



**HAL**  
open science

## TGING : graphic information processing for teaching

P. Orús, G. Villarroya

► **To cite this version:**

P. Orús, G. Villarroya. TGING : graphic information processing for teaching. Jun 2003, Reims, France. edutice-00001366

**HAL Id: edutice-00001366**

**<https://edutice.hal.science/edutice-00001366>**

Submitted on 13 Jan 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# TGINF: GRAPHIC INFORMATION PROCESSING FOR TEACHING

P. Orús, G. Villarroya. Department of Mathematics. University Jaume I Castellón. Spain

**ABSTRACT:** We present TGINF, a software application for graphic information processing designed in Visual Basic (taking Bertin's works as a reference) to introduce data processing in teaching. We will describe the software, the rationale behind it, how it works and an example of its actually being used on a matrix of Boolean data to show the possibilities and limits of TGINF.

**RESUMÉ:** TGINF, est un logiciel de traitement de l'information, réalisé en Visual BASIC, et qui utilise comme repère les travaux de J. Bertin (1977) pour introduire, le traitement des données, dans l'enseignement.

Nous présentons: le programme TGINF, sa motivation et son fonctionnement, et un exemple d'utilisation sur un tableau des données booléennes, en montrant les possibilités et les limites de TGINF

## 1. GRAPHIC INFORMATION PROCESSING IN A SCHOOL SITUATION: DIDACTIC REASONS FOR A DECISION

In R. Gimeno's book (1980) about the use of graphics in classification in infant schools, it is shown that in all disciplines graphics (*Le graphique*) introduce a motivating factor in teaching. They aid pupil's reasoning and interpretative writing skills and bring to the fore the level of intelligence of so-called 'weak pupils'. Gimeno applies the virtues of graphic processing described by J. Bertin in the school situation. Bertin propounds processing data in a visual way so as to highlight the relationships among them aside from any subjective interpretations by readers or researchers. The meanings of a graphic construction (graphic semiology) are derived from the set of visual oppositions that appear among the very elements the graphics syntax resorts to, and from the logical relationship which is established visually between the representation of each datum and the whole.

Other authors, such as P. Orús (1992) in her PhD dissertation, apply Gimeno's works to the primary school context using classificatory analyses within the classroom as a didactic instrument with which to manage pupils' natural reasoning in the production and use of formal reasoning.

More recently, I. Pitarch (2001) takes Gimeno's and Orús's works a step further by introducing graphic information processing into secondary studies. This work also analyses the interest and feasibility of incorporating deductive methods into statistical reasoning and is a proposal that combines logical, deductive and statistical work through data processing.

Pupils are introduced to performing and interpreting multivariate analysis by using graphic methods without the need to understand the statistical knowledge that upholds it. The relations that exist between multivariate analysis and graphic analysis are shown in the following table:

| TYPOLOGICAL ANALYSIS  | PHASES OF DECISION   | GRAPHIC INTERVENTION   |
|---|--|--|
| <ul style="list-style-type: none"><li>- Reduction of data</li><li>- Exploration of data</li></ul>                             | <ul style="list-style-type: none"><li>- Definition of problem</li><li>- Construction of data table</li></ul> | <ul style="list-style-type: none"><li>- Matrix analysis of the problem (by defining the questions)</li></ul> |
| <ul style="list-style-type: none"><li>- Classification of data</li><li>- Validation of hypotheses</li></ul>                   | <ul style="list-style-type: none"><li>- Data processing</li></ul>  | <ul style="list-style-type: none"><li>- Graphic information processing (to find out the answers)</li></ul>   |
| <ul style="list-style-type: none"><li>- Forecasting based on the groups obtained</li><li>- Generation of hypotheses</li></ul> | <ul style="list-style-type: none"><li>- Interpretation</li></ul>   | <ul style="list-style-type: none"><li>- Graphics of communication</li></ul>                                  |

As classification and data analysis are not really included in school curricula as such, no software is currently available for use in schools, and commercial software requires prior knowledge of statistics. In some cases, these applications are also difficult to use in secondary studies and are completely beyond the reach of pupils at primary and infants' schools. This was what really led us to develop a software application with these characteristics: TGINF.

PITARCH, I. (2002). *Estudio sobre la viabilidad y el interés didáctico del tratamiento de la información en la ESO*. Trabajo de investigación del tercer ciclo presentado en la Universitat Jaume I y dirigido por Pilar Orús Báguena. (Ed: UJI, Castellón).

## 2. THE SOFTWARE: TGINF (GRAPHIC INFORMATION PROCESSING)

TGINF is a programme that is based on Jacques Bertin's work on graphic information processing. It was designed in Visual Basic and allows Boolean matrices (0's and 1's), which are represented graphically by a two-colour grid (black and white), to be manipulated on a computer screen. The method of working consists in taking a reference row or reference column that may be significant and trying to obtain *marks* (white or black zones). These zones are obtained by bringing similar rows or columns closer together by permutation. The ones that have the opposite characteristics (complementary values for the variables) are placed in the opposite area. The permutation process can be carried out either manually or automatically. If automatic processing is chosen, two options are available: classification using Lerman & Gras's similarity index or an automated version of Bertin's method.

Lerman's index is used to perform typological analyses (also called automatic analyses). This kind of analysis allows us to break down a population of subjects or objects, which are described by variables or characteristics, into a specific number of subgroups. In this index, Lerman propounds measuring the similarity  $s_{ij}$  between two objects  $i, j$  (without there being any relationship of dependence between them) by the probability of observing a number  $K$  of co-presences below a number  $P_{ij}$  of actually observed co-presences:  $s_{ij} = P(K < P_{ij})$ . The objective behind using Lerman's index is to obtain a similarity index that neutralises the *size effect* presented by similarity indices defined for logical matrices. This size effect arises from the fact that two objects that have a high number of attributes will be closer than two objects that have a low number of attributes. Thus, Lerman's index gives preferential treatment to the co-presence of *rare* criteria or variables, which are uncommon bearing in mind the size of the sample that was studied. In this programme, we use the adaptation of Lerman's index carried out by R. Gras and used in his CHIC software application.

Taking Poissonian approximation we can obtain similarity index  $s_{ij}$  by the expression:

$$\text{Pr ob}[card(X \cap Y) = k] = e^{-\frac{n_i n_j}{n}} \frac{\left(\frac{n_i n_j}{n}\right)^k}{k!}$$

### Files and Data

The current version of TGINF interprets any EXCEL97 data table. The first row of this table must show the position from which it possible to start reading the contents of the table, the number of rows and the number of columns (in the example: 3,B,9,16). Subjects are placed in columns and are indicated by capital letters A,B..., while variables are placed in rows.

|    | A | B | C | D | E | F  | G | H | I | J | K | L | M | N | O | P | Q | R |
|----|---|---|---|---|---|----|---|---|---|---|---|---|---|---|---|---|---|---|
| 1  |   | 3 | B |   | 9 | 16 |   |   |   |   |   |   |   |   |   |   |   |   |
| 2  |   | A | B | C | D | E  | F | G | H | I | J | K | L | M | N | O | P |   |
| 3  | 1 | 0 | 0 | 0 | 0 | 0  | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |   |
| 4  | 2 | 0 | 1 | 1 | 1 | 0  | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |   |
| 5  | 3 | 0 | 0 | 0 | 0 | 0  | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |   |
| 6  | 4 | 1 | 1 | 0 | 0 | 1  | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |   |
| 7  | 5 | 0 | 1 | 1 | 1 | 0  | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |   |
| 8  | 6 | 1 | 1 | 0 | 0 | 1  | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |   |
| 9  | 7 | 0 | 0 | 0 | 0 | 0  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |   |
| 10 | 8 | 0 | 0 | 1 | 0 | 0  | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |   |
| 11 | 9 | 0 | 1 | 1 | 1 | 0  | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |   |
| 12 |   |   |   |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |

Table 1: Data about town planning (Bertin, 1977 p.59)

In the example, we can read the data starting from the third row, column B, and we work with a matrix made up of 9 rows and 16 columns. The translation of this matrix into a black and white grid is as follows:



Table 1G

### 3. AN EXAMPLE OF TGINF IN USE

We show the application of various methods of automatic data processing for the data represented in Table 1.

We applied Bertin's and Gras's methods to classify the initial Boolean matrix by rows. The following results were obtained:

#### CLASSIFICATION OF MATRIX 1 BY ROWS

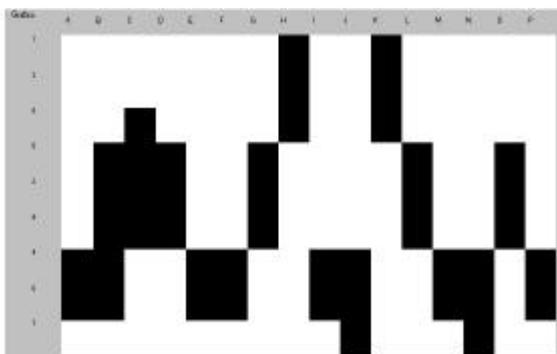


Table 1R BERTIN

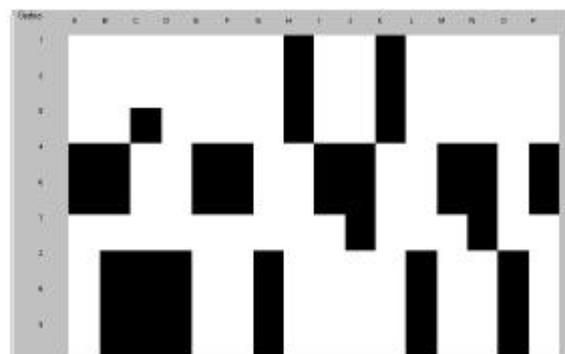


Table 1R GRAS

## CLASSIFICATION OF MATRIX 1 BY ROWS AND COLUMNS

When we apply the two methods described above, it can be seen graphically that the groups obtained are practically identical, with the same variables characterising the same group of subjects. The difference is due to the distinct location of these groups on the plane.

If the classification by columns is applied to the matrices obtained, the result is:

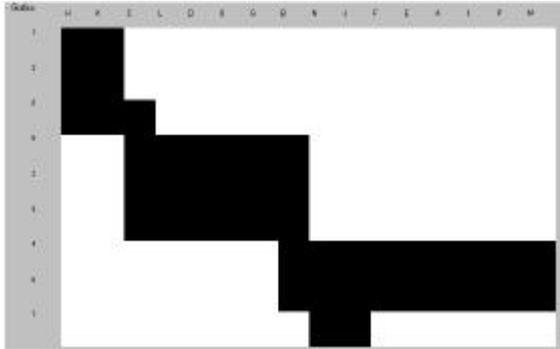


Table 1R+C BERTIN

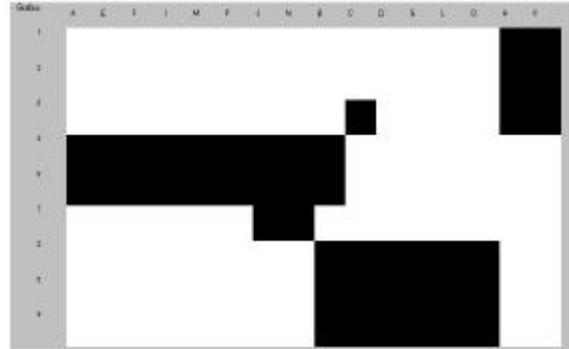


Table 1R+C GRAS

By analysing the results obtained, it can be observed that the resulting classifications are the same when either method is applied. The variables make up classes {Var1, Var3, Var8}, {Var4, Var6, Var7}, {Var2, Var5, Var9} and the subjects make up classes {H, K}, {J, N}, {A,E,F,I,M,P}, where B and C are subjects of uncertain classification in the classifications that were established.

## 4. CONCLUSIONS: POSSIBILITIES AND LIMITATIONS OF TGINF

There are a number of different didactic applications for our programme:

- It updates Gimeno's method by using the new technologies to perform graphic classification.
- The programme is easy to use. In the school context it allows manual image processing, that is, data can be handled directly. Permutation of previously selected rows or columns can be carried out by clicking with the right button of the mouse and an undo feature allows the last actions to be cancelled.
- It offers reversibility, that is, we can go back to the initial situation (when the file was loaded) at any time by clicking on the icon:



- It is possible to save any matrix obtained during the process of permuting rows and columns.
- It offers the teacher solutions to the problem being dealt with by allowing manual processing and automatic processing to be compared.
- It allows two kinds of automatic processing to be used:
  - BERTIN'S METHOD, which involves grouping marks around a row or column
  - LERMAN & GRAS'S SIMILARITY INDEX, which was adapted and used by Gras in the CHIC programme.

- It can be used to process up to 730 columns and as many rows as EXCEL can handle, but when we have more than 20 rows and 30 columns, the legends with the name of the variables and their characteristics cannot be read on screen. We are currently seeking to resolve this problem.
- The programme enables us to perform all the actions described above but it has not yet been tried in a school situation. We consider that it is essential for teachers and pupils to see TGINF running and how it works in order to demonstrate its didactic capabilities as well as its possible limitations. Our work will continue in this direction and will enable us to introduce the improvements needed for it to work at an optimal level.

## REFERENCES

- BERTIN J. (1977) *La Graphique et le traitement graphique de l'information*, Paris:FLAMMARION
- COUTURIER R. (2000) *Traitement de l'Analyse Statistique dans CHIC*. Caen: *Actes des Journées: LA FOUILLE DANS LES DONNEES PAR LA METHODE D'ANALYSE STATISTIQUE IMPLICATIVE Applications et traitement par CHIC*
- GIMENO R. (1980) *Apprendre à l'école par la graphique*, Paris :RETZ
- GRAS R. (1995) *Traitement "A la main" de tableaux de donees par des methodes classificatoires*. Atelier animé per R. Gras. Caen: *Actes du colloque METHODES D'ANALYSES STATISTICS MULTIDIMENSIONNELLES EN DIDACTIQUE DES MATHEMATIQUES*.
- ORUS P. (1992) *Le raisonnement des élèves dans la relation didactique; effets d'une initiation à l'analyse classificatoire dans la scolarité obligatoire*. Thèse présentée à l'Université de Bordeaux-I. (Ed: IREM de Bordeaux).
- PITARCH, I. (2002). *Estudio sobre la viabilidad y el interés didáctico del tratamiento de la información en la ESO*. Trabajo de investigación del tercer ciclo presentado en la Universitat Jaume I y dirigido por Pilar Orús Bágüena. (Ed: UJI, Castellón).