Supporting teachers when diagnosing their students in algebra

Elisabeth Delozanne, Brigitte Grugeon, Dominique Previt, Pierre Jacoboni

Supporting teachers when diagnosing their students in algebra

Elisabeth DELOZANNE [1,2], Brigitte GRUGEON [3,4], Dominique PREVIT [1,5] , Pierre JACOBONI [1]
[1]*LIUM, Avenue Olivier Maessian, 72085 LE MANS Cedex 09, France*
[2]*IUFM de Créteil, Rue Jean Macé, 94861 BONNEUIL Cedex*
[3]*DIDIREM - Paris VII, 2, Place Jussieu, 75 251 PARIS Cedex05, France*
[4]*IUFM d'Amiens 49, boulevard de Châteaudun 80044 AMIENS CEDEX, France*
[5]*IUFM de Bretagne 153, rue de Saint-Malo CS 54310 35043 RENNES CEDEX, France*

Abstract. In this paper we will describe our current work on the Lingot project. We will focus on the description of a software called « Pépite », the objective of which is to help teachers diagnose their students' algebra competencies. Then we will show how different uses of Pepite by teachers lead us to define several types of diagnosis in order to fit better various users' needs.

1. Introduction

Teachers are asked to be more efficient and to develop innovative learning scenarios in order to tailor instruction to students, to manage heterogeneous classes, to personalise learning programs, to motivate students that have given up school. In particular in mathematics, these new tasks require a fine analysis of the students' work in order to understand the coherence of the personal conceptions they have built. In our opinion this is a condition to have a better regulation of learning by organising tailored activities in classrooms in order to destabilise wrong or unsuitable conceptions and to strengthen or to develop right ones. The Lingot project is based on three main pedagogical assumptions:

- Students' answers show coherences in their reasoning.
- It is possible to analyse their answers to identify the coherences the students have built (correct, partial, inappropriate).
- Identifying those coherences would help teachers give appropriate tasks to destabilize them (if inappropriate) and to make them evolve.

The key point of our approach is that students' answers to mathematical problems are not analysed as errors or as lack of skills but as indicators of incomplete, naive and often inaccurate conceptions that students have built and that interfere with and sometimes prevent from rather than support learning. "Because they are strongly rooted in personal experience and could not be tested using available tools, faulty conceptions endured" [12]. Detecting these conceptions is a very difficult task. Teachers are not trained to that kind of diagnosis. Very often they interpret students' answers as mistakes and their strategy to remediate consists in giving the right rules and repeating the same type of exercises. Sometimes these can reinforce inaccurate conceptions rather than correct them. We assume that ITS can be very helpful assistants for teachers in such news tasks with the condition they are designed of course with great attention paid to software modelling and to students' cognitive development, but, also, to teachers' activity and teachers' personal development [4, 15, 1].

We chose algebra to test our hypothesis by designing and implementing a prototype.

The first reason is that algebra is a key domain to enter mathematics and scientific studies but learning algebra rises such difficulties for students that “for many, algebra acts more like a wall than a gateway, presenting an obstacle that they find too difficult to cross”[5]. The second reason is that Mathematical Educational Research has made a great deal of work studying how to improve the learning and teaching of algebra [5, 10, 3]. Thus it is possible to ground the ITS design on a careful a priori engineering of the teaching-learning process in school algebra in order to build models with empirical and theoretical foundations. Reciprocally, implementing prototypes lead to stronger modelling and to validation, discussion, testing, supplementing, and systematizing or inflecting the educational analysis.

The Lingot project is based on three research assumptions:

- Educational research provides foundations for ITS design and reciprocally ITS help to operationalise educational research and can help to disseminate research results in schools.
- Prototyping allows experimentation in ecological situations and then help to generate new research hypothesis.
- The design process of the ITS is crucial for the acceptance and integration of ITS in schools practice.

“For the teacher’s standpoint classroom technology often is itself the problem not the solution”[4]. In this situation it is necessary to pay particular attention to usability problems but actually designing a good artefact is not sufficient to make a “tool” or an “instrument” for users, this especially if the software leads to radically new ways of doing familiar things. According to Instrumented Mediated Activity framework developed by Rabardel [13], an artefact offers a field of possible activities but also a field of constraints on human activity. One given user has to make his own instrument with an artefact by adapting his activity to the constraints and possibilities of the artefact. This theoretical framework seems to be fruitful to study ITS design, especially if the ITS will lead to modify teachers approach to learning, teaching and class management. In that case it seems necessary to have a participative design approach in order to implement the software with the teachers. We assume that teachers’ role is very important in designing ITS not only as a specialist of the context of usage and classroom teaching [11] but as an innovator in the sense that teachers will invent new ways of using the software in order to tailor it to their classroom practice: the design process is considered as “an activity distributed between designers and users”[13]. In the Lingot project we focus on a design process to bridge the gap between classroom expertise and software development expertise. The design process is thus based on the following assumptions:

- Designing an ITS to support teachers’ activity must coordinate participatory design approach with detailed observation of classroom practises.

This paper studies our current work on the Lingot project and illustrate our approach more specifically on the description of a software called « Pépite », the objective of which is to help teachers diagnose their students’ algebra competencies. Then we will show how different uses of Pépite by teachers lead us to define several types of diagnosis in order to fit better different users’ needs in different contexts.

2. The Lingot project

The Lingot project is a multidisciplinary project involving computer scientists, educational researchers in mathematics, cognitive psychologists, mathematic teachers trainers and mathematics teachers. The objective of the Lingot project is to design situations including ITS for learning algebra in secondary schools. On the one hand, the aim is to allow teachers to take into account their students' cognitive diversity in order to manage the classroom and regulate individual learning. On the other hand the aim is to provide researchers with tools for studying systematically and on the long term, the effects of teaching strategies on learning.

The basic idea is to start from a multidimensional model of algebraic competencies in secondary schools to study the instruction given in schools. The individual competencies (accurate or not) a student has built are situated in relation with competencies the institution is expecting. From these analyses the research team build learning situations in order to introduce new concepts or to make students' conceptions evolve to the expected skills. The key point is to allow the teacher to manage different learning processes which is a crucial issue, especially for students with big difficulties. The metaphor is to look for the nuggets (in French "les pépites") from which we can build ingots (in French "les lingots").

Three main issues are then considered in the Lingot project: diagnosing students' competencies, designing learning situations adapted to the students' cognitive profiles, designing instruments to support teacher activity. In this paper we focus on the first issue.

3. The Pépite project

This is the diagnosis part of the Lingot project. The basic idea is that students' answers to a set of well chosen problems show coherences in their reasoning and computing. Understanding these coherences will help teachers to tune the learning situations better.

3.1 Didactical foundations

The foundation of the project is a research work in mathematical education [3,8]. From a review of theoretical and empirical work published and from a long term study of student algebraic activity, Grugeon built a *multidimensional model of algebraic competencies* expected from students in French secondary schools (15-16 years old). The main dimensions considered are: mastered skills, meaning of letters, processing algebraic expressions, translation between different representations, types of justifications. In order to situate students in comparison with this model, she proposed a paper and pencil diagnosis tool. Students' algebraic competencies and difficulties are analysed with three entries: type of problems (algebra as a tool to solve arithmetical problems, generalisation problems, proving problems, modelling problems), the objects of algebra and formal computing. She proposed *a test* : a set of about twenty exercises with closed questions and open questions. Then *an analysis grid* derived from the model of competencies is used by a human (a teacher or more probably a researcher) to code the students' answers to each exercises of the test. Then carrying out a global analysis of the coding results, the teacher (or researcher) builds *a cognitive profile* of the student in algebra. Figure 3, 4 and 5 show the Lauren's cognitive profile built by a teacher with the software that supports this diagnosis method. It is a three levels description of the student's competencies:

- A quantitative description expressed by success rates on mastered skills,
- A qualitative description expressed by characteristics of students in giving the meaning of letters, processing algebraic expressions, translating between representations, type of rationality,
- A description of flexibility between representations expressed by a diagram indicating the links between representation modes the student mastered.

This paper and pencil diagnosis tool is too complex to be used by teachers in every day classrooms. So the first research work of the Pépite project was to explore the idea of the automation of the paper and pencil tool. This work [9] intended to demonstrate:

- the possibility to collect with a computer data on students' competencies from which experts could build the students' profiles,
- the possibility to automate this diagnosis (at least partly),
- the possibility for teachers to make decisions in their classrooms from these cognitive profiles.

3.2 The first Pépite software

Pépite is a free software available at [16]. The first version was implemented in Delphi by S. Jean. Then a second version improving the automatic diagnosis and including users' suggestions for improving usability and effectiveness has been implemented in Java. Following the diagnostic strategy proposed by Grugeon, Pépite software is made of three modules.

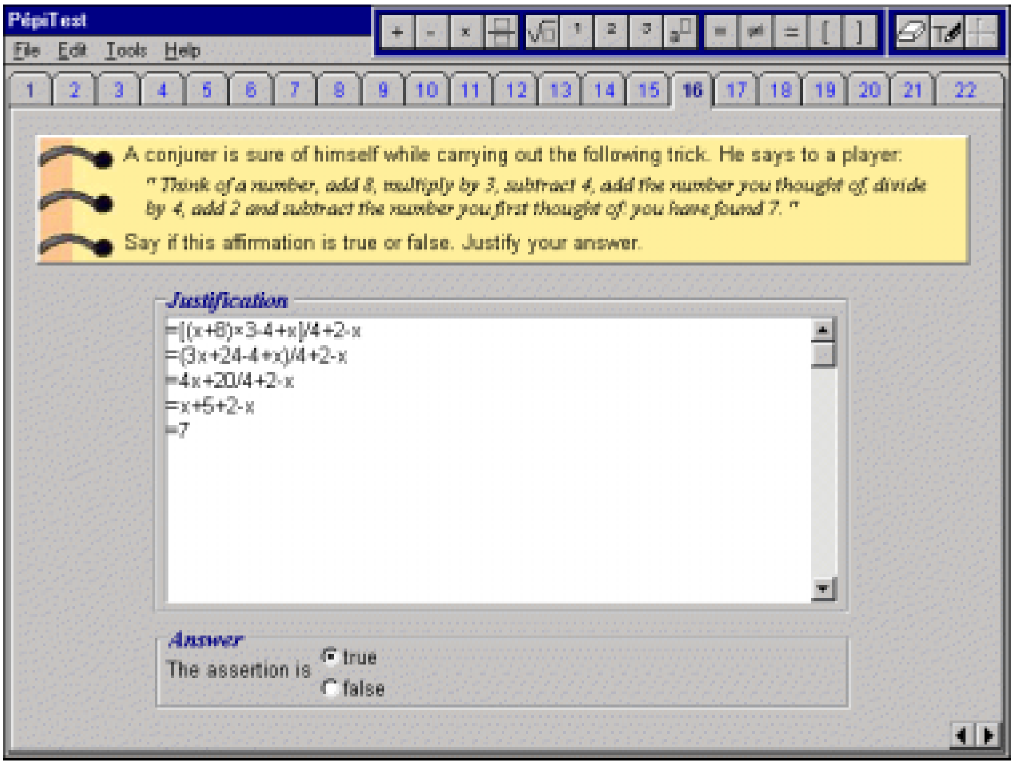

Figure 1 : Lauren's answer to an exercise in PepiTest

PepiTest is the students' software. It proposes 22 exercises derived from the paper and pencil tasks and it gathers students' answers to problems. It contains closed questions with Multiple Choices answers or more interactive answering techniques (for instance matching of clickable parts of a graphic), but with a limited number of possible answers). It contains also open questions requiring the student to produce algebraic expressions or natural language answers or mixed answers (we call "mathural language"). For educational researchers it is important for the diagnostic that the students can provide explanations by themselves. A great attention has been paid to HCI design issues because it is crucial for the diagnostic that the data collected be indicators of students' competencies and not indicators of interface manipulation problems. In particular, the mathematical educational researchers in the project team were very suspicious at the beginning about the modifications in the mathematical tasks due to the difficulties to enter algebraic expressions with a keyboard and a mouse. Figure 1 shows a student's answer to an exercise.

PepiDiag is the analysis module; it "interprets" students' answers to every exercise of

PepiTest. Like in the paper and pencil tool, it matches every student's answer with an item of diagnostic. We call that operation coding student's answers. PepiDiag automatically fills a "diagnostic matrix" of 55 questions and 35 items derived from the multidimensional model of algebraic competencies. In fact it partially fills the matrix. The closed answers and algebraic expressions are analysed. Answers in natural language and mixed answers are very partially analysed by key words analysis. So 75 % of the students' answers are automatically analysed.

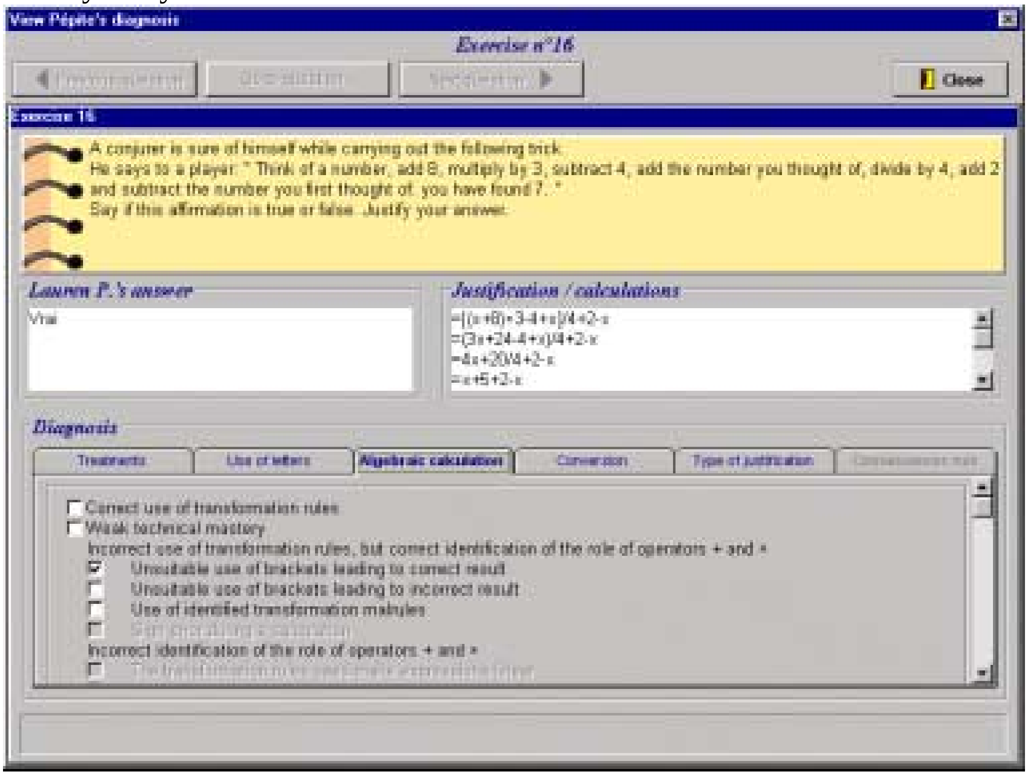

Figure 2 : PepiProf teachers' coding interface with the six dimensions of the diagnosis

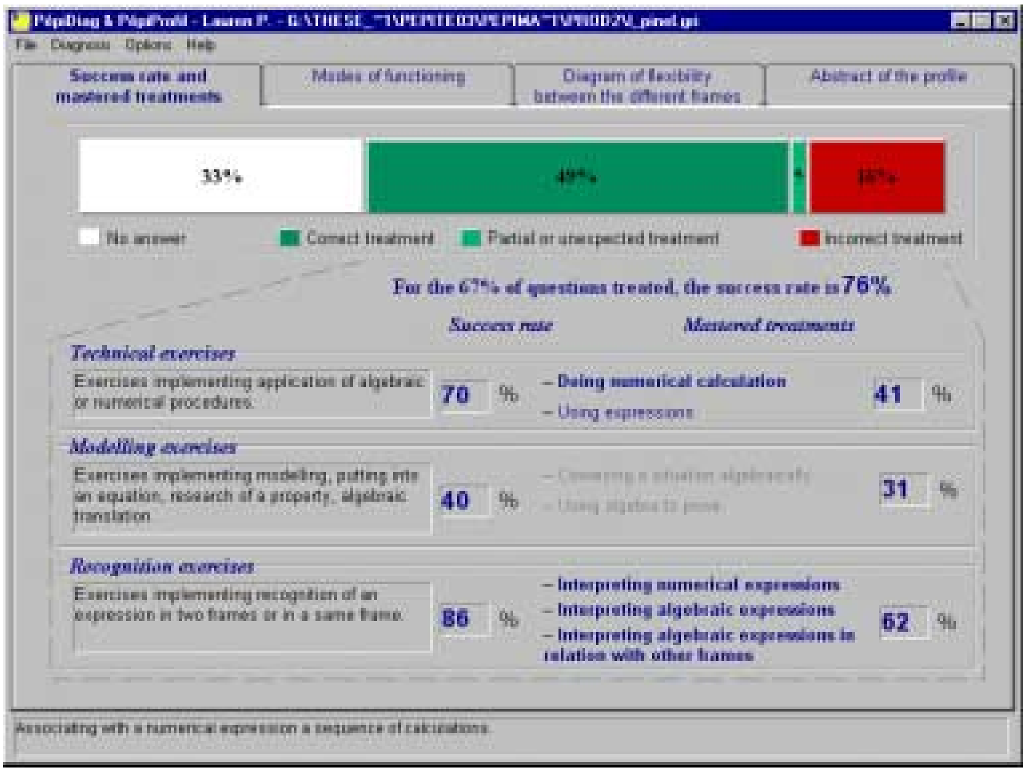

Figure 3 : Lauren's cognitive profile, quantitative description

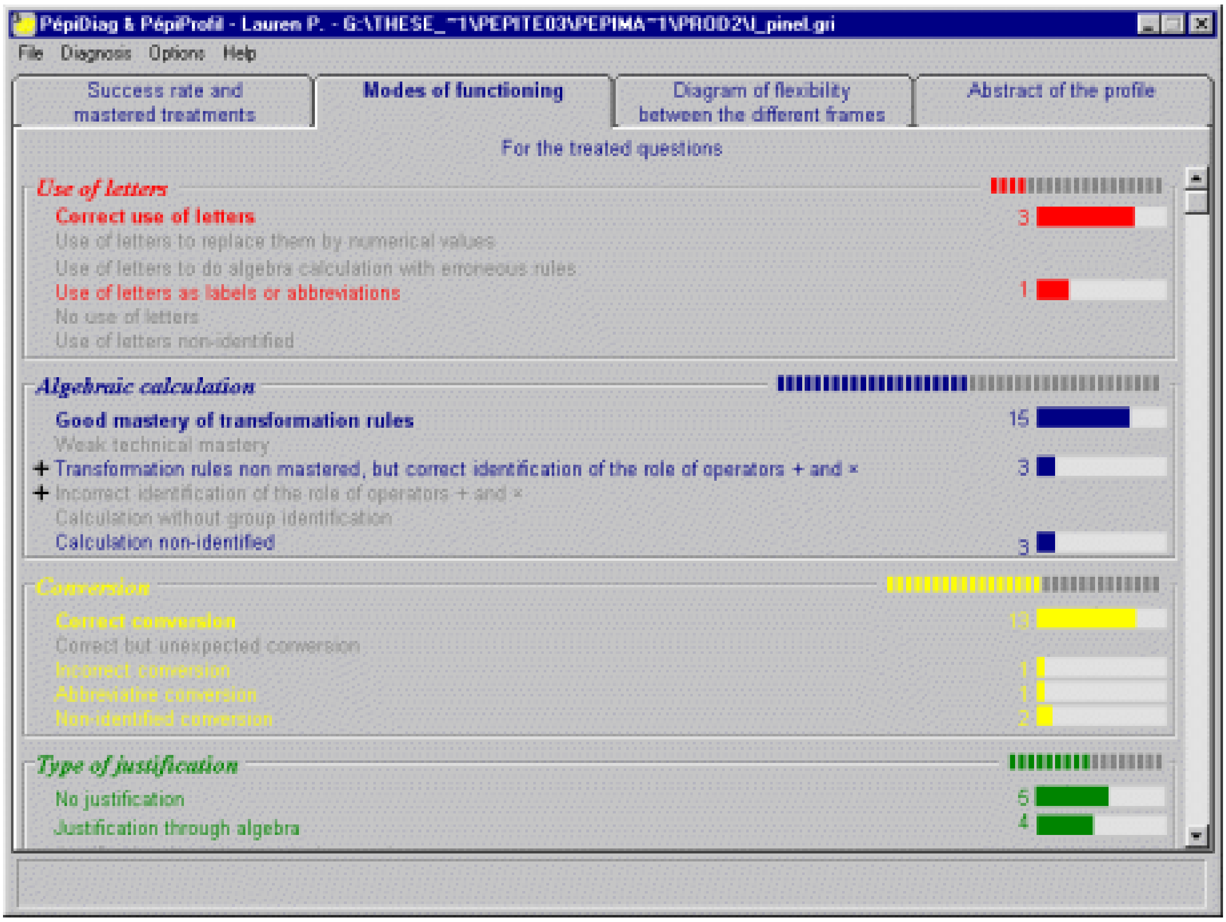

Figure 4 : Laurens' cognitive profile, the qualitative description

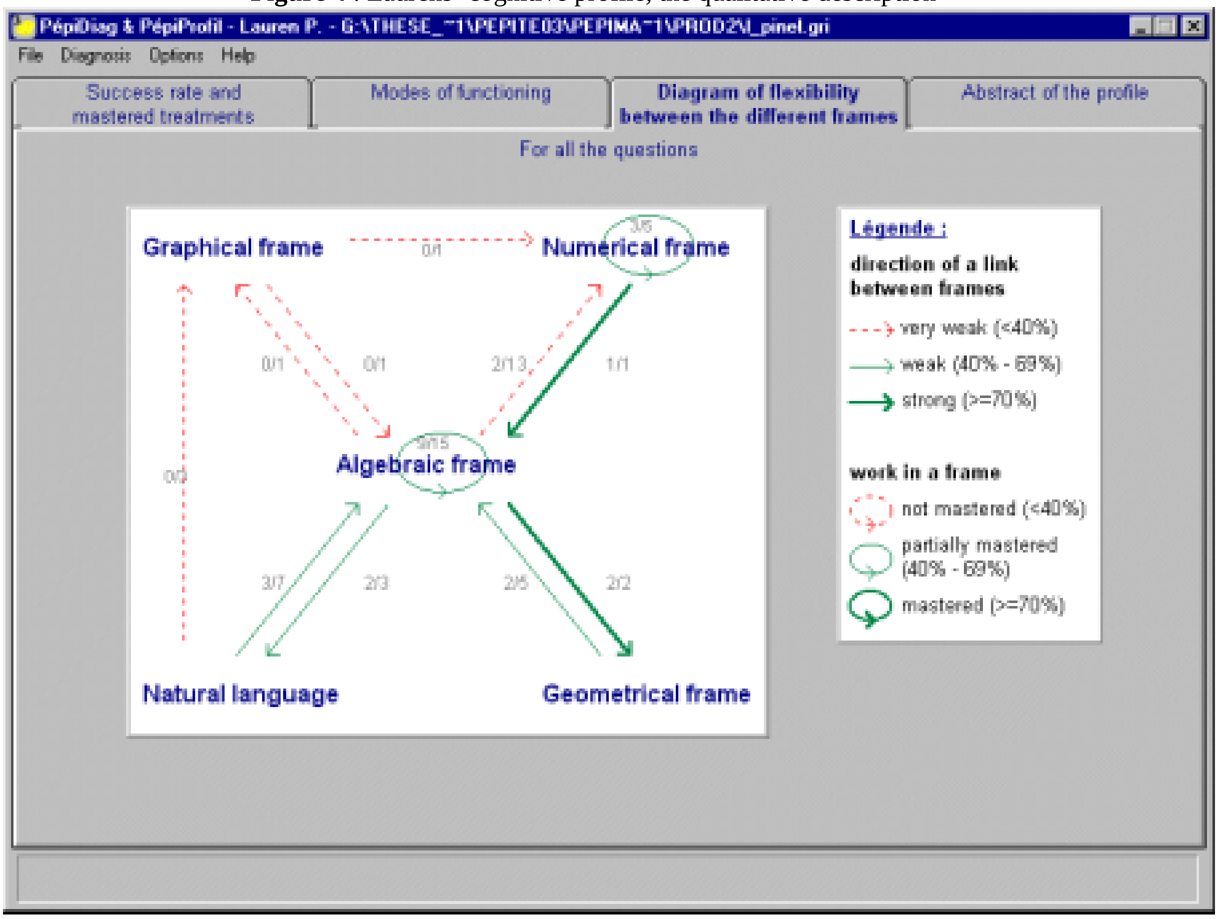

Figure 5 : Lauren's cognitive profile, the diagram of flexibility between representations

PepiProf is the teacher’s software; it establishes the student’s profile from the filled matrix by transversal analysis and presents it to the teacher. It also provides an interface to modify the coding of student’ answers (i.e. to modify the diagnosis matrix without showing it to the teacher) in order to allow the teacher to control the software coding and to correct or complete it when necessary.

Figure 2 shows this interface. In Lauren’s answer shown in figure 1, PepiDiag was successful when diagnosing: incorrect treatment, correct uses of letters, unsuitable use of brackets leading to correct result, correct translation from natural language to algebraic expression, explanation through algebra. The teacher can modify the coding if necessary. Figure 3, 4 and 5 shows a student’s profile with the three descriptions instantiated with Lauren’s profiles.

3.3 Uses of the first Pepite software

We have observed uses of Pépite in different contexts. First we have tested PepiTest in lab, videotaping one student. Then more than two hundred students from French secondary schools (grade 9 or 10) had the test. After that educational researchers used Pépite for educational studies in order to identify patterns of regularities in the 200 students’ profiles. Pépite has been used in workshops by educational researchers or teachers trainers. It had been used in teachers training sessions by pre-service teachers or by experimented teachers in professional development programs. Two pilot studies have been led with volunteer experienced teachers. Finally some teachers told us about using Pépite without any observer. Table 1 synthesize these contexts.

Table 1 : different contexts of use for Pepite 1

Context	Situation	Users	Number	Data collected
Testing students	classrooms	students	200	Students’ answers Questionnaire Observations Research Reports
Educational research	Detecting regularities in students’ profiles	researchers	3	List of usability problems, of bugs
Teachers trainers development	Studying one student	Teachers trainers	40	List of tradeoffs Questionnaires
Teachers training	Studying one student, studying algebraic competencies	Pre-service and in service teachers	100	Questionnaire Observations
Pilot session	Classroom (Individual support and performance appraisal)	teachers	3	Observations reports
Spontaneous uses	Classroom	teachers	9	report

4. Results

All those tests show several results.

PepiTest gathers data that can be used for the diagnostic. First, researchers can study students’ answers to PepiTest and manually build a cognitive profile. Secondly if we consider 50 students’ answers to PepiTest exercises we obtain every kind of answers predicted by the didactical analysis. That means that the software does not reduce the variety of answers in comparison with the paper and pencil test. Thirdly, as expected, students have problems to enter algebraic expressions but those difficulties do not prevent from producing them. Fourthly, one research study shows that, in spite of local differences, experienced teachers recognise their students' coherences of reasoning in comparison with paper and pencil performance. Finally in pilot sessions and spontaneous uses, teachers report that the software shows competencies or fragilities they had not noticed their students had. The main reason is that PepiTest proposes more various exercises than those usually done in mathematic classrooms. So teachers' trainers think it is a useful tool to increase awareness on different aspect of algebra competencies.

Usages of PepiDiag demonstrate some inconsistencies in the analysis grid when used systematically by the software. Researchers have to complete the 25 % of answers that are not analysed and to correct about 10 % of Pépite coding of student's answers. Thus in the second Java version, we have modified the analysis grid to simplify it and correct inconsistencies and we have improved the analysis of algebraic expressions to detect the diagnosis items better.

PepiProf supports two teachers' tasks: completing and studying the coding of student's answers and scrutinizing the student's profiles. The interface supporting the coding task is adapted and used by every category of users:

- It gives teachers a framework to interpret the students' answers.
- They understand the diagnosis items when they see them in the context of students' answers.

It is particularly appreciated by pre-service teachers and by teachers' trainers because it gives an entry to understand students difficulties in algebra.

The interface supporting scrutinization of student's profile is used by researchers .They specially appreciate the possibility to access the diagnosis items by different ways : from the students answers, from the general item, from the list of questions related to this items. But this interface raises difficulties when used by teachers, especially with experienced teachers:

- They had difficulties to understand diagnostic items when not in the context of the tasks and students' answers.
- This interface implements a big didactical expertise that can be in opposition with spontaneous diagnosis practises they have developed.
- In the present state of the software development, there are no teaching strategies associated with the cognitive profile, so teachers are not motivated to scrutinize the profile without a precise objective.

Globally, we observed several spontaneous usages of PepiTest as learning activities instead of diagnosing activities, especially to support discussions in classrooms or in pairs. From the questionnaires we obtained several suggestions. As an assessment tool, teachers think that Pépite is time consuming. They would like to choose exercises. They ask for having several tests, according to different levels and in order to assess learning evolutions. Most of them ask for a feedback to students given by the software, because they think it is impossible for them to give a personal feedback. Some of them ask for a "profile of the class" instead of individual profiles especially to organise learning activities in the beginning of a year or to create working groups. Most of them ask for proposition of teaching strategies to remediate to the difficulties that have been diagnosed. Some of them reported that Pépite demonstrates competencies for some students with big difficulties that the teachers had not noticed before and then increases the teacher's confidence in the student's chance of success. By the way, it increases students' confidence in the teacher.

From the usage of Pépite by researchers we notice a very interesting fact. Experts do not practice diagnostic as they describe it in the method they have proposed and that is implemented in Pépite. First they use an adaptive testing. They examine students' answers to one meaningful exercise and according to the answer they state a general hypothesis and go to confirm it and to refine it on very few other exercises. Then in the context of giving a teaching strategy to remediate for a particular student, they give their diagnosis with diagnostic items of the analysis grid but contextualised to exercises and not in global and abstract terms like those used in the cognitive profiles description.

In relation with the whole Lingot project, those first experiments confirm the interest to find classes of cognitive profiles related to teaching strategies.

5. Lessons and questions

As far as students' activity is concerned the model of the algebraic activity issued from educational research allowed us to implement a software that gathers on a computer data that help teachers to identify their students' difficulties in algebra. Even when teachers do not know the didactical model, the way is has been implemented in the test and in the coding interface is well accepted and used. The limitations of this model are that it is predefined and specific to a school level (15-16 years old students in French secondary schools). These lead us to new questions :

- How to allow teachers to adapt testing to their classroom practice? How to determine the parameters?
- Is it possible to define a model of competencies at each school level?
- Is it possible to define patterns of exercices from which teachers could make their own tests?

Concerning the diagnostic system, we have proved that it was possible to automate partially the diagnostic by implementing Grugeon's diagnostic model. But how can we improve the diagnostic in order to be able to study automatically big corpus or to provide feedback to students? Can linguistic or statistical methods help? Is it possible to define diagnostic patterns related to exercise patterns to generate the diagnostic when a teacher has defined a test?

Concerning the teacher's software, the only part that has been used by teachers in classrooms is the coding interface. To improve usability it is important that the diagnosis software has reflexive abilities to evaluate a degree of confidence for the coding produced. How can we introduce these abilities in the diagnosis software?

Concerning the whole system, it seems that the model of diagnosis proposed by Grugeon's work is too rational and prescriptive to be followed by humans. We noticed that Grugeon herself uses a dynamic and adaptive diagnosis related to the students' answers but also to the objective of the diagnosis: to define a teaching strategy. So the model she had proposed is a model of predicted task not of effective task. This explains that it is well accepted by beginners but not by experts. Differences between beginners and experts teachers can be seen in the terms used to describe the student's profile. Adaptive testing could save time [2] and could also be closer to experienced teachers diagnosing methods.

This work conducted with mathematics teachers and researchers showed that there are several methods to diagnose and they are linked to the people that diagnose and to the usage to be done with the diagnostic. Is it possible to identify several classes of diagnostic usages, and for each class a specific software to support it ? Is Grugeon's model of students' competencies robust enough to found all these software? In particular, in this first work, we have focussed on transferring paper and pencil tasks on a computer. How to modify the model to manage with algebraic task without equivalent in paper and pencil context?

We are working on these questions, especially

- on defining several scenarios of diagnosing situations for several types of users in different contexts;
- on modelling exercices and diagnostic methods in order to generate patterns of tests that could be adapted by teachers;

- on defining classes of cognitive profiles and teaching strategies related to each class of profile;
- on studying the feedback to students.

6. Conclusions

In this paper we studied on one example how an ITS can be designed by coordinating educational issues, software design issues and teachers development issues. We highlighted how the design and experiment in different contexts of a prototype show results and rise new research directions. We show that the solution of competencies diagnosis with ITS not only relies on the intelligence of the machine to understand students but also relies on the intelligence in the interaction with users to support them in their daily activities.

Acknowlegment

This research has been partially funded by the "Programme Cognitique, école et sciences cognitives". Numerous colleagues from IUFM of Creteil and teachers are acknowledged for testing Pépite with their classes.

References

[1] I. Arroyo, A. Schapira, B. P. Woolf, Authoring and sharing word problems with AWE, in J.D. Moore et al. (Eds.), Artificial Intelligence in Education, IOS Press, 2001, 527-529.

[2] I. Arroyo, R. Conejo, E. Guzman, B. P. Woolf, An Adaptive Web-Based Component for Cognitive Ability Estimation, in J.D. Moore et al. (Eds.), Artificial Intelligence in Education, IOS Press, 2001, 456-466.

[3] M. Artigue, T. Assude, B. Grugeon, A. Lenfant, Teaching and Learning Algebra : approaching complexity trough complementary perspectives, In Helen Chick, Kaye Stacey, Jill Vincent, John Vincent (Eds), The future of the Teaching and Learning of Algebra, Proceedings of 12 th ICMI Study Conference, The University of Melbourne, Australia, December 9-14, 2001, 21-32.

[4] J. M. Carroll, G. Chin, M. B. Rosson, D. C. Neale, The development of Cooperation: Five Years of Participatory Design in the Virtual School,, in John M. Carroll (ed.), Human-Computer Interaction in the New Millenium, Addison Wesley, 2001, 373-418.

[5] H. Chick, K. Stacey, J. Vincent, J. Vincent (Eds), The future of the Teaching and Learning of Algebra, Proceedings of 12 th ICMI Study Conference, The University of Melbourne, Australia, December 9-14, 2001.

[6] É. Delozanne, B. Grugeon, P. Jacoboni, "Analyses de l’activité et IHM pour l’éducation ", In Proceedings of IHM’2002, International Conferences Proceedings Series, ACM, 2002, Poitiers, France 25-32

[7] É. Delozanne, P. Jacoboni, S. Jean, B. Grugeon, Assessing Students' Competence in Algebra, Workshop "Learning Algebra with the Computer, a transdisciplinary Workshop”, ITS 2000, Montreal, June 2000.

[8] B. Grugeon, Etude des rapports institutionnels et des rapports personnels des élèves à l’algèbre élémentaire dans la transition entre deux cycles d’enseignement : BEP et Première G, thèse de doctorat, Université Paris VII, décembre 1995.

[9] S. Jean, E. Delozanne, P. Jacoboni, B. Grugeon, A diagnostic based on a qualitative model of competence in elemantary algebra, in S. Lajoie, M. Vivet, AI&ED’99, IOS Press, Amsterdam, , Le Mans (1999) 491-498

[10] C. Kieran, The Learning and Teaching of School Algebra, Handbook of research on Mathematics Teaching and Learning, Douglas Grouws Ed. Macmillan publishing company, 1992.

[11]K. R. Koedinger, J. R. Anderson, Intelligent Tutoring Goes to School in the Big City, IJAIED (8), 30-43, 1997.

[12] S. M. Land, M. J. Hannafin, Student-Centred Learning Environments, in D.H. Jonassen, S. Land (Eds.), Theoritical Foudations of Learning Enviroments, Lawrence, Erlbaum Associates, 2000, 1-19

[13] P. Rabardel Instrument Mediated Activity in Situations, in Blandford A., Vanderdonckt J., Gray Phil (eds.), Joint Proceedings of HCI’2001 and IHM’2001, 17-30

[14] D. Rasseneur, E. Delozanne, P. Jacoboni, B. Grugeon, Learning with virtual agents: Competition and Cooperation in AMICO, Proceedings of ITS’2002, Biarritz (France), 5-8 juin 2002. Cerri S., Gouardéres G., Paraguaçu F. (eds.), Springer-Verlag, p. 61-70.

[15] M.E. Verona, D. Curtis, D. Shaffer, Supporting Teacher development in Enacting the RiverWeb Water Quality Simulator, in J.D. Moore et al. (Eds.), Artificial Intelligence in Education, IOS Press, 2001, 87-98.

[16] http://pepite.univ-lemans.fr