uma estratégia de formação que favorece a articulação da teoria e da prática precisa propiciar ao professor-aluno vivenciar uma situação concreta de modo que possa observar, comparar e refletir sobre a aprendizagem do aluno. Em outras palavras, uma situação concreta em que o professor-aluno possa colocar em ação e constatar na ação a teoria abordada nas leituras de textos e nos debates com os colegas e docentes.

É interessante observar que foi a partir das atividades práticas e das inquietações que emergiam das interações do professor-aluno com o computador e com os alunos que os pressupostos educacionais ganharam sentido, elucidando sua compreensão sobre uma situação concreta de aprendizagem na ação. Esse encadeamento de reflexões e de articulações, por sua vez, aparece ancorado na categoria-emergente REFLEXÃO SOBRE A

PRÓPRIA APRENDIZAGEM [11PA, NÍVEL 2], completando a CLASSE principal da Árvore de Similaridade (figura 1)

Exemplo 2: Foco na Categoria-emergente

INDÍCIO DE MUDANÇA DA CONCEPÇÃO DE INFORMÁTICA NA EDUCAÇÃO [04RC, NÍVEL 3], é a categoria-emergente focalizada que se apresenta na árvore de similaridade da situação TEORIZANDO NA AÇÃO na Classe denominada CONTEXTUALIZAÇÃO E RECONTEXTUALIZAÇÃO, como mostra a figura abaixo:

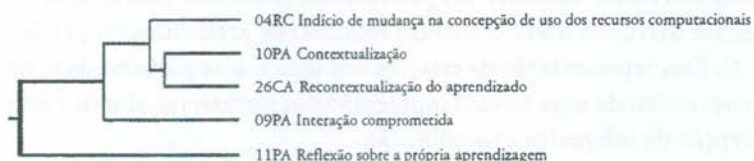


Figura 5 – CLASSE – CONTEXTUALIZAÇÃO E RECONTEXTUALIZAÇÃO

Esse conjunto de relações nos dá uma pista interpretativa forte, de que a contextualização mediada pelas interações virtuais pode potencializar o professor-aluno para recontextualizar aquilo que está aprendendo na sua realidade prática de atuação, tal como mostra o extrato abaixo:

As potencialidades que favorecem a interação deste curso tem vários aspectos, um deles tem que ver com nosso dia a dia, com a atividade áulica, a relação que nós mantemos com os alunos e a possibilidade de analisar "in situ" os resultados dos projetos implementados, refletindo com o grupo e compartilhando experiências, sendo no mesmo momento professores-alunos, e vivenciando nossas próprias aprendizagens, recontextualizando a prática. Outra questão tem que ver com a colaboração intergrupo, é sentir-se acompanhada por outras personas, quasi mais que em num curso presencial, que mesmo na distancia ficam perto da gente.... [Registro textual de A.A., 12/2000].

A reflexão do professor-aluno revela a importância de aprender no contexto real do seu trabalho e, que essa aprendizagem é favorecida pela qualidade da interação que se estabelece no ambiente virtual do curso.

pedagógico dos recursos computacionais. A ressignificação da prática demonstra uma compreensão da informática na educação numa perspectiva integradora da tecnologia nas atividades pedagógicas, favorecendo ao professor-aluno recontextualizar o seu aprendizado. No entanto, essa recontextualização se dá num nível mais elaborado do que acontece na situação TEORIZANDO NA AÇÃO, pois nesse momento são anunciadas novas práticas pedagógicas, tal como mostra o extrato abaixo:

...Hoje, sinto-me mais segura para atuar como coordenadora do laboratório, principalmente como professora. Lembro-me do primeiro projeto que tivemos de desenvolver neste curso juntamente com os alunos... devido a intervenção da formadora/monitora, fez-me analisar o quanto eu, alunos e professora da disciplina estávamos "equivocados" no momento de destacar as dúvidas provisórias (dos alunos), isto é, eles iriam, na verdade, usar os computadores, principalmente para realizar uma atividade/pesquisa de Geografia através dos computadores e não vice-versa. Essa experiência inicial foi de suma importância para minha prática pedagógica, inclusive, estou conseguindo orientar melhor meus colegas de trabalho quanto à questão de elaboração e execução de projetos [Registro textual de T.K., 12/2000].

O início de mudança da prática pedagógica aparece atrelado ao conjunto de relações dessa Classe e isso acena uma probabilidade de que os professores-alunos estejam projetando ou Tateando modificações na prática pedagógica. O fato de as reflexões dos professores-alunos evidenciarem uma nova compreensão da informática na educação e novos encaminhamentos nas situações reais da sua atuação profissional (enquanto multiplicadores) nos dá pistas mais assertivas para dizer que alguns professores-alunos estão (re)construindo o conhecimento prático e que mudanças parciais e contextuais podem estar ocorrendo.

Algumas considerações

Nesta trajetória metodológica em busca do entendimento das relações entre as categorias-emergentes e as possibilidades de aprendizagem do professor-aluno na abordagem do curso de formação desenvolvido no

ambiente virtual, a utilização do *software* CHIC foi fundamental. Isto porque o meu envolvimento com os dados, ao mesmo tempo que revelava aspectos interessantes acerca das reflexões dos professores-alunos sobre as questões relacionadas à aprendizagem e à prática pedagógica, instigava-me a olhar para as categorias-emergentes a partir de uma nova perspectiva de análise. Era preciso me distanciar dos dados e da realidade do curso para interpretar novas possibilidades de aprendizagem que a educação a distância pode oferecer para a formação contextualizada dos professores.

E isso só foi possível com o uso do *software* CHIC, que propiciou identificar relações inesperadas, as quais desvelaram novas compreensões sobre o processo de aprendizagem dos professores-alunos. Essa compreensão, por sua vez, é que aponta novos horizontes para a formação do professor reflexivo, capaz de reconstruir a sua prática pedagógica integrando os recursos computacionais. E, talvez o mais interessante, é que essa formação possa acontecer integrando os novos elementos do contexto virtual.

Referências

- ALMEIDA, M. E. B. de (2000). *O computador na escola: contextualização e a formação de professores*. Tese de doutorado. Programa de Pós-Graduação em Educação: Currículo. São Paulo, Pontifícia Universidade Católica.
- ALMOULOU, S. A. (1997). *L'Ordinateur, outil d'aide à l'apprentissage de la démonstration et de traitement de données didactiques*. These de Docteur, U. F. R. de Mathematiques. Rennes, França, Université de Rennes I.
- PAPERT, S. (1985). *Logo: computadores e educação*. São Paulo, Brasiliense.
- _____. *Computer Criticism vs. Technocentrism*. Massachusetts, CA. *E&L MENO*, n. 1.
- _____. (1994). *A máquina das crianças: repensando a escola na era da informática*. Porto Alegre, Artes Médicas.
- SCHÖN, D. A. (1983). *The Reflective Practitioner – How Professionals Think in Action*. New York, Basic Books.

- SCHÖN, D. A. (1992). Formar professores como profissionais reflexivos. In: NÓVOA, A. (coord.). *Os professores e a sua formação*. Lisboa, Publicações Dom Quixote/ Instituto de Inovação Educacional.
- VALENTE, J. A. (1993). "Por que o computador na educação". In: VALENTE, J. A. (org.). *Computadores e conhecimento: repensando a educação*. Campinas, SP, Gráfica da Universidade Estadual de Campinas.
- _____. (1999). "Formação de professores: diferentes abordagens pedagógicas". In: VALENTE, J. A. (org.). *O computador na sociedade do conhecimento*. Campinas, SP, Unicamp-Nied.
- _____. (2000). "Educação a distância: uma oportunidade para mudança no ensino". In: MAIA, C. (coord.). *ead.br: educação a distância no Brasil na era da Internet*. São Paulo, Anhembi Morumbi.
- _____. (2002). "A espiral da aprendizagem e as tecnologias da informação e comunicação: repensando conceitos". In: JOLY, M. C. (ed.). *Tecnologia no ensino: implicações para a aprendizagem*. São Paulo, Casa do Psicólogo.

Recebido em abr./2002; aprovado em jun./2002

Formação de professores para inserção do computador na escola: Inter-relações entre percepções evidenciadas pelo uso do *software* CHIC

MARIA ELIZABETH BIANCONCINI DE ALMEIDA*

Resumo

O presente trabalho apresenta a metodologia de análise de dados qualitativos desenvolvida com o uso do *software* CHIC para construção de árvores hierárquicas entre temas em análise evidenciando as inter-relações entre as percepções dos sujeitos participantes de uma ação de formação de professores para a inserção do computador na prática pedagógica, contextualizada na realidade da escola pública e na prática pedagógica do professor. Inicialmente, descreve o contexto em que a formação de professores se realizou e as respectivas concepções teóricas. Em seguida, discute as dificuldades enfrentadas para analisar qualitativamente a imensa massa de dados disponível levando em conta convergências, divergências, ambigüidades e idiosincrasias encontradas nos depoimentos dos sujeitos da formação. Finalmente, apresenta as árvores construídas pelo CHIC e os respectivos resultados de análises, evidenciando que as práticas pedagógicas de uso do computador ocorrem freqüentemente fora do âmbito da sala de aula, não tendo se incorporado às práticas cotidianas da classe. Uma conquista revelada é a consciência de desenvolver um ensino comprometido com a aprendizagem significativa do aluno e a importância de que os programas de formação tenham como eixo o contexto de atuação do professor e a realidade da escola, direcionando-se para a criação de uma cultura que permita ao educador tornar-se usuário crítico da tecnologia, utilizá-la em sua prática pedagógica e assumir-se como um agente de mudança de sua própria atuação e contexto.

Palavras-chave: análise de dados multidimensional; *software* CHIC; formação de professores contextualizada; construção do conhecimento; tecnologia na educação.

Abstract

The present work shows the methodology for qualitative data analysis developed with the use of the CHIC software for the construction of hierarchical trees among themes under analysis. The analysis revealed the interrelations between subjects' perceptions in a teachers' development action work

* Professora do Departamento de Ciência da Computação e do Programa de Pós-Graduação em Educação: Currículo, da PUC-SP. E-mail: bbalmeida@uol.com.br.

focusing on computer introduction in pedagogical practice, contextualized both in the reality of the public school and in the teacher's pedagogical practice. First, the article describes the context in which the teachers' development was carried out and its theoretical conceptions. Next, it discusses the difficulties faced during the qualitative analysis of the immense amount of available data, taking into account convergences, divergences, ambiguities and idiosyncrasies found in the subjects' accounts. Finally, it presents the trees built by CHIC and their respective analysis results, showing that pedagogical practices using the computer frequently occur outside the classroom environment, hence have not been incorporated into the daily classroom practices. One of the positive outcomes was the awareness regarding the development of a teaching practice that is committed to the students' meaningful learning and regarding the importance of having development programs that focus on the teacher's practice context and on the school's reality. Thus, these programs can be directed towards the creation of a culture that enables educators to become critical users of technology, able to utilize it in their practice, aware of their role as agents of change of their own action and context.

Key-words: *multidimensional data analysis; CHIC software; contextualized teachers' development; knowledge building; technology in education.*

Introdução

Comecei a me dedicar a pesquisas e formação de educadores para o uso do computador na prática pedagógica no período em que trabalhei como docente na Universidade Federal de Alagoas – UFAL.¹ Em 1991, coordenei o Seminário Nacional de Informática na Educação, que teve a participação de conferencistas nacionais e internacionais de renome no cenário da área, entre os quais o Prof. Dr. Régis Gras, da Universidade de Nantes, que apresentou uma palestra sobre análise de dados qualitativos com o uso da estatística multidimensional.

As possibilidades vislumbradas, de construir árvores hierárquicas que evidenciavam as inter-relações entre os dados através da metodologia e do *software* CHIC, desenvolvidos pelo Prof. Regis, permitiam antever contribuições efetivas à análise de informações oriundas de programas de formação de professores desenvolvidos na época. No entanto, outros acontecimentos centravam minha atenção nas ações de formação e as análises voltavam-se à ação-reflexão-depuração²-nova ação, aplicando diferentes metodologias de análise, sem utilizar o CHIC.

1 De 1982 a 1995, fui professora do Departamento de Matemática Aplicada da UFAL.

2 Depuração é um conceito usado na atividade de programação que se refere à análise das ações representadas para encontrar os erros e corrigi-los. Dessa forma, o erro deixa de ser objeto de punição e passa a ser objeto de investigação, revisão e construção.

Algum tempo depois, já na Pontifícia Universidade Católica de São Paulo, onde fazia o doutorado, continuei a investigar a formação de professores para a inserção do computador na prática pedagógica e adotei como estudo de caso um programa de formação de grande porte, desenvolvido pela PUC-SP em parceria com a Secretaria de Educação do Estado de São Paulo, SEE.

Este artigo, inspirado em uma tese de doutorado defendida na PUC-SP (Almeida, 2000), descreve a metodologia de uso do *software* CHIC desenvolvida na análise da formação de professores para a inserção do computador na prática pedagógica, contextualizada na realidade da escola pública e na prática pedagógica do professor.

Contexto analisado

O uso crescente da tecnologia de informação e comunicação, TIC, em escolas brasileiras é um fato cada vez mais comum, tanto em escolas privadas como no sistema público. Neste último caso, vários projetos foram executados com importantes contribuições, destacando-se o caráter inovador da inserção do computador na escola com ênfase na mudança educacional e na aprendizagem do aluno (Almeida, 1999; Valente e Almeida, 1997; Valente, 1999). Embora a almejada transformação do sistema educacional não tenha se concretizado nesses projetos, eles lançaram as bases para a formação de uma massa crítica de pesquisadores que influenciou iniciativas posteriores.

Hoje, programas governamentais destinados à introdução da TIC na educação começam a tornar-se realidade, não mais como experimento-piloto, e atingem um percentual considerável de escolas. Programas de âmbito federal, estadual ou municipal encontram-se em execução, sendo marcados pela ênfase na formação do professor para utilizar a tecnologia de informação e comunicação integrada ao processo de ensino e de aprendizagem.

Entre os programas de inserção do computador na prática pedagógica em escolas públicas, o subprojeto Informática na Educação do Programa de Educação Continuada – Inovações no Ensino Básico

(PEC-IEB),³ desenvolvido pela Secretaria de Estado da Educação de São Paulo, de setembro de 1997 a novembro de 1998, no qual a PUC-SP realizou as ações de formação, caracteriza-se como uma atividade educativa entre educadores no exercício de suas funções em noventa e quatro (94) escolas da região metropolitana de São Paulo,⁴ tendo como participantes oito (8) professores de cada escola, sendo sete (7) de distintas áreas de conhecimento, além do coordenador pedagógico.

Tendo como objetivo do subprojeto desenvolver uma metodologia de formação continuada do educador para o uso do computador no ensino e na aprendizagem, as ações de formação se realizaram nos laboratórios de informática das escolas, em três módulos de oficinas teórico-práticas, com seminários para trocas de experiências nos intervalos entre os módulos realizados no auditório da PUC-SP.

Nas oficinas, eram criadas as condições para que o professor pudesse explorar recursos computacionais aplicáveis à educação, enquanto era provocado a refletir sobre teorias, percepções e crenças a respeito de conhecimento, ensino e aprendizagem e a estabelecer inter-relações com sua prática pedagógica tendo em vista compreender como, quando, por que e para que integrar o computador a essa prática. A ação-reflexão-depuração-nova ação foi enfatizada em “um processo que articula a técnica, a reflexão, a prática pedagógica e as teorias educacionais subjacentes à prática, segundo a natureza da situação contextual” (Pellegrino et alii, 1998).

A formação constituiu uma atividade educativa entre educadores no exercício de suas funções e com efetiva participação na formação, contribuindo com suas experiências, competências, potencialidades e representações. Assim, após a exploração de determinado *software* e análise de suas potencialidades, os professores desenvolviam atividades de uso do computador com seus alunos, as quais eram acompanhadas pelos formadores, que orientavam os professores/alunos no sentido de ajudá-los a

3 Maiores detalhes em <http://www.educacao.sp.gov.br/> (capturado em maio de 2000).

4 O PEC-IEB da Secretaria de Estado da Educação de São Paulo, SEE, realizou-se com o apoio financeiro do Banco Mundial e envolveu diversas áreas de conhecimento, definidas segundo as demandas das Delegacias de Ensino da região de abrangência do Programa. A PUC-SP assumiu a parceria do Programa PEC-PUC-SP no Pólo 4, que envolvia as seguintes Delegacias de Ensino: 3ª e 4ª DE da Capital, 1ª e 2ª DE de Guarulhos e DE de Caieiras.

compreender os desafios e a buscar alternativas para sobrepujá-los. Os relatos retornavam para discussão no grupo em formação, servindo de espelho para todos reverem as próprias práticas nas análises das experiências dos colegas.

Trata-se de formação segundo uma concepção de continuidade, de processo, não buscando um produto pronto, mas sim a criação de um movimento, cuja dinâmica se cria na reflexão na ação e na reflexão sobre a ação (Shön, 1992), ação esta experienciada durante a formação, recontextualizada na prática do formando e refletida pelo grupo em formação, realimentando a formação, a prática de formandos e formadores e as teorias que a fundamentam.

Não se tratava de uma formação voltada para atuação no futuro, mas sim de uma formação direcionada para o presente, tendo como pano de fundo a ação imediata do educador. Procurava-se estabelecer uma congruência entre o processo vivido pelo educador formando e sua prática profissional.

A partir da convivência com os desafios e outros fatores que interferiam no trabalho educativo, na busca conjunta de alternativas para sobrepujar as dificuldades, no compartilhamento de conquistas e fracassos, nas reflexões na e sobre a própria ação, o educador tinha a possibilidade de compreender o que, como, por que e para que (Imbernón, 1998) empregar o computador em sua ação.

Após a conclusão da formação observou-se um recuo na concretização de práticas pedagógicas com o computador em relação ao que ocorria durante o período da formação, o que evidenciou a relevância da problemática como objeto de investigação.

Concepção da formação

A formação de educadores para a incorporação da TIC à prática pedagógica parte de pressupostos voltados para a mudança da escola favorecida pela apropriação e utilização pedagógica dessa tecnologia, de modo a propiciar ao formando condições de desenvolver crítica e reflexivamente um estilo próprio de atuar com a tecnologia, baseado na espiral de aprendizagem descrita por Valente como *descrição-execução-reflexão-*

depuração (Valente, 1999), na qual a intervenção de um agente de aprendizagem é essencial para provocar a reflexão e a depuração do processo em desenvolvimento e propiciar a aprendizagem.

No processo de aprendizagem, pode-se recorrer a qualquer momento a um dos elementos da espiral, mas a recursividade ascendente pode não ocorrer espontaneamente ou ocorrer de forma inadequada quando o aprendiz (o professor/aluno) encontra-se diante de um conflito e não está conseguindo identificar seus equívocos. Cabe ao professor-mediador (ou ao colega, parceiro da atividade) intervir na atividade do aprendiz, incentivá-lo, provocar questionamentos para ajudá-lo a explicitar objetivos, identificar conhecimentos e estratégias empregados, interpretar resultados, compreender e corrigir equívocos e voltar a realizar a atividade, possivelmente com um nível superior de compreensão.

O formador não indica caminhos, mas ajuda o professor/aluno a seguir sua própria trajetória, orienta-o na busca de conceitos, fornece informações pertinentes e evita situações em que ele possa sentir-se constrangido, incapaz e até abandonar o trabalho. Por meio da interpretação, articulação entre informações com conhecimentos anteriormente adquiridos, reflexão e depuração, é possível reorganizar as estruturas cognitivas e construir o novo conhecimento.

Portanto, a formação do educador procura propiciar-lhe condições de mergulhar na própria aprendizagem, refletir sobre como se aprende e se ensina e como inserir a tecnologia computacional em sua prática pedagógica com vistas à aprendizagem do aluno. Daí a abordagem de formação contextualizada nas problemáticas existentes no tempo e no espaço da instituição educacional. Porém, isso não significa, necessariamente, que a formação deva se realizar fisicamente na escola e, sim, que as necessidades da formação emergem do contexto educacional no qual se busca desenvolver uma cultura que permita ao educador tornar-se usuário crítico da tecnologia, incorporá-la à prática pedagógica e assumir-se como um agente de mudança de sua própria atuação e de seu contexto. Evidencia-se então a epistemologia da prática mediante um processo em que prática e teoria integram-se em um *currículo orientado para a ação*.

Formadores e formandos constituem um grupo em formação. Todos têm a oportunidade de aprender em *situações problemáticas contextualizadas*, as quais são analisadas pelo grupo em um processo

contínuo de investigação, interação, cooperação e socialização, caracterizando um enfoque de formação contextualizada. Cria-se assim, um clima propício para uma atuação que favorece a transformação social e a emancipação humana, estabelecendo uma *práxis* contextualizada, cuja frequência das interações e comunicações indica a ocorrência de mudanças *gestadas* nas escolas (Imbernon, 1998).

Reelaborar a vivência da formação para a atuação em sala de aula usando o computador com seus alunos permite ao professor/aluno recontextualizar as atividades experienciadas na formação para sua prática pedagógica. Essas práticas recontextualizadas eram analisadas pelo grupo em formação, cuja reflexão propiciava a busca de teorias para melhor compreendê-la e a reconstruí-la. Desta forma, a teoria iluminava a compreensão e a reelaboração da prática, assim como a prática refletida conduzia à reconstrução da teoria.

O ambiente de formação continuada não se restringe a situações agendadas para cursos, oficinas, conferências e reuniões, pois a perspectiva de continuidade é fundamental para instaurar um processo de reflexão, investigação, revisão e intervenção sobre situações-problema existentes na prática real. As bases teóricas da formação são "reinterpretadas à luz da solução que requer, ou seja, levando em consideração a modificação de determinada realidade" (Imbernon, 1998, p. 84).

O estabelecimento de espaços coletivos de discussão e reflexão permite que o professor/aluno participante da formação se posicione como aprendiz em ambientes de aprendizagem informatizados, em sintonia com o processo de aprendizagem de seu aluno. O formando tem a oportunidade de compreender suas próprias potencialidades, conflitos e dificuldades e, ao mesmo tempo, refletir sobre seu trabalho e sua atuação com uma *crítica justa e objetiva que o acompanha, associada ao desejo de acertar, de fazer sempre o melhor* (Ribas, Carvalho e Alonso, 1999, p. 50).

A tecnologia propicia a interação entre os participantes (formandos e formadores) e cria condições favoráveis para o desenvolvimento da autonomia para utilizá-la na busca de informações, na construção de conhecimentos de distintas áreas do saber, na representação e resolução de um problema contextual ou no desenvolvimento de um projeto.

Assim, a formação se desenvolve a partir dos fundamentos da formação contextualizada na realidade da escola e na prática pedagógica do

professor, articulada ao domínio do computador e sua incorporação a essa prática.

A fim de subsidiar a realização de outras ações correlatas, torna-se relevante desenvolver investigações que levem a compreender os avanços dessa formação, suas idiossincrasias, limitações e potencialidades.

Metodologia da pesquisa

A problemática da pesquisa consistia em buscar compreender a formação de professores para a inserção do computador na prática pedagógica e suas respectivas conseqüências, procurando indícios de possíveis transformações na prática docente e suas influências na escola.

A pesquisa foi realizada com o objetivo de analisar o subprojeto Informática na Educação do PEC/PUC-SP (1997/1998) e as respectivas influências nas práticas pedagógicas dos professores que participaram da formação, tendo em vista:

- buscar indícios de mudanças na atuação do professor;
- refletir sobre avanços, potencialidades, desafios, dificuldades, incoerências e ambigüidades;
- recontextualizar a teoria sobre a formação de professores para a incorporação do computador à prática pedagógica.

Para melhor situar o caso em estudo, evidencia-se a necessidade de relacionar aspectos quantitativos e aspectos qualitativos, sem uma oposição entre os dois, mas com a preocupação de articulá-los de modo a melhor compreender o fenômeno, fortalecer a argumentação e auxiliar na elucidação e análise crítica (Thiollent, 1984, p. 48).

Na etapa de seleção das escolas-foco das análises, parte das informações foram processadas por computador, empregado como meio para facilitar a análise dos dados, tanto quantitativos, quanto qualitativos. Assim, o envolvimento da tecnologia nessa etapa da pesquisa foi *de caráter mais técnico do que conceitual*, o que, segundo Bogdan e Biklen (1994, p. 43), é *uma inovação significativa*.

A coleta de dados para aprofundamento das análises ocorreu seis meses após a conclusão das ações de formação, quando foram selecionadas três escolas,⁵ nas quais foram entrevistados quatorze professores, sendo quatro efetivos na escola, dois estáveis e oito contratados para suprir carências eventuais. Os outros dez não atuavam mais na escola em que participaram da formação.

O plano da pesquisa caracterizou-se como um projeto aberto e flexível, que foi se estruturando durante a investigação, realizada após a conclusão das ações de formação. Procurava-se não identificá-lo com idéias preconcebidas ou com um plano prévio esmiuçado detalhadamente. As questões iniciais foram formuladas com o objetivo de serem reelaboradas à medida que os estudos evoluíam. Os dados e as informações obtidos indicavam os passos a seguir. O plano, em seu todo, assim como a coleta e a análise dos dados, foram continuamente revistos, refletidos e depurados durante a investigação.

A fim de evitar o direcionamento provocado pelas definições prévias de hipóteses e de objetivos que *delimitam o enfoque e reduzem o interesse pela situação e circunstância*, adotou-se como unidade de análise desse estudo de caso instrumental (Stake, 1998, p. 26) as categorias e os respectivos temas indicadores decorrentes dos marcos teóricos da formação.

5 Foram selecionadas três escolas (3), levando-se em conta os seguintes critérios:

1. O capacitador foi a mesma pessoa nos três módulos;
2. O laboratório de informática estava instalado e em funcionamento na escola desde o início das ações do PEC/PUC-SP;
3. A escola foi sede da capacitação em pelo menos um dos módulos da formação;
4. Houve participação dos professores da escola com apresentação de trabalhos nos dois seminários para apresentação das experiências, realizados pelo PEC/PUC-SP;
5. Os coordenadores e dirigentes (diretor ou vice-diretor) da escola participaram das oficinas realizadas para os dirigentes;
6. Os dirigentes manifestaram interesse e aceitaram a participação da escola na pesquisa;
7. A escola situa-se em local de fácil acesso e comunicação via telefone ou fax.

O software CHIC na análise dos dados

Considerando-se que “estudo de caso não é um pacote metodológico padronizado, isto é, não é um método específico de pesquisa, mas uma forma particular de estudo, cuja metodologia é eclética” (André, 1984, p. 52), na coleta de dados foram utilizadas descrições, entrevistas, gravações em fita cassete e vídeo, questionários, escritos dos alunos.

Todo o conteúdo coletado revela as manifestações dos sujeitos baseadas em sua memória sobre a percepção dos fatos relacionados a um programa de formação. Essas manifestações foram emitidas a um entrevistador, que foi também o responsável pelas sucessivas leituras e asserções, além de ser sujeito da pesquisa com um papel preponderante no processo em análise. Portanto, não existe um coeficiente independente, que permita medir o grau de confiabilidade dos dados, que poderiam vir a ter interpretações e significados diferentes para outros pesquisadores.

A partir dos instrumentos de registro dos dados, foram realizadas sucessivas leituras em busca de uma compreensão das falas e do contexto em análise. Com o movimento de releitura dos depoimentos, foram extraídos os conceitos considerados relevantes até chegar aos temas emergentes, considerando-se a presença ou a ausência de conteúdos relacionados com os temas indicadores oriundos das categorias ou a possível presença de outros temas não previstos inicialmente.

Para melhor identificar tais temas, além da análise elaborada preliminarmente, os dados obtidos nas entrevistas e nos questionários foram submetidos a outro investigador, que utilizou os marcos teóricos e respectivas categorias para identificar os temas por ele considerados emergentes. Em seguida, realizou-se o cruzamento entre os temas emergentes levantados pelos dois investigadores, analisando convergências e divergências até chegar a um entendimento consensual entre a definição das categorias e a identificação dos temas emergentes.

Nesse momento, evidenciou-se a problemática desse tipo de análise de dados, agravada pela enorme quantidade de dados coletados: a leitura e interpretação dos dados exige a definição de uma metodologia que articule o qualitativo ao quantitativo com o intuito de identificar as operações adequadas para transformar os dados em objetos de conhecimento (Maraschin, 1996).

Na busca de encontrar um método de análise que permitisse sistematizar e visualizar conexões entre dados processados por recursos computacionais, identificaram-se dois grandes grupos de métodos de análise: os métodos de análise fatorial e os métodos de classificação automática (Volle, 1985, p. 17). Considerando-se que “as análises multidimensionais permitem obter sínteses de dados, em visão holográfica, dos fatores discriminantes (as análises fatoriais), das tipologias (classificação hierárquica), das hierarquias (análise implicativa e hierarquia implicativa de classes) etc”. (Almouloud, 1997, p. 165), optou-se pela análise hierárquica de similaridade ou semelhança para a organização dos dados e a respectiva análise, de acordo com sua intersecção, com o uso do software CHIC – Classification Hiérarchique Implicative et Cohésitive.

A análise multidimensional é uma importante ferramenta para as pesquisas em ciências humanas, cujas apreciações qualitativas correm o risco de limitar-se a expressões demasiadamente vagas, do tipo: “os professores disseram que...”, “os professores acreditam que...”, “nós pensamos que...”, como poderia ser o caso da presente investigação, baseada em respostas dos professores a questionários e entrevistas semi-estruturadas. Para evitar essa limitação, além da intuição e do cruzamento das interpretações de dois pesquisadores sobre os fatos observados, o uso do *software* CHIC viabilizou construir e visualizar as significações a partir de aproximações, semelhanças, contradições ou repetições, revelando as concepções dos sujeitos e fornecendo informações inacessíveis por métodos simétricos clássicos (Gras, 1996).

Os temas indicadores foram transformados em variáveis, cuja presença ou ausência foi analisada em correspondência com os temas emergentes e identificadas as interseções que representavam convergências entre os temas, ou seja, a vizinhança entre dois ou mais temas. A vizinhança entre os temas é maior quanto mais próximos eles estiverem da condição “ $A \cup B$ ” ou “*A união com B*” (Maraschin, 1996, 1995, p. 101). Os temas emergentes que não tinham correspondência com os temas indicadores relacionados às categorias não foram analisados pelo *software*, mas sim tratados qualitativamente por meio de interpretações.

As sucessivas aplicações do CHIC permitiram vislumbrar os temas que sobressaíram na percepção dos sujeitos e nas respectivas articulações. Desse modo, a questão sobre o estudo de caso instrumental do

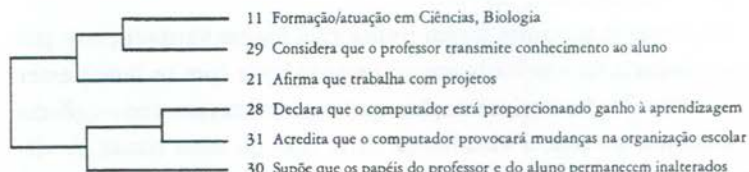
PEC/PUC-SP pôde ser tratada em sua complexidade e diversidade típica de caso, compreendido por meio da análise das similaridades e singularidades das três escolas e pelas percepções dos sujeitos envolvidos.

Resultados da análise

Para efeito das análises das convergências entre os depoimentos dos professores, foram empregadas as árvores de similaridade fornecidas pelo CHIC, cujas associações significativas foram representadas em classes, conforme as relações que se explicitavam com maior força.

Classe P1: Tensão entre estabilidade e mudança

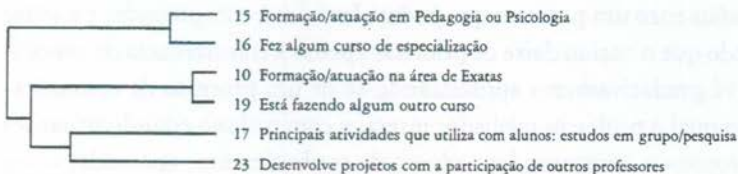
Esta classe mostra professores com formação em Ciências ou Biologia, para os quais a inserção do computador está proporcionando ganhos à aprendizagem do aluno e deverá provocar mudanças na organização escolar. Não há indicações de perceberem alterações nos papéis do professor e do aluno, embora afirmem trabalhar com projetos, dando a impressão de que essas mudanças ainda estão fora do âmbito de sua atuação. Portanto, esses professores parecem ter uma atuação que oscila entre uma perspectiva inovadora e uma prática conservadora, convivendo na tensão entre a mudança e a estabilidade.



Classe P2: Perspectiva inovadora da prática pedagógica

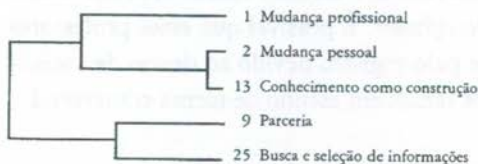
Esta classe diz respeito a professores com formação e/ou atuação em Pedagogia, Psicologia ou em disciplinas da área de exatas, que anunciavam uma prática próxima de uma perspectiva inovadora, não explicitando contradição em seus depoimentos. Há indícios de que esses professores

buscam uma melhoria em sua formação, transferem possíveis vivências da formação para a sua prática pedagógica, já fizeram algum curso de especialização e estavam participando de outro curso durante a realização da capacitação do PEC/PUC-SP.



Classe P3: Mudança em construção

A classe P3 indica a existência de professores cujos depoimentos articulam três importantes ingredientes de um ambiente de aprendizagem com a presença do computador: a abordagem pedagógica identificada pela congruência entre mudança pessoal e conhecimento como construção; o estabelecimento de parcerias na busca e seleção de informações; a mudança profissional. Essa classe fornece indícios de predisposição para a mudança da prática, processo esse que parece encontrar-se em andamento. Os professores estão assumindo o discurso de uma atuação inovadora, embora possa existir uma distância considerável entre essa atuação anunciada e a sua prática concreta.

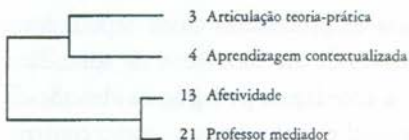


Classe P4: Mediação contextualizada

A classe P4 apresenta nós que complementam a classe anterior. Os depoimentos dos professores indicam uma atuação como mediador da aprendizagem do aluno, no qual a afetividade se faz presente, junto com aspectos sociocognitivos. Dessa forma, o professor demonstra levar em

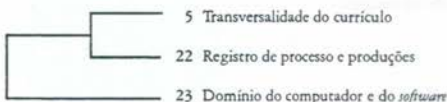
conta as contribuições do aluno em relação ao conhecimento de seu cotidiano, procurando escutá-lo, desafiá-lo e incitá-lo a compreender o seu contexto, isto é, a transformar o conhecimento do senso-comum em conhecimento científico.

Para os professores que se expressaram dessa forma, parece estar em andamento um processo que poderá levá-los a compreender e a atuar de modo que o ensino deixe de priorizar apenas a transferência de conteúdos e vá gradativamente aproximando-se de um processo de comunicação, no qual o professor mediador interage com o aluno criando situações que propiciam compreender o objeto de conhecimento, apreender o seu significado e atuar em seu mundo, transformando-se na convivência com o outro.



Classe P5: Registro da transversalidade

Essa classe indica a existência de professores que percebem a importância do registro de processo e produções no computador relacionado ao trabalho com temas transversais ao currículo e a respectiva articulação entre conteúdos de distintas áreas. Para isso, consideram necessário o domínio do computador e do *software*. É possível que esses professores tenham demonstrado interesse pelo registro devido ao desejo de socializar suas experiências e tratar os temas em estudo de forma transversal.



A preocupação com o registro, que aparece na classe P5, pode estar implícita nas duas classes anteriores, P3 e P4, ou pode indicar uma ambigüidade devida à preocupação com a cultura dos produtos e respectivas

divulgações. Em várias situações observadas nas escolas, evidenciou-se o incentivo dos formadores em relação ao registro das produções no computador, mas não havia até então a cultura de registrar o processo.

Assim, há indícios de coerência parcial nos depoimentos dos professores em relação aos fundamentos teóricos da formação desenvolvida. Essa coerência pode ser observada pelos seguintes aspectos: processo de mudança pessoal e profissional em andamento; tentativa do professor de atuar como mediador na promoção da aprendizagem do aluno; conhecimento considerado fruto de um processo de construção no qual a afetividade é um dos componentes; aquisição relativa do domínio de recursos computacionais; valorização do registro para o estudo e a representação de temas transversais; tomada de consciência sobre a importância das conexões entre pensamento, emoções e sentimento, tendo em vista o desenvolvimento humano em sua inteireza.

A identificação de temas que não apareceram nas árvores de similaridade permitiu analisar o que estava presente nas intenções e pressupostos do subprojeto PEC/PUC-SP e não ocorreu na formação ou não emergiu nos depoimentos dos sujeitos. Assim, a primeira constatação que salta aos olhos diz respeito a temas relacionados com a autonomia, que aparecem apenas parcialmente nos depoimentos investigados.

A ênfase da formação relacionava-se com o domínio do computador e do *software* e o uso da tecnologia com alunos, o que se revela nos depoimentos, denotando ser essa uma das prioridades com forte influência nas ações dos professores. A maioria dos professores que estava utilizando o computador com alunos demonstrava maior preocupação com o domínio da tecnologia em si mesma, ou com o conteúdo em estudo, do que com a metodologia do trabalho em desenvolvimento, sem preocupar-se se deveria trabalhar com o desenvolvimento de projetos em sala de aula ou com outras atividades, desde que significativas para o aluno.

Há indícios de que o professor reconheceu a necessidade de dominar os recursos computacionais para usá-los na criação de situações de aprendizagem significativas ao aluno, bem como para fornecer informações pertinentes sobre as operações e ferramentas do *software*, as fontes e os respectivos conteúdos, a proposição de desafios, a intervenção e a orientação das ações do aluno, etc. No entanto, não se pode afirmar que essa tomada de consciência venha a concretizar a incorporação do computador à prática pedagógica.

Conclusão

As análises realizadas neste estudo evidenciaram que um programa de formação contextualizada de professores para a incorporação da tecnologia de informação e comunicação à prática pedagógica tem condições de criar ambientes de aprendizagem que favoreçam os participantes:

- aprender a aprender para resolver problemas com que se depa-ram na vida e na profissão;
- assumir uma atitude de abertura para o novo, o inesperado e o imprevisível;
- dominar recursos dessa tecnologia, usá-los em sua prática conforme os objetivos pedagógicos e orientar os alunos para a seleção de recursos mais adequados à atividade em desenvolvimento;
- compreender como se aprende e como se ensina com o uso da tecnologia;
- descobrir processos e resultados positivos do uso da tecnologia em sua prática;
- criar ambientes de aprendizagem, nos quais a tecnologia é utilizada pelo aluno para busca, articulação e troca de informações e experiências, resolução de problemas e reconstrução contínua do conhecimento, reflexão, interação e colaboração;
- questionar crenças e práticas institucionais;
- investigar a própria ação e formação, tomar consciência de suas dificuldades e das estratégias adotadas para sobrepujá-las;
- desenvolver a autonomia para tomar decisões em relação aos recursos e às metodologias a utilizar em sua prática pedagógica e para intervir no processo de aprendizagem individual e grupal.

Ao colocar o professor em relação direta com a própria prática e acentuar o seu papel como agente de mudança e de aprendizagem ao longo da vida, a formação altera o contexto da formação. O educador torna-se co-autor do planejamento, desenvolvimento, reflexão (avaliação processual e de resultados) e depuração (reformulação de metodologias) da formação continuada.

O deslocamento do eixo da formação para o contexto escolar apresenta grande complexidade de operacionalização e seu sucesso depende

diretamente de uma ação colaborativa que envolva um contingente considerável de professores e gestores educacionais comprometidos com esse processo. O subprojeto Informática na Educação do PEC/PUC-SP teve atenção especial para com esse aspecto, mas a maioria dos professores participantes não tinha um vínculo efetivo na escola e pouco pôde fazer em termos da disseminação da formação. As escolas com maior quantidade de professores efetivos conseguiram dar continuidade às atividades de uso do computador com alunos, apesar das diversas mudanças de direção. Seus professores mostravam-se mobilizados e insistiam na criação de um espaço coletivo para o diálogo, a discussão de suas propostas, a troca de experiências e a disseminação da formação com os colegas.

Embora as ações efetivas de uso do computador com alunos em várias situações continuassem como ação dicotomizada do âmbito da sala de aula, não tendo se incorporado às práticas cotidianas, uma conquista revelada na análise dos dados explicita a condição essencial do ensino comprometido com a aprendizagem como processo de construção do conhecimento.

Na abordagem de formação contextualizada baseada no ciclo *descrição-execução-reflexão-depuração*, quando a tecnologia começa a adentrar os espaços educacionais, há um momento em que seu uso se apresenta como prática de um pequeno grupo de educadores e alunos. Esses autores de primeira mão vão gradativamente conquistando outros parceiros, até que exista um contingente de profissionais e de alunos trabalhando, que pode induz a inserir tais práticas como uma das atividades inovadoras contempladas no plano político-pedagógico da instituição e assumidas pela coletividade.

Portanto, a formação contextualizada de educadores para a incorporação da tecnologia de informação e comunicação tem como eixo norteador a escola e as experiências, conhecimentos e práticas do educador, propiciando a este tornar-se um investigador reflexivo da própria prática, cuja formação ocorre na *práxis*, favorecendo mudanças pessoais e profissionais que influem em mudanças na prática pedagógica e na escola. Formadores e formandos constituem um grupo em formação, que compartilha práticas, reflexões, conquistas e restrições. Trata-se de uma formação que se desenvolve na transversalidade do currículo; inter-relaciona formação, ação e reflexão, realidade e conteúdo, homem e máquina, arte e tecnologia, teoria e prática, razão e emoção.

Referências

- ALMEIDA, M. E. B. (2000). *O computador na escola: contextualizando a formação de professores*. Tese de doutorado. Programa de Pós-Graduação em Educação: Currículo. São Paulo, Pontifícia Universidade Católica.
- _____. (1999). O aprender e a informática: a arte do possível na formação do professor. *Cadernos Informática para a Mudança em Educação*. MEC/SEED/ ProInfo.
- ALMOULOU, S. A. (1997). Fundamentos da didática da matemática e metodologia de pesquisa. Volume III. *CEMA: Caderno de Educação Matemática*. Programa de Estudos Pós-Graduados no Ensino de Matemática.
- ANDRÉ, M. E. D. (1984). Estudo de caso: seu potencial na educação. *Cadernos de Pesquisa*, n. 49, pp. 51-54.
- BOGDAN, R. e BIKLEN, S. (1994). *Investigação qualitativa em educação. Uma introdução à teoria e aos métodos*. Portugal, Porto Editora.
- GRAS, R. (1996). *Nouvelle Méthode Exploratoire de Données*. Paris, La Pensée Sauvage.
- IMBERNÓN, F. (1998). *La formación y el desarrollo profesional del profesorado. Hacia una nueva cultura profesional*. Barcelona, Editorial Graó, de Serveis Pedagògics.
- MARASCHIN, C. (1996). *O escrever na escola: da alfabetização ao letramento*. Tese de doutorado. Programa de Pós-Graduação em Educação. Porto Alegre, Universidade Federal do Rio Grande do Sul.
- _____. (1995). *Mapeando ecologias cognitivas*. In: VII CONGRESSO INTERNACIONAL LOGO e I CONGRESSO DE INFORMÁTICA EDUCATIVA DO MERCOSUL. *Anais*, Porto Alegre, RS, LEC/UFRGS.
- PELLEGRINO, C. N.; SCHLÜNZEN, E. T. M.; SCHLÜNZEN, K. Jr.; MORELATTI, M. R. M.; ALMEIDA, M. E. e HERNANDES, V. K. (1998). *Informática na Educação: vivenciando novas experiências nas escolas da rede pública de São Paulo* [cd-rom]. V CONGRESSO ESTADUAL PAULISTA SOBRE FORMAÇÃO DE EDUCADORES. *Anais*, Águas de São Pedro.

- PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO (1998). *Inovações em educação pública: memórias, utopias e práticas*. São Paulo, Estação Palavra.
- RIBAS, M.; CARVALHO, M. A. e ALONSO, M. (1999). O caráter emancipatório de uma prática pedagógica possível. In: ALONSO, M. (org.). *O trabalho docente: teoria e prática*. São Paulo, Pioneira.
- SCHÖN, D. A. (1992). Formar professores como profissionais reflexivos. In: NÓVOA, A. (org.). *Os professores e sua formação*. Lisboa, Dom Quixote.
- STAKE, R. E. (1998). *Investigación com estudio de casos*. Madrid, Ediciones Morata.
- THIOLLENT, M. (1984). Aspectos qualitativos da metodologia de pesquisa com objetivos de descrição, avaliação e reconstrução. *Cadernos de Pesquisa*, n. 49, pp. 45-50, maio.
- VALENTE, J. A. (org.) (1999). "O computador na sociedade do conhecimento". In: VALENTE, J. A. (org.). *O computador na sociedade do conhecimento*. Campinas, Unicamp/Nied.
- VALENTE, J. A. e ALMEIDA, F. J. (1997). Visão analítica da informática no Brasil: a questão da formação do professor. *Revista Brasileira de Informática na Educação-SBIE*, n. 1.
- VOLLE, M. (1985). *Analyse des Données*. 3 ed. Paris, Economica.

Recebido em abr./2002; aprovado em maio/2002

THE UNIVERSITY OF CHICAGO
DIVISION OF THE PHYSICAL SCIENCES
DEPARTMENT OF CHEMISTRY

REPORT OF THE
COMMISSION ON THE
STRUCTURE OF THE
ATOMIC NUCLEUS
AND THE
PROPERTIES OF
NUCLEAR PARTICLES

BY
RICHARD FEYNMAN
AND
MORSE TELLER

Submitted to the
National Academy of Sciences
and the National Bureau of Standards
for their consideration

in connection with the
National Science Foundation
and the Atomic Energy Commission
program on the structure of the
atomic nucleus

CHICAGO, ILLINOIS
1951

Published by the
University of Chicago Press

Price \$10.00

Dissertações defendidas no segundo semestre de 2002*

ARAÚJO, C. C. V. B. *A matemática no livro didático de álgebra linear*
(*The Mathematics in the Didactic Book of Linear Algebra*)

Palavras-chave: álgebra linear; livro didático; base de um espaço vetorial.
Key-words: linear algebra; didactic book; base of a vectorial space.

ARBACH, N. *O ensino de geometria plana: o saber do aluno e o saber escolar*
(*Ensino de Plain Geometria: Saber of the Pupil and Pertaining to School Saber*)

Palavras-chave: ensino; geometria; demonstração.
Key-words: education; geometry; demonstration.

DUARTE, A. R. S. *Henri Poincaré e Euclides Roxo: subsídios para a história
das relações entre filosofia da matemática e educação matemática*
(*Henri Poincaré and Euclides Roxo: Subsidies for the History of the Relations
between Philosophy of the Mathematics and Mathematical Education*)

Palavras-chave: educação matemática; história da matemática; Euclides
Roxo.
Key-words: mathematical education; history of the mathematics;
Euclides Roxo.

* As dissertações completas estão divulgadas no site: www.pucsp.br/pos/edmat

GODOY, E. V. *Matemática no ensino médio: prescrições das propostas curriculares e concepções dos professores*

(Mathematics in Average Ensino: Lapsing of the Curricular Proposals and Conceptions of the Professors)

Palavras-chave: currículos; ensino médio; contextualização; interdisciplinar.

Key-words: resumes; average education; contextualização; to interdisciplinar.

LELLIS, M. C. T. *Sobre o conhecimento matemático do professor de matemática (On the Mathematical Knowledge of the Professor of Mathematics)*

Palavras-chave: formação de professores; conhecimento matemático; significados matemáticos.

Key-words: formation of professors; mathematical knowledge; mathematical meanings.

MACHADO, R. C. G. *Uma análise dos exames de admissão ao secundário (1930-1970): subsídios para a história da educação matemática no Brasil*

(An Analysis of the Examinations of Admission to the Secondary one (1930-1970): Subsidies for the History of the Mathematical Education in Brazil)

Palavras-chave: provas de matemática; exame de admissão; legislação, ensino.

Key-words: tests of mathematics; examination of admission; legislation, education.

MAIOLI, M. *Uma oficina para formação de professores com enfoque em quadriláteros*

(A Workshop for Formation of Professors with Approach in Quadrilaterals)

Palavras-chave: formação de professor; teoria das situações; registros.

Key-words: formation of professor; theory of the situations; registers.

NOTARI, A. M. *Simplificação de frações aritméticas e algébricas: um diagnóstico comparativo dos procedimentos*

(Simplification of Arithmetical and Algebraic Fractions: A Comparative Diagnosis of the Procedures)

Palavras-chave: equivalência de expressões algébricas; análise de erros.

Key-words: equivalence of algebraic expressions; error analysis.

POSSANI, R. A. R. *Apreensões de representações planas de objetos espaciais em um ambiente de geometria dinâmica*

(Apprehensions of Plain Space Object Representations in an Environment of Dynamic Geometry)

Palavras-chave: apreensão operatória; diferenciação de planos.

Key-words: operatória apprehension; differentiation of plans.

SANTOS, A. M. *Mensuração, Algarismos significativos e notação científica: um estudo diagnóstico do processo ensino-aprendizagem, considerando o cálculo e a precisão de medidas*

(Significant Mensuração, Numbers and Scientific Notation: A Diagnostic Study of the Process Teach-learning, considering the Calculation and the Precision of Measures)

Palavras-chave: mensuração; algarismos significativos; notação científica.

Key-words: significant measurement; numbers; scientific notation.

SANTOS, E. P. *Função afim $y=ax+b$: a articulação entre os registros gráfico e algébrico com o auxílio de um software educativo*

(Similar function $y=ax+b$: The joint between the registers graphical and algebraic with the aid of an educative software)

Palavras-chave: função afim; registro de representação; software educativo.

Key-words: similar function; register of representation; educative software.

TRALDI JÚNIOR, A. *Sistema de inequações do 1º grau: uma abordagem do processo ensino-aprendizagem, focando os registros de representações*

(System of Inequações of 1º Degree: A boarding of the Ensino-Aprendizagem Process, focando the Registers of Representations)

Palavras-chave: resolução de problemas; seqüência didática; otimização.

Key-words: resolution of problems; didactic sequence; otimização.

THE UNIVERSITY OF CHICAGO
DIVISION OF THE PHYSICAL SCIENCES
DEPARTMENT OF CHEMISTRY
5780 SOUTH CAMPUS DRIVE
CHICAGO, ILLINOIS 60637

MEMORANDUM FOR THE RECORD
SUBJECT: [Illegible]

DATE: [Illegible]

TO: [Illegible]

FROM: [Illegible]

Normas para publicação

Pesquisadores interessados em contribuir com publicação nesta revista deverão preparar o texto e enviá-lo segundo as regras que se seguem.

Preparação para envio – uma cópia do texto em disquete(s) com os nomes dos autores e sem numeração de página. Outras quatro (4) cópias impressas, sendo que uma deve ser idêntica à(s) do(s) disquete(s) e as outras três (3) devem ter numeração de página e não trazer os nomes dos autores.

Versão – programa Word 6.0 for Windows, para ser lido em PC.
Formatação

Título – centralizado, em letras maiúsculas e em negrito.

Nomes dos autores – em uma só das vias impressas e no disquete, separar os nomes dos autores do título por um espaço simples entre linhas. Os dados de cada autor deverão ser colocados conforme exemplo, abaixo do título.

Ex: Maria Dolores da Silva

Mestre em Educação Matemática – PUC-SP

Professora do Curso de Matemática – PUC-SP

e-mail: dolores@pucsp.br

Resumo – em português e inglês ou francês, com, no máximo, 10 linhas, espaço duplo, mesma fonte do texto, em itálico, acompanhado de três palavras-chave.

Corpo do texto – Papel tamanho A4

Margem superior e inferior com 2,5 cm

Margem direita e esquerda com 3,0 cm

Fonte Times New Roman,

Tamanho da letra 12 pontos.

Espaçamento entre linhas 1,5 linha

Alinhamento justificado

Referências bibliográficas – de acordo com as normas da ABNT em vigor.

Exemplos:

- Livro
GOMES, L. G. F. (1998). *Novela e sociedade no Brasil*. Niterói, EdUFF (Coleção Antropologia e Ciência Política, 15).
- Tese
BARCELOS, M. F. P. (1998). *Ensaio tecnológico, bioquímico e sensorial de soja e guandu enlatados no estádio verde e maturação de colheita*. Tese de doutorado em Nutrição. Campinas, Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas.
- Artigo de revista
GURGEL, C. (1997). Reforma do Estado e segurança pública. *Política e Administração*, Rio de Janeiro, v. 3, n. 2, pp. 15-21, set.

Citações no texto – citações no texto devem vir acompanhadas de sobrenomes(s) do(s) autor(es) em corpo menor e entre parênteses, acrescido do ano de publicação e página.

Tabelas e gráficos – deverão ter como elementos: número, título, data de referência, fonte e nota.

Impressão – em jato de tinta ou em laser. Páginas impressas só numa face.

Os trabalhos devem ser enviados para:

Revista Educação Matemática Pesquisa

Programa de Estudos Pós-Graduados em Educação Matemática

Rua Marquês de Paranaguá, 111 - Consolação - SP - CEP 01303-050

Fone: (11) 3124-7210

Fax: (11) 3159-0189

e-mail: pgedmat@exatas.pucsp.br

EDUCAÇÃO MATEMÁTICA PESQUISA

REVISTA DO PROGRAMA DE ESTUDOS PÓS-GRADUADOS
EM EDUCAÇÃO MATEMÁTICA PUC-SP

Educação Matemática Pesquisa publica trabalhos voltados para as linhas de pesquisa: *A Matemática na estrutura curricular e a Formação de Professores; Epistemologia e Didática da Matemática; Tecnologias de Informação e Didática da Matemática*. Também está aberta para outros campos do conhecimento, que venham proporcionar um diálogo com a área, como a Epistemologia, a Psicologia Educacional, a Filosofia, a História das Ciências e a História Disciplinar.

INFORMAÇÕES PARA AQUISIÇÃO

Anexo cópia do depósito em conta no **Banco Bradesco**, agência 3227-1, c/c 1285-8, favorecido **Sonia B. C. Iglioni**, para aquisição dos seguintes exemplares de *Educação Matemática Pesquisa*:

- | | | |
|------------------------------------|------------------------------------|------------------------------------|
| <input type="checkbox"/> v. 1 n. 1 | <input type="checkbox"/> v. 1 n. 1 | <input type="checkbox"/> v. 3 n. 1 |
| <input type="checkbox"/> v. 1 n. 2 | <input type="checkbox"/> v. 1 n. 2 | <input type="checkbox"/> v. 3 n. 2 |
| <input type="checkbox"/> v. 2 n. 1 | <input type="checkbox"/> v. 2 n. 1 | <input type="checkbox"/> v. 4 n. 1 |
| <input type="checkbox"/> v. 2 n. 2 | <input type="checkbox"/> v. 2 n. 2 | <input type="checkbox"/> v. 4 n. 2 |
- R\$ 40,00 R\$ 24,00 R\$ 30,00
R\$ 24,00 R\$ 30,00

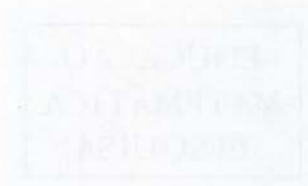
Número avulso: _____ R\$ 18,00 (cada)

Nome: _____

Endereço: _____

Cep: _____ Cidade: _____ Estado: _____

Telefone: _____ Fax: _____ Ocupação: _____



INSTITUTO DE PESQUISA E DESENVOLVIMENTO
EM CIÊNCIAS E TECNOLOGIA

Este trabalho foi desenvolvido no âmbito do projeto de pesquisa "Estudo de Fatores Determinantes da Qualidade da Produção em Empresas de Pequeno Porte", financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pelo Instituto de Pesquisa e Desenvolvimento em Ciências e Tecnologia (IPDCT). O autor agradece aos membros da comissão avaliadora e aos colaboradores que possibilitaram a realização deste trabalho.

INFORMAÇÕES PARA AQUISIÇÃO

Antes de adquirir este livro, consulte o preço no Banco Brasileiro de Dados e Informações (BBDI), endereço: Rua B. C. Ligon, post. 200, 111-1, 13080-000, Campinas, SP.

0-1001 [100] 0-1001 [100]
0-1001 [100] 0-1001 [100]
0-1001 [100] 0-1001 [100]

Nome: _____
Cidade: _____
Estado: _____
País: _____
Endereço: _____
Código Postal: _____



Impressão de miolo e acabamento:

Gráfica da PUC-SP

Rua Ministro Godói, 965 – Perdizes – SP

Tel.: 3670-8366

SUMÁRIO

Editorial

L'analyse statistique implicative:
ses bases, ses développements
Régis Gras

Une étude diagnostique en vue de la formation des
enseignants en géométrie
Saddo Ag Almouloud

A implicação estatística usada como ferramenta
em um exemplo de análise de dados multidimensionais
Régis Gras e Saddo Ag Almouloud

Possíveis relações entre o processo de formação e a
prática pedagógica: proposta de procedimentos de pesquisa
Ana Lúcia Manrique e Marli E. D. A. André

O uso do CHIC na análise de registros textuais
em ambiente virtual de formação de professores
Maria Elisabeth Brisola Brito Prado

Formação de professores para inserção do computador
na escola: inter-relações entre percepções evidenciadas
pelo uso do software CHIC
Maria Elizabeth Bianconcini de Almeida