Design and Validation of a Model based Diagnosis System

Stéphanie Jean, Elisabeth Delozanne, Pierre Jacoboni, Brigitte Grugeon

Design and Validation of a Model based Diagnosis System

Stéphanie JEAN*, **Élisabeth DELOZANNE*, Pierre JACOBONI***
Brigitte GRUGEON**

** LIUM - Université du Maine*
72 085 LE MANS Cedex 9 (France)
Tel.: 33 2.43.83.38.58
Email: Stephanie.Jean@lium.univ-lemans.fr
http://www-ic2.univ-lemans.fr/~jean/

***DIDIREM - Paris VII*
2 place Jussieu
75 251 PARIS Cedex 05 (France)

Abstract

PÉPITE is a multidisciplinary project involving researchers in computer science and in didactics for mathematics. It consists in building a computerised environment able to model the competence of 15-year-old students of French secondary schools in elementary algebra. This work is based on a didactical analysis whose result is a multidimensional model of competence in algebra and a validated pencil and paper diagnosis tool. The PÉPITE project aims to automate this tool.

In this paper we present the prototype developed to automatically build students' profiles and the results that we have already obtained. We put emphasis on the methodology adopted to design and validate this tool and the model of student's competence built by PÉPITE.

Keywords

Assessment, student modelling, interface design, multidisciplinary approach, validation criteria, elementary algebra.

1 Introduction

The aim of the PÉPITE project is to develop a computerised environment able to model 15-year-old students' competence in elementary algebra. We propose a methodology based first on a close collaboration with educational researchers from the very beginning of the project, second on a user-centred design approach and third on a definition of validation criteria from the early stage of the design.

We first present the multidimensional model of algebraic competence and its pencil and paper tool, which are the bases of the project. They were developed by educational researchers in mathematics [GRUGEON, 1997]. We then describe the different components of PÉPITE we have designed to build the students' profiles. For each step we present our methodology of validation and the results we have obtained. We lay stress on the importance of the interface to obtain reliable observations as well as to propose profiles easily usable by teachers.

The basis of the PÉPITE project is an educational research starting from the hypothesis of knowledge building: students have built their own knowledge, which may be different from reference knowledge [SFARD, 1991]. Consequently, students' productions present coherence, which correspond to their knowledge. From our point of view, this study had two main results.

First it proposes a model of competence in algebra that includes quantitative and qualitative descriptions. This model considers four dimensions (use of letters, algebraic calculation, conversion from one context to another and type of explanation).

Second a tool has been built to analyse students' productions and to give a description of student's functioning in algebra. The aim is to find points to rely on to modify students' knowledge. The diagnosis tool combines a set of pencil and paper exercises with a diagnosis matrix linking questions and dimensions of analysis. This very detailed description of the student's behaviour requires a higher level description: the cognitive profiles, which are built by transversal analysis of the diagnosis matrix. See [JEAN & AL, 1999] for a summary of this educational research and [GRUGEON, 1997] for more details.

These profiles have three levels of description: a quantitative description of algebraic competence in terms of *success rates*, a qualitative description of *functioning coherence* per modalities and a description of *flexibility* between the considered frames (algebraic, numerical, graphical, geometrical and natural language) represented by a diagram.

Manually building the profiles is a difficult task, even with experienced teachers. It requires much time and personal involvement to understand how the diagnosis works and also much time to analyse each student's production. In order to make this pencil and paper tool accessible to teachers we decided to automate it. The automation of the diagnosis is the aim of the PÉPITE project.

The computerised version of the diagnosis consists of three parts corresponding to the three steps of the manual building of profiles with the pencil and paper tool:

– The student interface proposes exercises and gathers student's answers.

– The analysis module "interprets" the students' answers and fills the diagnosis matrix.

– The teacher interface establishes the profiles from the filled matrix by transversal analysis and presents them to the users (users or researchers).

These different parts are detailed in the following sections.

2 The test interface

PÉPITEST proposes an adaptation of the pencil and paper tasks to the computer and collects students' answers to 22 different exercises (figure 1 shows an example of exercise in PÉPITEST).

This first module of our project, data collector for the diagnosis is an integral part of the diagnosis. Indeed, the design of the student interface plays a significant role in the quality of the diagnosis. Taking into account Human-Computer Interface issues ensures better results in the analysis by allowing students to behave as they usually do and therefore by furnishing more reliable answers (data for the diagnosis). Quality of data is decisive in the assessment process. That's the reason why we took much care in building the student interface.

Design

The main difficulties lie in the adaptation of the pencil and paper exercises and also of usual students' tools to the computer platform. Indeed, doing mathematics in a pencil and paper environment or with PÉPITEST is not the same. Annotating figures or writing algebraic expressions and text is much more difficult in a computerised environment.

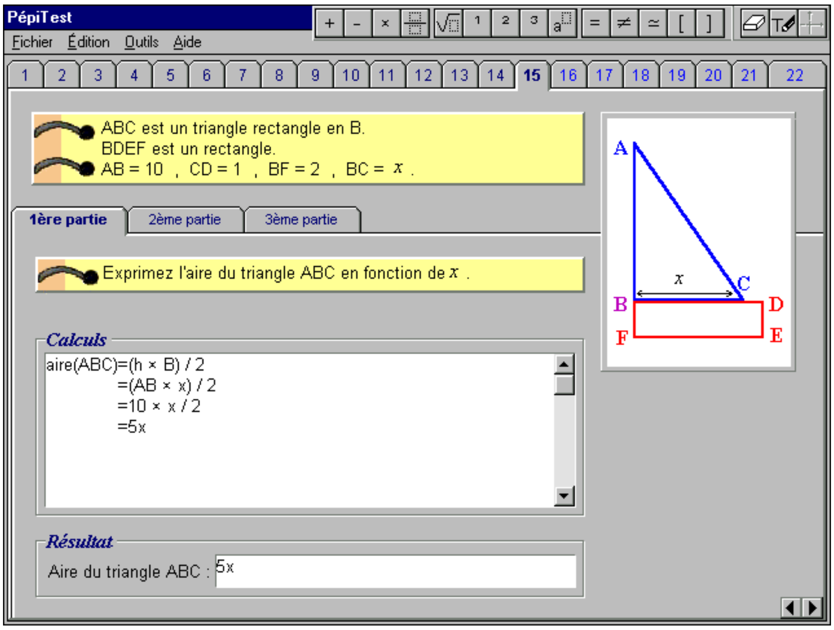
Exercises

Each exercise gives special design difficulties linked to the nature of the cognitive activities in play. So we paid much attention to ergonomics.

Algebraic expressions

The problem of producing algebraic expressions is particularly important. There is inevitably a great difference between the way students usually write expressions and the way they have to “write” them on a computer with only a keyboard and a mouse. Figures 1 and 2 show an example of student’s pencil and paper production: $\mathrm{A(ABC)} = \dfrac{B \times h}{2}$ and a PÉPITEST’s one: $\mathrm{aire(ABC)} = (h \times B)/2$; the writing of the fraction differs [ARTIGUE, 1995]. This classical problem of linear and spatial writing is one of the problems linked to the producing of algebraic expressions.

To limit students' difficulties, we are developing an algebraic expressions editor which should enable students to write expressions more naturally and to represent them as they usually do. With this editor, manipulation of expressions should be closer to students’ usual practice: either in the production of expressions or in their representation.



Statement:
ABC is a B right-angled triangle.
BDEF is a rectangle.
AB=10, CD=1, BF=2, BC=x.

Question of the 1st part:
Express the area the ABC triangle according to x.

Figure 1: An example of exercise in PÉPITEST: exercise 15 and translation of its statement

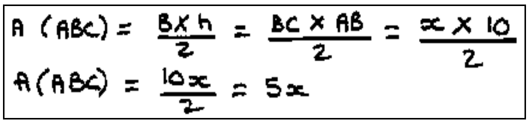


Figure 2: An example of student’s pencil and paper production for exercise 15.

Proposed tools

PÉPITEST proposes several tools to replace the students' usual pencil and paper tools (pen, rubber, roughbook, annotation tool, graphic representation tool, etc.). These tools have been carefully designed because their use has to be as close as possible to the use of the pencil and paper ones, keeping in mind that we can’t propose an exactly equivalent environment.

Validation criteria

The validation of PÉPITEST will be ensured:

– By the equivalence between the data: the pencil and paper productions and PÉPITEST’s ones, both at a micro-level and at a macro-level (① on figure 3).

– By the equivalence between the results: the manually built profiles from pencil and paper students’ productions and the manually built profiles from students’ productions with PÉPITEST (② on figure 3).

We explain how we measure those equivalencies in the following section.

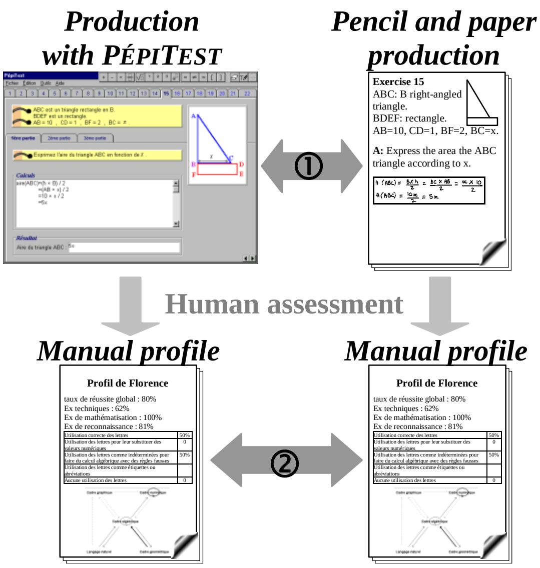


Figure 3: PÉPITEST validation criteria.

Experimentation and results

PÉPITEST is completed and has now been tested several times with a total of 75 students from 3 different classes.

For this module we mainly obtained the following results:

At the macro-level, first, students had difficulties in producing algebraic expressions as we expected. But, those difficulties did not prevent them from writing such expressions. Therefore it shows that the expressions editor will be welcomed and useful but may be bypassed. Second, for each question of the test, we have found every kind of expected answer of the didactical analysis. This is an important result: PÉPITEST does not reduce the variety of students' productions for each analysis component.

At the micro-level, for the same student there can be some differences between his productions with PÉPITEST and his usual pencil and paper productions. PÉPITEST disturbs some students and let them make errors they usually don’t do. In most cases, those errors correspond to weak knowledge needing to be reinforced, PÉPITEST can then be considered as a microscope, highlighting problems. However teachers detected a significant difference between a student’s production with PÉPITEST and their conceptions about her knowledge once. But teachers consider that productions are mostly equivalent.

Thus those two levels ensure our first validation criterion.

Educational researchers could fill the diagnosis matrix from students' answers to PÉPITEST problems and then manually build the profiles. The teacher of the class could thus confirm the obtained profiles. So it corresponds to our second validation criterion.

This clarifies the way of measuring the validity of our behavioural model according to [BALACHEFF, 1994].

3 The diagnosis

PÉPIDIAG is the diagnosis module of PÉPITE. The diagnosis is not made in real time but in batch mode. There are two reasons for this:

– Only the final students’ answers are analysed either by PÉPIDIAG or by a teacher,

– The profiles are built for teachers. They are neither used in a learning environment nor directly by students. Those are the reasons why we don’t have to make real-time profiles.

PÉPIDIAG interprets and codes students' productions, from the data (students' productions) furnished by PÉPITEST in a text file. It fills the diagnosis matrix (55 questions × 35 possible codes) and defines values given to the functioning modalities tested by the exercise.
The difficulty to establish this diagnosis comes from the variety of exercises and from the variety of answers given by students. In the diagnosis, for each question there corresponds a set of class of answers usually given by students (this is the result of the didactical research). To each of these answers corresponds a code in the diagnosis matrix. The system has to interpret students' answers to associate them with the right code. This association is easy to do with closed questions but is much more difficult with totally open questions where the students mix natural language and mathematics.

Validation criterion

PÉPIDIAG will be validated if for a same student's production a manually filled matrix is totally equivalent to an automatically filled one (① on figure 4).

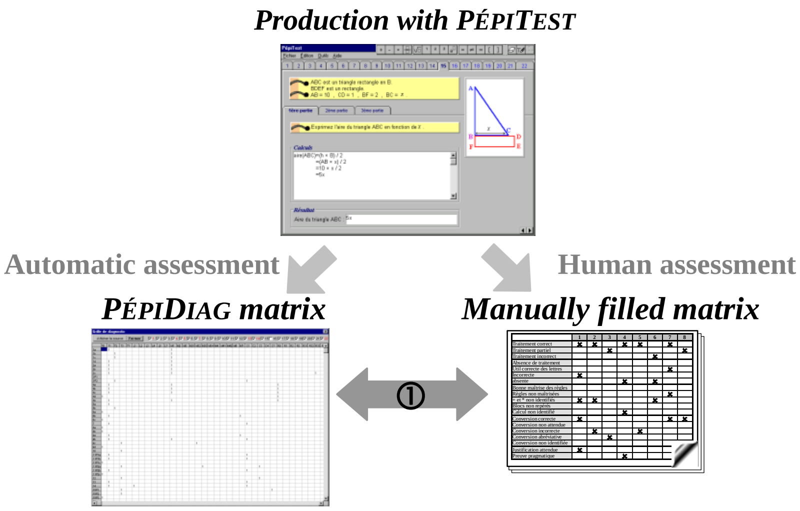


Figure 4: PÉPIDIAG validation criterion.

Results

A first prototype of PÉPIDIAG already exists. It analyses automatically closed answers and open answers with simple algebraic expression (by simple we mean made of only one expression). The first results are as follows:
This prototype analyses about 75 percent of students' answers to PÉPITEST questions. That means that we can already partially automate the diagnosis. We ran PÉPIDIAG on all 75 student's answers gathered with PÉPITEST and the system filled-in the diagnosis matrices. In order to correlate this partial diagnosis with human assessment, we choose five students with different levels of competence and we asked an expert to manually fill the diagnosis matrices. For almost every student the matrix by PépiDiag matched the manual one produced by the human assessor, even if locally there were still minor differences, the matrices lead to similar profiles. We have previously found the same results concerning human assessment. This explains the way of measuring the validity of our epistemic model according to [BALACHEFF, 1994].

4 The profiles

PÉPIPROFIL, the teacher interface has two tasks. First it establishes the students' profiles from the filled matrix by a transversal analysis which corresponds to an algorithm. Second, it presents the profiles to the users (teachers or researchers). In this section we present this second task more precisely. We describe and explain the different levels of

description of the profiles (quantitative description of algebraic competence, qualitative description of functioning coherence and description of flexibility between contexts) and show the way we present them to make them easier to understand by teachers.
The diagnosis has to be adapted to teachers to be accepted by them. It has to be easy to use and to understand. The diagnosis should be adapted to teachers' needs and habits and not the opposite.
We used two different ways to make the diagnosis acceptable to teachers. First of all, we use a multidisciplinary and user-centred approach. Teachers were present at each step of the project. PÉPITE results from an educational research based both on a theoretical work and on an extensive empirical work. That is the reason why PÉPITE is close to teachers' needs from the beginning. Second PÉPIPROFIL gives control to teachers over the tool and over the results: they can verify, complete and modify the analysis proposed by PÉPIDIAG. This aspect is crucial for having teachers accept a form of assessment done by the system.

The different parts of the profiles

The profiles corresponds to the pencil and paper ones, they gather much information on students' functioning in algebra. In order to let teachers take advantage of the richness of the information given by the profile, PÉPIPROFIL gives either a quick insight of the students' functioning in algebra or all the available information.
So we paid much attention to the way we present the profiles in order to make them easy to understand and use.
In the following section, we present the different parts of the profiles. Those can be adapted by teachers to their use in different ways in PÉPIPROFIL:
– By modifying numerous parameters (for instance in the diagram of flexibility between contexts, the percentage defining the limit between weak and good flexibility can be changed).
– By explaining the results by showing how the profile has been made, showing PÉPIDIAG analysis in the context of the exercise.
– By giving teachers control on the results: teachers have the possibility to modify PÉPIDIAG's analysis for certain exercises or modalities and to complete the analysis for answers that PÉPIDIAG didn't succeed in analysing.

Quantitative description of algebraic competence

The first part of the profile aims to give a quick overview of student's algebraic competence by a quantitative description with regard to expected competence.
This part consists in presenting:
– The global success rate for treated exercises.
– The percentage of unanswered questions, correct treatments, unexpected or partial treatments and incorrect treatments represented by a bar chart.
– For each kind of exercises (technical, recognition and modelling exercises), a definition of the type, the student's success rate and each treatment (s)he masters.

Qualitative description of functioning coherence by modalities

The second part of the profile presents a more qualitative description of student's functioning in algebra. In the four functioning modalities studied in the test (using of letters, algebra, conversion from a context to another and type of explanation) several components represent the different possible students' functioning.
For each modality this part shows the components corresponding to the student's answer, the proportion represented by the component in the modality (by a percentage

and a coloured band of corresponding width) and the number of questions where the component was concerned in the test.

This representation clearly shows teachers the most usual student's functioning by identifying the most important components in each modality.

Description of flexibility between contexts

In this third part of the profile, a diagram represents the flexibility between the frames considered in the test (algebraic, numerical, graphical, geometrical and natural language). Dotted lines represent the links considered in the test but not mastered at all by the student. Normal lines show weak links and the bold ones represent a good flexibility of the student in the related frames. The same coding is applied on the circles corresponding to the work in the contexts considered in the test (of course algebra, which is the subject of the test, but the numerical context is also represented in some exercises). This graphical representation is very efficient: teachers immediately see the essential points of the students' functioning.

Summed up profile

In addition to these tree parts, PépiProfil proposes an additional part with a text summing up the most important points of the profile. This can be considered to be a help given to teachers to interpret the profiles. The aim is also to propose several representations for the same information: teachers can choose the representation they prefer (or use several).

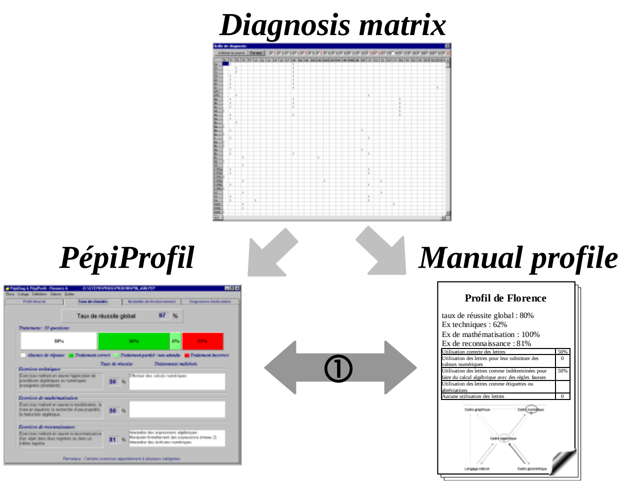


Figure 5: PépiProfil validation criterion.

Validation criterion

The building of profiles part of PépiProfil will be validated if for a same matrix a manually built profile and PépiProfil's one are totally equivalent (① on figure 5).

PépiProfil's validation concerning the presentation of the profiles to teachers will consist of Human-Computer Interface's assessment (with ergonomic validation, tests with experts and larger-scale test) [Bastien & Scapin, 1993], [Shneiderman, 1992] already implemented for PépiTest [Jean & al, 1998].

Results

PépiProfil is now completed; we obtained two main results:

With a manually filled matrix, PépiProfil computes the same profile as a teacher does.

From the partial 75 matrices yet filled by the system, corresponding to the 75 students of our corpus, PépiProfil builds partial profiles that have been confirmed by teachers.

Concerning the presentation part, tests with teachers are planned but not done yet.

5 Conclusion

The three modules of PÉPITE are now completed to a large extent and our validation criteria are clearly defined. From the beginning, we integrated teachers and educational researchers in our design team. This multidisciplinary and user-centred approach allows us to propose a test adapted to students and to build profiles adapted to teachers' needs. Results are already numerous even with the partial diagnosis.

PÉPITE already supplies reliable data and builds partial but usable profiles.

Future work includes: First, the algebraic expressions editor has to be completed and integrated into PÉPITEST. This improvement would certainly increase the reliability of data. Second, PÉPIDIAG has to be completed in order to analyse all the answers and not only 75%. Meanwhile, we keep in mind that difficulties linked to interpreting answers to open questions would probably prevent us from analysing 100% of answers. Furthermore, PÉPIPROFIL has to be widely tested with teachers to validate its interface and usefulness.

PÉPITE gives us the opportunity to build and use cognitive profiles on a large scale.

The profiles can be used:

– By the teacher or the teacher team to form groups relying more on students' knowledge than only on marks and to adapt their teaching. The teacher might be helped in that by a software able to build a profile of the class, to form groups from similar profiles and to propose adapted working subjects.

– In a Computer Supported Learning Environment proposing activities, advises and explanations adapted to the student's knowledge, but also in distance learning activities and in larger-scale assessments. In these uses, profiles must be perfectly reliable because there is no teacher to control them.

The methodology we have adopted to design and validate our prototype seems to be re-usable in others CSLE: early and close collaboration with educational researchers, user-centred design and early definition of validation criteria. Those criteria specify how to measure equivalencies between the "behavioural models" (observed data) and between the "epistemological models" (student's profiles) built by the system and by human assessors [BALACHEFF, 1994].

6 References

[ARTIGUE, 1995] Michèle ARTIGUE, Une approche didactique de l'intégration des EIAO à l'enseignement, in D. GUIN, J.-F. NICAUD et D. PY, EIAO, Tome 2, pp. 17-28, Eyrolles, Paris, 1995.

[BALACHEFF, 1994] Nicolas BALACHEFF, Advanced Educational Technology: Knowledge Revisited, Journal of Educational Technology Systems, 23 (2), pp 98-106, 1994.

[BASTIEN & SCAPIN, 1993] Christian BASTIEN and Dominique SCAPIN, *Ergonomic Criteria for the Evaluation of Human-Computer Interfaces*, RT n°156, INRIA, June 1993.

[GRUGEON, 1997] Brigitte GRUGEON, *Design and Using of a Multidimensional Structure of analysis in elementary algebra*, Recherches en Didactique des Mathématiques, vol 17 (2), pp 167-210, 1997. (In French)

[JEAN & al, 1998] Stéphanie JEAN, Élisabeth DELOZANNE, Pierre JACOBONI and Brigitte GRUGEON, *Cognitive profile in elementary algebra: the PÉPITE test interface*, IFIP TC-3 Official Journal special issue « Education and Information Technology », no 3, Chapman & Hall Ltd, pp 1-15, 1998.

[JEAN & al, 1999] Stéphanie JEAN, Élisabeth DELOZANNE, Pierre JACOBONI and Brigitte GRUGEON, *A Diagnosis Based on a Qualitative Model of Competence in Elementary Algebra*, AI-ED 99, France, 1999 (to be published).

[SFARD, 1991] A. SFARD, On the Dual Nature of Mathematics Conceptions: reflections on processes and objects as different side of the same coin, Educational Studies in Mathematics, Vol. 22, pp 1-36, 1991.

[SHNEIDERMAN, 1992] Ben SHNEIDERMAN, Designing the user Interface: Strategies for Effective Human-Computer Interaction, Addison-Wesley Publishers, 1997 (3rd edition).