

Student's modelling with a lattice of conceptions in the domain of linear equations and inequations

Jean-François Nicaud, Denis Bouhineau, Hamid Chaachoua, Marilena Bittar

► To cite this version:

Jean-François Nicaud, Denis Bouhineau, Hamid Chaachoua, Marilena Bittar. Student's modelling with a lattice of conceptions in the domain of linear equations and inequations. AI-ED, 2005, Amsterdam, Netherlands. hal-00190223

HAL Id: hal-00190223 https://telearn.hal.science/hal-00190223

Submitted on 23 Nov 2007

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.

Student's modelling with a lattice of conceptions in the domain of linear equations and inequations

Jean-François NICAUD*, Hamid CHAACHOUA**, Marilena BITTAR***, Denis BOUHINEAU*

* Université Joseph Fourier, Leibniz-MeTAH, 46 Ave Viallet, 38031 Grenoble, France Jean-Francois.Nicaud@imag.fr, Denis.Bouhineau@imag.fr ** IUFM-Leibniz-MeTAH, 46 Ave Viallet, 38031 Grenoble, France, Hamid.Chaachoua@imag.fr *** Universidade Federal de Mato Grosso do Sul, Campo Grande, Brasil, marilena@nin.ufms.br

Abstract. We present a student's modelling process in algebra which consists of two phases. The first phase is a local diagnosis where a student's transformation of an expression A into an expression B is diagnosed with a sequence of rewriting rules. A library of correct and incorrect rules has been built for that purpose. The second phase uses a lattice of conceptions built for modelling students more globally. Conceptions are attributed to students according to a mechanism using the local diagnoses as input.

This modelling process has been applied off-line to data gathered in France and Brazil with 13-16 years old students who used the Aplusix learning environment. The results are described and discussed. The process will be included later inside Aplusix in order to model students on-line and to provide the students' conceptions to the teachers.

Keywords: student's modelling, algebra, rewriting rules, conceptions.

1. Introduction

Since 2003, we are engaged in a research work¹ devoted to automatic student's modelling in algebra. Our general model is based on correct and incorrect *rules* allowing to interpret the local behaviour of the student, and on *conceptions* allowing to model the student in a more global way. Our methodology contains five components: (1) Gathering a large set of data concerning students' calculations; (2) Construction of a library of correct and incorrect rules; (3) Automatic diagnosis of the student's calculations in term of sequences of rules; (4) Construction of a library of conceptions; (5) Automatic diagnosis of the student's calculations in term of sequences.

According to Artigue [1], a conception is related to a concept and is characterized by three components: (1) a set of situations which give meaning to the concept; (2) a set of significations (mental images, representations, symbolic expressions); (3) tools (rules, theorems-in-act, algorithms). This paper is mainly devoted to theorems-in-act associated to the concept of movement in (in)equations. An example of theorems-in-act is the following: *When a sub-expression is moved from one side to the other side in an (in)equation, its sign is always changed*. Note that this is correct for additive movements (e.g., $x-3 = 5 \rightarrow x = 5+3$) and incorrect for multiplicative movements (e.g., $-3x = 5 \rightarrow x = 5/3$). This theorem-in-act is not a rewriting rule: it applies to all the movement rules (and generally, several theorems-in-act of the term theorem-in-act for fluidity reasons. The reader needs just to remember that conception means, in the paper, a theorem-in-act part of a conception.

¹ This work is funded by the programme 'ACI, Ecole et Sciences Cognitives' of the French Ministry of Research.

Several research works have been devoted to conceptions and some conceptions have been built by hand [1, 3, 4]. However, automatic diagnosis of conceptions is a difficult problem which has been little investigated. This is the part of our research which is mainly described in this paper. Examples of previous works in this field are Pépite [7] and Baghera [10]. We will compare our work with Pépite in the last section.

For gathering the data to be analysed, we have used the Aplusix learning environment [8, 12] which allows students to freely make calculation steps, as they do in the paper environment, and which records all the students' actions in log files. Except for 6 classes, Aplusix has been used in an ecological context, i.e., by students supervised by their mathematics teachers during the normal schedule of the class. Data have been gathered in France and Brazil from about 3000 students of grades 8 to 11.

The long term goal of this work is twofold: First, we aim at building a map of conceptions covering the domain of elementary algebra (algebra for grades 8 to 11) and to diagnose thousands of students from many countries, to get a repartition of the obtained conceptions; this is our academic goal. Second, we plan to encompass the map of conceptions and the diagnosis mechanism in the Aplusix system to improve its interest by providing to the teachers the conceptions of their students, and by selecting the best exercises to give to the students according to the calculated conceptions (in particular to correct the erroneous parts of these conceptions); this is our applied goal which is similar to Cognitive Tutors [6] in a larger domain. This paper presents the ongoing work on the movement concept in (in)equations. A detailed description of the different parts of the work is available in a research report [11].

2. Modelling student's calculations with rewriting rules

Most of the calculations in formal algebra concern the application of rewriting rules, according to the "replacement of equals" principle. Given two expressions A and B, a rewriting rule R: $A \rightarrow B$ can be applied to an expression E if there is unification between A and a sub-expression U of E. The application of the rule, when possible, consists of replacing U in E by B. Rewriting rules usually come from algebraic identities provided by axioms and theorems. For example, the identity A(B+C)=AB+AC produces two rewriting rules, one is the expansion rule

 $A(B+C) \rightarrow AB+AC$, the other is the factorisation rule $AB+AC \rightarrow A(B+C)$. Most of the rewriting rules of elementary algebra can be classified in *reduction and simplification rules* (numerical calculations, like term collection, e.g., $3x+5x \rightarrow 8x$, etc.), *factorisation rules* (which provides additional factors e.g., $AB+AC \rightarrow A(B+C)$), *expansion rules* (inverses of factorisation rules), and *rules on relations* (see below). When a student solves an exercise in formal algebra, he/she produces calculation steps. Our model interprets these steps as the application of correct or incorrect rules. For example, if the student transforms $2x(3x^2-4)$ into $6x^3-4$, we can interpret this calculation step by the application of the incorrect rule $A(B+C) \rightarrow AB+C$ followed by the application of correct reduction rules.

The main strategy for solving linear equations and inequations consists of expanding both sides, if necessary, then isolating the variable, using rules which carry out identical operations on both sides or using movement rule that move an additive or multiplicative expression from one side to the other. Examples of these rules are: $A=B \rightarrow A+C=B+C$, addition to both sides; $A+C=B \rightarrow A=B-C$, additive movement; $AC=B \rightarrow A=B/C$ ($C\neq 0$), multiplicative movement. There are 12 movement rules for equations (4 additives and 8 multiplicative) and also 12 for \neq inequations. For the other inequations ($\leq \geq \geq$), multiplicative movement rules are duplicated according to the semantic sign of C, because when this sign is negative, we have an inversion of the inequality (e.g., $AC \le B \rightarrow A \ge B/C$ (C<0)). Therefore, there are 20 movement rules for each sort of inequations (4 additives and 16 multiplicative). In the school practice, operations on both sides of (in)equations are progressively replaced by movement rules which are the fundamental rules for solving linear (in)equations. Other rules are expansion and reduction rules, which do not apply to the entire (in)equation but to a sub-expression.

The very large quantity of movement rules (104 = 2*12+4*20) questions the relevance of the rule model. Is a student supposed to have all these detailed rules in mind when he/she solves an (in)equation? Certainly not in this form. Hence, we propose another model which emphasizes the main features and we consider at the same time incorrect movements. A movement has an argument which is the element moving from one side to the other side. In an additive movement, the argument is located in a sum or is the whole side. When the movement is carried out correctly, the argument remains additive, on the other side, with an opposite syntactic sign. In a multiplicative movement, the argument can be multiplicative within a numerator, or multiplicative within a denominator. When the movement is carried out correctly, the argument is negative in the other side, and its place changes (denominator \leftrightarrow numerator). In the case of an inequation, the orientation of the inequation changes if the semantic sign of the argument is negative. We can describe movements with only one general rule, entitled "Movement", associated to the vector presented in table 1.

Dimension	Possible values				
Symbol of relation	$= \neq < \leq > \geq$				
Horizontal orientation	LeftToRight, RightToLeft				
Vertical orientation	NumeratorToNumerator, DenominatorToNumerator,				
	NumeratorToDenominator, DenominatorToDenominator				
Initial position of the argument	InitAdditive, InitMultiplicative				
Final position of the argument	FinAdditive, FinMultiplicative				
Change of syntactic sign of the argument	SynSignChanged, SynSignUnchanged				
Change of the inequality orientation	OrientationChanged, OrientationUnchanged				

Table 1. The seven dimensions of the general Movement rule.

For example, the incorrect transformation $2x-4 < 5 \rightarrow 2x > 5-4$ is represented by a Movement with the argument "-4" and the vector (<, LefToRight, NumeratorToNumerator, InitAdditive, FinAdditive, SignUnchanged, OrientationChanged).

3. A local diagnosis algorithm

We have implemented an algorithm to diagnose student's local algebraic transformations in terms of rewriting rules. The algorithm uses a library of 260 correct and incorrect rules. This library has been obtained from a cognitive analysis (we observed students' actions in various situations using the replay system of Aplusix) and from an epistemic analysis (as the one carried out for the movement rule), see details in [9, 11]. The incorrect rules of the library currently covers operations on (in)equations, reductions (except operations on fractions and square roots) and expansions. Fractions, square roots and factorisations will be studied later.

The local diagnosis algorithm has three phases. The first phase isolates the subexpressions where the transformation occurs. As a result, an expression A has been transformed into an expression B by the student. The second phase is a heuristic search algorithm which develops a tree from the starting expression A. At each step, the node of the search space which is the closest to B, according to a distance between expressions, is chosen and developed.

The development consists of applying the applicable rules of the library. When the

development produces the expression B, the goal is reached and the path from A to B in the tree is a sequence of rules that explains the transformation of A into B. The second phase stops when a chosen number of nodes have been developed or when no more rules are applicable. So this phase can fail. It can fail because of: a missing incorrect rule, an early stop of the process, a student's behaviour that has not to be understood. The third phase consists of the evaluation of the different diagnoses and of the choice of the best one. For example, the transformation of 2x-6 = 7x-8 into -5x = -14 is diagnosed with 4 rules: (1) Incorrect additive move of 6 leading to 2x = 7x-8-6; (2) Correct additive move of 7x leading to 2x-7x = -8-6; (3) Correct additive reduction, leading to -5x = -14.

The current performance of the local diagnosis, in terms of success/failure, for classes of grades 8 and 9, in France (540 students) and Brazil (2500 students), is the following: Between 90% and 100% of success for correct transformations, depending of the class, and between 74% and 93% of success for incorrect transformations. Note that a failure is sometimes the best diagnosis (researchers don't always explain a student's transformation). The correctness of the diagnoses has been studied by three researchers for two French classes, grades 8 and 9, for incorrect expansions, incorrect reductions and incorrect transformations on in(equation): between 82% and 97% diagnoses have been considered to be correct, depending of the class and the category of rule (expansions, etc.).

4. A lattice of conceptions

We have designed a set of conceptions for the concept of movement which is organised as a lattice. Each conception which is not a micro-conception has two subconceptions and is the union of these two subconceptions. The conceptions at the first level are called "global conceptions" with respect to an aspect of the concept of movement. We have considered three aspects of the concept of movement: the *sign aspect* (whether the sign of the argument is changed or not), the *inequality orientation* for inequations (whether the orientation of the inequality is changed or not) and the *operator evolution* (what happens to the operator linking the argument to the (in)equation in the movement). Figure 1 shows a part of this lattice. We have defined five global conceptions for the *sign aspect*:

- CorrectSign: Correct treatment of the sign of the argument;
- AbsoluteValue: Change of the sign of the argument if and only if this sign is "-";
- SemiAbsoluteValue: Change of the sign of the argument if and only if: this sign is "–" and the argument is multiplicative; or the argument is additive;
- SaveSign: Never change the sign of the argument;
- ChangeSign: Always changes the sign of the argument.

Of course, CorrectSign is the only correct conception in that list. The other conceptions produce correct or incorrect calculations depending on the context. The *inequality orientation* and the *operator evolution* aspects lead to similar decompositions in global conceptions.

Let us detail the ChangeSign global conception of the sign aspect. This global conception consists of "always changing the sign" of the argument of a movement. It can be decomposed in partial conceptions: ChangeSign-eq for "always changing the sign in an equation" and ChangeSign-ineq for "always changing the sign in an inequation". Again, ChangeSign-eq can be decomposed in more partial conceptions ChangeSign-eq-add for "always changing the sign in an equation for an additive argument" and ChangeSign-eq-mult for "always changing the sign in an equation for a multiplicative argument". In this framework, some students may have a ChangeSign-eq and not a ChangeSign-ineq

conception. Those who have both ChangeSign-eq and ChangeSign-ineq conceptions have the ChangeSign conception.

For ChangeSign-eq-add (Y-eq-add in figure 1), we have two micro-conceptions named 1b and 2b. The first one is changing the sign when it is "+" in an equation and when the argument is additive; the second one is changing the sign when it is "-" in an equation and when the argument is additive.

In order to have a lattice, we add some nodes considered as abstract conceptions: three nodes containing the aspects (SignAspect, InequalityOrientation and OperatorEvolution) and having the global conceptions as direct descendants, a "Top" node having these three nodes as direct descendants, and a "Bottom" node being direct descendant of every micro-conception. The lattice has a total of 146 nodes.

Figure 1. Part of the lattice for two global conceptions of the sign aspect: CorrectSign (X in the figure) and ChangeSign (Y in the figure). The abstract nodes are not represented. X has an Xbis conception which is another decomposition of the same conception (X is first decomposed in eq/ineq then add/mult; Xbis is first decomposed in add/mult then eq/ineq). X and Xbis are active at the same time. Xbis is necessary to allow to have X-add and X-mult in the lattice.



5. Computation of the student's conceptions

The computation of the student's conceptions uses first the local diagnoses described in section 3, and an intermediary construction: the Local Behaviour Vector (LBV) which links movement rules and micro-conceptions. The movement rules of the local diagnoses which match LBVs are counted for each LBV. When a condition is verified for a LBV, a corresponding micro-conception is activated. Then a propagation mechanism is launched to activate other conceptions.

For building the Local Behaviour Vectors, we use some elements of the vector of the unique movement rule and some other useful elements. This is the decision of the observer, as emphasised by [3]: "Modelling behaviour requires a first level of interpretation, that of the organization of reality". The LBVs we have chosen for the *sign aspect* is described in table 2.

Dimensions (or variables)	Possible values		
Type of the exercise	Equation, Inequation		
Initial position of the argument	InitAdditive, InitMultiplicative		
Sign of the argument	SignArgPlus, SignArgMinus		
Change of syntactic sign of the argument	SynSignChanged, SynSignUnchanged		

Table 2. The seven dimensions of the Local Behaviour Vector for the sign aspect.

The first three variables define the context, the last one corresponds to the action. When the first three variables are chosen, the two possible values of the last one determine a couple of opposite LBVs. An example of a couple of opposite LBVs is the following: the context is (Equation, InitAdditive, SignArgMinus); one LBV has the action SignChanged, the other has the action SignUnchanged.

When we have a list of rules attributed to a student by the local diagnoses of a set of transformations, we match each movement rule to each LBV and count the number of occurrences. A student who has a rational behaviour, with respect to the model, would mark only one LBV of a couple of opposite LVB. However, we cannot expect to have this level of rationality. As a consequence, we have chosen the following mechanism to activate LBVs: Let LBV1 and LBV2 be two opposite LBVs. Let n1 and n2 be the respective numbers of occurrences of LBV1 and LBV2 for a given student:

IF n1+n2 = 0 THEN there is no LBV activation ELSE IF $n1 / (n1+n2) \ge 2/3$ THEN LBV1 is activated with coefficient n1 / (n1+n2)ELSE IF $n2 / (n1+n2) \ge 2/3$ THEN LBV2 is activated with coefficient n2 / (n1+n2)ELSE there is no LBV activation

An activated LBV becomes a micro-conception which can be expressed by: "In the context of this LBV, the student generally performs the action of the LBV". Generally is expressed by the coefficient which is a sort of *Certainty Factors* [5].

The micro-conceptions are the lowest real conceptions in the lattice (having Bottom as direct descendant). After the activation of the micro-conceptions, each upper conception is determined by calculating recursively its coefficient as the geometrical average of the coefficients of its two direct descendants (if the coefficients of the two direct descendants are a and b, the result is sqrt(ab)).

A micro-conception is correct or incorrect. When it is incorrect, any behaviour that matches the micro-conception is incorrect. A conception which is not a micro-conception is correct if and only if its two direct descendants are correct. When a conception, which is not a micro-conception, is incorrect, some behaviours that match the conception are incorrect but, generally, others are correct. For example, the ChangeSign conception in which "the student always changes the sign" contains correct and incorrect behaviours.

6. Experimental study

Since 2003, we have conducted several experiments with classes in France and Brazil of grades 8, 9 and 10 and we have recorded thousands of hours of students' activities. We have recently applied the modelling process to a part of the data. It produced a description of each student in terms of a list of conceptions attributed to this student, and a summary table containing the number of occurrences of each conception.

An experiment with a group of 342 students of grade 9 was conducted in Campo Grande (Brazil) in 2004 with 20 minutes of use of Aplusix for familiarisation and 1 hour of use in the test mode where no feedback was given to the students. The modelling process has been applied to the data corresponding to the test phase. The analysis of the distribution of the conceptions with respect to the context shows what follows:

• Type of exercise: 56% of conceptions concern equations (97% correct and 3% incorrect); 44% of conceptions concern inequations (71% correct and 29% incorrect).

Initial position of the argument: 62% of conceptions concern an additive position (95% correct and 5% incorrect); 38% of conceptions concern a multiplicative position (68% correct and 32% incorrect).

Most of the conceptions concern equations with an additive position of the argument. The high rate of correct conceptions cannot be viewed as certitude of a good result because the level of generality of the conceptions has to be taken into account, the more general ones being the "global conceptions" defined in section 4. Actually, we had only 2% of correct global conceptions (e.g., CorrectSign). At the other levels (the level is the depth in the lattice, the global conception having level 1), we have 32% correct conceptions for level 2 (e.g., CorrectSign-eq), and 66% correct conceptions concerning very specific contexts of levels 3 and 4. These results are coherent with hand analysis made for a few students and with the general opinion of the teachers of the classes. The distribution of the number of conceptions per student is a Gaussian distribution, see table 3.

Table 3. Distribution of the numbers of conceptions. Note that the ideal student has 3 conceptions: the correct global conception of each of the 3 aspects of the movement concept.

Number of conceptions	0	1	2	3	4	5	6	7	8	Total
Number of students	11	37	58	80	38	47	48	22	1	342
Percentage	3.2	10.8	16.9	23.4	11.1	13.7	14	6.4	0.3	100

The distribution of the conceptions with respect to the aspects of the movement concept is shown in table 4. There are many correct conceptions for *Sign aspect* and *Operator evolution*, but just a few of them are at level 1. There is an important amount of incorrect conceptions at level 1 for *Inequality orientation*. This is coherent with the fact that these students have had many exercises about equations and not many about inequations in the preceding school year.

Level of the	Sign a	aspect	Inequality	orientation	Operator evolution		
conception	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect	
1	20	0	0	37	0	0	
2	158	0	85	4	158	0	
3	129	11	0	0	186	0	
4	0	0	0	0	335	9	

Table 4. Distribution of the conceptions with respect to the three aspects of the movement concept.

We have applied the modelling process to 221 French grade 10 students, obtaining similar results. The main result for both populations is the following: incorrect conceptions concern two contexts: (1) inequations, (2) multiplicative initial position of the argument.

Last, we have modelled 30 French grade 10 students who used Aplusix during the whole school year 2003-2004. The analysis data collected at the end of the school year shows:

- A total of 171 conceptions with 18% for level 1, 15.2% for level 2 and 30% for level 3;
- 70% students have the CorrectSign correct global conception;
- \circ 40% have the SaveOrientation incorrect global conception (never change the sign of an inequality).

We note that these students, who had a longer training, have more general conceptions, and that the sign aspect is rather well acquired but the orientation aspect of the inequality is not.

7. Discussion and future work

This work is a significant step towards the achievement of the goals we have presented in the introduction. The obtained results are coherent with opinions of teachers and with analyses "by hand" of a few students' data, but a deeper study of the coherence is necessary and will be carried out. However, we need to analyse in depth the data that are not captured by the process. For example when we find that a student has 3 conceptions, we have an interesting result, but in order to achieve a completion goal, we would like to have an opinion about the

behaviours of this student that do not participate to these 3 conceptions. Some may be sleeps, other random behaviour, other rational behaviour not captured by the model.

Let us summarize our work and make a comparison with Pépite [7]:

- We collect very large sets of data in ecological situation; Pépite collect limited sets of data in specific situations;
- Our diagnoses are based on implemented detailed knowledge which means that our system has a deep algebraic understanding of the phenomena; this is not the case of Pépite;
- We are building a wide map of conceptions understandable by teachers. So does Pépite which domain is larger than ours, including other registers than the formal one.
- We are preparing an operational process in which the students will learn with Aplusix in ecological situations and in which Aplusix will calculate on-line the students' conce-ptions and tell them to the teacher; this is not the case of Pépite. As Aplusix has proved to help students learn algebra, and as publishers are very interested in selling Aplusix, we hope that many students of many countries will benefit of this work in a few years.

This way of modelling students is not a dynamic way in the sense of building automatically pieces of knowledge to attribute to students like [2]. There are several reasons for that. First, we think that algebra of grades 8 to 10 is a too complex domain for that goal, when a deep modelling is expected; this complexity led us to consider two levels: the rule level and the conception level. Second, we are fundamentally interested by the map of conceptions and by a capacity to provide understandable descriptions of conceptions to the teachers.

Last, we began to think of a new goal which consists of making a benchmark for student's modelling in algebra. This means to give access to the data we collect and to ask the interested research teams to use these data to model students with their own methods and tools. Results of the researches would be published on a Website and compared. Two levels would be possible: a rule level and a conception level. Tools made by teams could be accessible on the Website.

References

- [1] Artigue M. (1991). Epistémologie et didactique. Recherche en didactique des mathématique, 10 (2/3).
- [2] Baffes P., Mooney R. (1996). Refinement-Based Student Modeling and Automated Bug Library Construction. Journal of Artificial Intelligence in Education, 7(1), p 75-117.
- [3] Balacheff N., Gaudin N. (2002). Students conceptions: an introduction to a formal characterization Cahier du laboratoire Leibniz # 65, <u>http://www-leibniz.imag.fr/LesCahiers/Cahiers2002.html</u>
- [4] Brousseau G. (1997). Theory of didactical situations in mathematics. Kluwer Academic Publishers, Dordrecht.
- [5] Buchanan B.G. and Shortliffe E.H. (1984). *Rule-Based Expert Systems, The Mycin experiments of the Stanford Heuristic Programming Project*. Addison-Wesley Publishing.
- [6] Koedinger, K. R., Anderson, J. R., Hadley, W. H., Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. International Journal of Artificial Intelligence in Education, 8, p. 30-43.
- [7] Jean S., Delozanne E., Jacobini P., Grugeon (1999). A diagnostic based on a qualitative model of
- competence in elementary algebra. AI-ED 1999, IOS Press, Amsterdam, p. 491-498.
- [8] Nicaud, J.F., Bouhineau, D., and Chaachoua H. (2004). Mixing microworld and CAS features in building computer systems that help students to learn algebra, in *Int. Journal of Computers for Mathematical Learning*, Volume 9 (2), Springer-Verlag.
- [9] Sander E., Nicaud J.F., Chachoua H., Croset M.C. (2005) From usage analysis to automatic diagnosis: The case of the learning of algebra. AIED Workshop on Usage Analysis in Learning Systems. Ibidem.
- [10] Webber C., Pesty S., Balacheff N. (2001). A multi-agent and emergent approach to learner modelling. Proceedings of ECAI.
- [11] Modélisation cognitive d'élèves en algèbre et construction de stratégies d'enseignement dans un contexte technologique. Project report of the "Ecole et sciences cognitives" research programme. Cahier #123 of the Leibniz laboratory, http://www-leibniz.imag.fr
- [12] The Aplusix Website. http://aplusix.imag.fr