Baghera Assessment Project, designing an hybrid and emergent educational society

Nicolas Balacheff, Ricardo Caferra, Michele Cerulli, Nathalie Gaudin, Mirko Maracci, Maria Alessandra Mariotti, Jean-Pierre Muller, Jean-François Nicaud, Michel Occello, Federica Olivero, et al.

Les cahiers du laboratoire Leibniz

Baghera Assessment Project, designing an hybrid and emergent educational society

edited by Sophie Soury-Lavergne

n° 81

Laboratoire Leibniz-IMAG, 46 av. Félix Viallet, 38000 GRENOBLE, France -

ISSN : 1298-020X

Avril 2003

Site internet : http://www-leibniz.imag.fr/LesCahiers/

INFORMATION SOCIETY TECHNOLOGIES
IST PROGRAMME

BAP

BAGHERA

Assessment Project

Designing an hybrid and emergent educational society

Contract n° IST-2001-33046
Final report
February 2003

http://www-baghera.imag.fr/

Contents

1 Introduction Outlines of the project

Author: Nicolas Balacheff

The Baghera Assessment Project (BAP) has the objective to explore a new avenue for the design of e-Learning environments. The key features of BAP's approach are: (i) the concept of emergence in multi-agents systems as modelling framework, (ii) the shaping of a new theoretical framework for modelling student knowledge, namely the cK¢ model. This new model has been constructed, based on the current research in cognitive science and education, to bridge research on education and research on the design of learning environments.

This document reports on the results, difficulties and prospective of BAP.

- First, we present the conceptual model cK¢ (for conception-knowledge-concept) which we shaped during the first period of the project. This model provides the background of the whole project. The section 2 presents it in details; a publication annexed to the report (Annex IV) proposes a more comprehensive introduction with the evidence of a complex example.
- Second, based on cK¢, we have explored the principles to be implemented for the design of BAP. This has been an intensive collaborative work between researchers in education and in computer science; it is at this stage that we have experience the full complexity of an integrated multi-disciplinary approach. In the end, we had to make a choice because of the complexity of the problem addressed and the available resources within a single year. Considering that teaching involves two critical tasks, which are (i) understanding students' understandings, and (ii) taking decision about the best feedback to provide them, we have decided to concentrate on the first aspect and to assess emergence on the issue of student modelling. This is presented in a section which introduces emergence and the targeted architecture (section 3), and an other one presenting the methodology and the mock-up (section 4).
- Third, the section 4 include the presentation of a parallel development of an other approach based on automated reasoning. This other approach is complementary to the emergence approach. Since emergence provides only a diagnosis at a behavioural level, while automated reasoning allows an epistemic diagnosis involving an analysis at a content level. In the targeted environment, emergence will suggest diagnosis of which automated reasoning will allow to evaluate the epistemic relevance.
- Fourth, as for the evaluation of the chosen approach, we have confronted human diagnosis to the machine diagnosis based on emergence. To do so, one team has been responsible for autonomously defining the evaluation methodology (the Pisa team), an other one has been responsible for its implementation (the Bristol team), but all teams have been involved in the gathering of data. This is presented in two sections, evaluation methodology (section 5) and evaluation process (section 6). Finally, a section briefly report on the meeting we have organised to get an external feedback on the project. A more formal evaluation has been asked to an expert of the field of ITS design, Peter Brusilovsky from CMI Pittsburgh (as indicated in the contract), this evaluation report is reproduced in section 7.

2 Conceptual framework

Leader of the task: Nicolas Balacheff

Authors: Nicolas Balacheff, Nathalie Gaudin

In the following, we use *knowing* as a noun to distinguish the students personal constructs from *knowledge* which refers to intellectual constructs recognised by a social body. This intends to keep the distinction made in French between "connaissance" and "savoir", or in Italian between "conoscente" and "sapienza".

2.1 cK¢, a knowledge model drawn from an understanding of students understanding

2.1.1 THEORETICAL INTRODUCTION

2.1.1.1 Coherence and sphere of practice

The co-existence of a rational thinking and of knowings which — from the observer's point of view — looks contradictory is a well known paradox. Bourdieu offered the following explanation:

> "In the diagram of the calendar, the complete series of the temporal oppositions which are deployed successively by different agents in different situations, and which can never be practically mobilized together because of the necessities of practice never require such a synoptic apprehension but rather discourage it through their urgent demands, are juxtaposed in the simultaneity of a single space. The calendar thus creates *ex nihilo* a whole host of relations […] between reference-points at different levels, which never being brought face to face in practice, are practically compatible even if they are logically contradictory." (Bourdieu 1990, p. 83).

The key elements of this explanation is time on one hand, and on the other hand the diversity of situations, expressed for a collective body, can be extended to the case of a single agent observed in different situations. Time organises the subject's actions sequentially in such a way that the contradictory knowings are equally operational because they appear at different periods of his history: contradictory knowings can then ignore each other. The issue of the diversity of the situations introduces an element of a different type. It is a possible explanation insofar as one recognises that each knowing is not of a general nature but that, on the contrary, it is related to a specific and concrete domain of validity on which it is acknowledged as an efficient tool. This emphasises that transfer from one situation to an other one is not an obvious process, even if in the eyes of an observer these situations are isomorphic.

Following Bourdieu, we will refer to *sphere of practice* in order to designate these domains of validity mutually exclusive in the history of the subject. Within a sphere of practice the rational subject is reconciled with the knowing subject: in a sphere of practice the subject is coherent, he or she is non-contradictory.

When identified, contradictions are recognised as such by an observer who is able to relate situations which are seen as independent and completely different by the subject herself. In the *observer referential system*, these different states of the observed systems [subject in a situation] should be labelled in the same way if the observer recognises some kind of morphism. So, one may like to speak of the subject knowing of decimal numbers, of continuity of functions or of line reflection even if later on one would complain that this knowing is not coherent.

Since to accept the existence of contradictory knowings is in radical contradiction with the theoretical principle which states that mental constructs are products of a process of adaptation ruled by criteria of reliability and of adequacy to problem-solving or task-performing, we introduce the word "conception" to speak about knowing-in-a-situation.

The following section shapes the meaning of "errors" and its consequences on our view of "knowing".

2.1.1.2. Errors and knowings

Errors and contradiction in learners' productions raises problems in education for a long while. Solutions to understand this phenomenon have been looked for in different directions, especially in the 80s with research on learners misconceptions. We will here comment on the most significant positions adopted and their evolution.

In a survey she presented at the 1986 annual meeting of the *American Educational Research Association*, Confrey links the development of research on misconceptions to the acknowledgement of a failure of teaching. Despite adequate scores on achievement measures, one had to recognise that many students held major misconceptions about fundamental concepts in mathematics and science. At the beginning, the education community had a rather pragmatic approach to this question:

> "misconceptions were defined empirically as documented failures of large numbers of students to solve problems which appeared to be related to fundamental concepts. […] Surprise, pervasiveness, resilience and deviation from the expected answers were its defining characteristics." (Confrey 1986 manuscrit p.4).

Confrey distinguishes (1990, revised version of 1986) three approaches within the misconception paradigm: the Piagetian genetic epistemology, the scientific epistemology and the information processing approach. Each of these three approaches aim at providing a *problématique* for the cases when the students' conceptions appear to be in contradiction with shared and recognised knowledge (ibid. p.4). In all three approaches the child-student is seen as a subject fundamentally different from the adult-expert who appears as the owner of the knowledge of reference (ibid. p.29)—in the case of the former one speaks of "naive theory", "private concepts", "beliefs" or even of the "mathematics of the child". But this view does not exclude the recognition of some sort of cognitive legitimacy of the so-called misconceptions:

> " […] a child may not be 'seeing' the same set of events as a teacher, researcher or expert. It suggests that many times a child's response is labelled erroneous too quickly and that if one were to imagine how the child was making sense of the situation, then one would find the errors to be reasoned and supportable " (ibid. p.29)

This remark, made in the case of the scientific epistemology, is in fact valid for all the three mentioned approaches, even for the Piagetian approach as we emphasise it in the following paragraph.

Within the student frame of reference — as opposed to an external frame of reference —these conceptions fall under the common rules of knowing:

> "a misconception does not require the postulation of an inadequate 'picture' of the world; it does require the notion of a successful completion of a number of problems wherein the cycle of problem formulation (expectation), problem-solving (action) and problem reconstruction (re-viewing) are successfully carried out." (Confrey 1986, p.28)

This statement of Confrey is expressed in close but different terms in her 1990 revised paper, referring to von Glasersfeld:

> "Accepting that we are trapped in our own human ways of knowing, he [von Glasersfeld 1984] suggests that we seek 'fit' rather than 'match' in our conceptual structure, as a key fits a lock. By using this metaphor, he is suggesting that we need to determine if our concepts seem practically viable, rather than objectively true. Therefore he argues that biological adaptation is more appropriate than correspondence for examining learning" (Confrey 1990 p.14)

To sum up: a misconception has a domain of validity otherwise it would not exist. So there is a very short distance between a misconception — as it is now understood — and a knowing. The key difference is that for a misconception there exists a refutation which is known, at least to the observer. But even when ascribing to a misconception the status of a knowing, what leads most authors to abandon the word itself, it remains as a corollary of the initial definition the existence of an *intrinsically correct knowledge of reference*, although such a position is clearly refuted by our current understanding of the history of science and mathematics. Let us notice, at this point, that considering that students' knowing and knowledge of reference are of a different nature, has as a consequence the impossibility of a model which would be a tool to give account of both.

Bachelard (1938, p.13) wrote that reality is never what one could believe, but is always what one should have thought of. This statement, stated in the first half of the last century, expressed that knowledge is always in progress. If we accept this, errors witness the inertia of the instrumental power of knowledge which has proved itself by its efficiency in enough situations, and its organism likelihood in an environment which changes it and which changes itself in its turn.

Aebli (1963), who developed the application of the psychology of Jean Piaget to didactics, characterised errors as witnesses of a student's misunderstanding or habits. He followed up, stating that the erroneous reactions which they could provoke in some problem-solving situation must be studied in detail with students so that they understand the reasons why some processes are not correct, and so that they capture differences and relations between the correct reaction and the error (ibid. p.101). To a certain extend we may suggest that the student is a cognitive subject but not yet a fully knowing subject.

> "*From the functional level, the child is identical to the adult, but with a mental structure which varies depending on the stages of development*" (Piaget 1969, p.224).
>
> "Engaged in a construction process the child *is always obliged to accommodate herself to an external reality, to the peculiarities of the environment from which she has to learn everything*" (ibid. p.225).

The content of the child mental structure has not yet completely the status of a knowing, even though all theoretical ingredients exist to allow to consider it as such. The Piagetian-Copernican revolution was not achieved at the beginning of the 70's.

The main evolution is to recognise that "errors are not only the effect of ignorance, of uncertainty, of chance […], but the effect of a previous piece of knowledge which was interesting and successful, but which now is revealed as false or simply not adapted." (Brousseau 1997, p.82). Salin (1976) proposed cognitive characteristics of errors: on one hand an error is a point of view of a knowing about another knowing (possibly for a subject, the evaluation of an ancient knowing from the point of view of a new one), on the other hand an error can be identified only if the feedback from the milieu can be "read" by the subject as the indication of a failure (a non satisfied expectation).

The thesis of Brousseau at the beginning of the 70s, goes beyond the fact of recognising mental constructs source of errors as knowings. It states that some of these knowings likely to be falsified are necessary to learning: the trajectory of the student may have to pass by the (provisional) construction of erroneous knowings because the awareness of the reasons why this knowing is erroneous is necessary to the construction and understanding of the new knowing.

Following Bachelard (1938 pp. 13-22), Brousseau calls *epistemological obstacles* these compulsory gateways to new understanding:

> "A piece of knowledge, like an obstacle, is always the fruit of an interaction between the student and her surroundings and more precisely between the student and a situation which makes this knowing 'of interest'. In particular, it stays 'optimal' in a certain domain defined by the numerical 'informational' characteristics of the situation." (Brousseau 1997, p.85).

The main difference between the previous position and the current one lies in their epistemological meaning; the status of knowing is different in each case. The first position implies the existence of a knowing-of-reference general and true. The second position, especially in the case of epistemological obstacles, requires only to establish a relationship between two knowings with the idea of an evolution, without judgement on them. Any knowing is what it is, whether it appears to be erroneous or not, partial or ill adapted; it is first of all the result of an optimal adaptation of the subject/milieu system following criteria of adequation and of efficiency.

As a consequence of the nature of the subject/milieu system, any knowing has a provisional character, or rather, any knowing could be revisited, its domain of validity can be modified as a result of some perturbations which it would be otiose to claim that they are unlikely. Indeed, one recognizes here the strong relation which links "knowing" and "problems" for which this knowing is a tool (allowing to get back to an equilibrium).

In the following section we shape the relation between knowing and problem, and we introduce a definition of conception.

2.1.1.3. Knowing, behaviour and conception

The only indicators we have of the good or of the bad subjects' understanding are their behaviours and productions which are consequences of the *knowing* they have constructed. But such evaluations are possible and their results are significant only in the case where one is able to establish a valid relationship between these observed behaviours and the invoked knowing itself. This relation between behaviours and knowing is crucial. It has been hidden as a result of the fight against behaviourism, but it has always been implicitly present in educational research at least at the methodological level.

The key issue is that the meaning of a piece of knowledge cannot be reduced to behaviours, that on the other hand meaning cannot be characterised, diagnosed or taught without linking it to behaviours.

Such a link was clearly pointed out by Schoenfeld in 1987, in a book he edited under the title "*Cognitive Science and Mathematics Education*". In the introduction to the book he associates the Cognitive Science approach to an effort toward a more detailed description of problem-solving behaviours so that they could then be taught and reproduced. This position of Schoenfeld is synthesised by the following indication about his own research at this time:

> "My intention was to pose the question of problem-solving heuristics from a cognitive science perspective: What level of details is needed to describe problem-solving strategies so that students can actually use them?" (ibid. p.18).

But this *problématique* leaves open two essential questions:

- on one hand, to which extent a finer granularity of a description would guarantee a better reliability of the transfer from one operator to an other? Or, rather, for any competency, does there exist a level of granularity which gives an intrinsic guarantee for the efficiency of such a transfer?
- on the other hand, to which extent a finer description of problem-solving behaviours informs about the relationships between behaviours and knowings?

Concerning the latter question, we must notice that Schoenfeld himself finally suggests that this relationship is essential in the chapter of his book devoted to constuctivism—but may be without drawing all the consequences from this remark.

The question of the relationships between behaviours and knowings is considered as fundamental to the theory of didactical situations (Brousseau 1997), since one of its postulates is that each problem-situation demands on the part of the student behaviours that are indications of knowing. This fundamental correspondence, established case by case, is justified by the interpretation of problem-situations in terms of games, and by the interpretation of behaviours in terms of indications of strategies the adapted nature of which must be demonstrated in the model, or of representations attributed to the student (Brousseau 1997, p.215). This postulate is shared by some approach in Cognitive Science: "All behaviour implies a knowing", writes Pichot (1995, p.206).

Indeed, this postulate justifies most of the educational experimental research since students' behaviours are the source of the corpus on which analyses are performed. But to "cut out" a behaviour from the observation of a so-called reality, which could be a classroom or a laboratory experiment, is both a methodological and a theoretical problem as Robert emphasises (1992, p.54).

An observed *behaviour is* not given by the "reality" but taken out of it as *a result of a decision taken by an observer*.

If a "behaviour" is the description of a material relationships between a person and her environment, then it depends on the characteristics of this person as well as on the characteristics of her environment. A good example is the case of instruments which at the same time facilitate action if the user holds the required knowing, and on the other hand limit this action because of their own limitations (Rabardel 1995, Resnick & Collins 1994, p.7). Actually, one may notice that these limitations could be related to material constraints as well as to the knowings involved in the design of these instruments.

"Person" and "environment" refer to complex realities not all of whose aspects are relevant for the type of questioning we are considering. What is of interest for us is the person from the point of view of her relationship to a piece of knowledge. For this reason we will refer from now-on to…

the *subject* as a reduction of the person to her cognitive dimension.

In the same way, we are not interested by the environment in all its complexity, but only by its features which are relevant with respect to a given piece of knowledge. We will call…

milieu such a subset of the environment of a subject;

the milieu is a kind of projection of the environment onto its epistemic dimension.

In many cases (e.g. mathematics), knowings are not only the consequences of the interaction between a subject and a material milieu, but also interactions with systems of signifiers produced by the subject herself, or by others. We must then extend the classical idea of *milieu* in order to integrate symbolic systems and social interaction as means for the production of knowings. This is the meaning of Brousseau's (1997 p.57) proposal to define…

the milieu as the subject's antagonist system in the learning process.

So, we do not consider knowing as a property, which can be ascribed only to the subject, nor only to the milieu. On the contrary we consider it as a property of the interaction between the subject and the milieu—its antagonist system. This interaction is meaningful because it succeeds in fulfilling the necessary conditions for the viability of the *subject/milieu system*. By viability we mean that the subject/milieu system has a capacity to recover an equilibrium following some perturbations; what implies that the perturbation is recognized by the subject (for example a contradiction or an uncertainty). In some cases the subject/milieu system may even evolve if the perturbations are such that this is necessary. This is, in other words, a formulation of Vergnaud's postulate that problems are the source and the criteria of knowings (Vergnaud 1981 p.220).

Problem means here a more or less serious perturbation of the subject/milieu system.

The existence of a knowing can then be evidenced by its manifestation as a problem-solving tool, what is reified as behaviours of the subject/milieu system as it overcomes the perturbations in order to satisfy its constraints of viability. These constraints do not address the way the equilibrium is recovered but the criterion of this equilibrium (we could also say that there is not only one way to know.) Following Stewart (1994 pp. 25-26) we would say that these constraints are *proscriptive*, what means that they express necessary conditions to ensure the system viability, and not prescriptive since they do not tell in detail in which way an equilibrium must be reconstructed (and we may add here that the description searched for by Schoenfeld are more prescriptive than proscriptive).

We can now propose a definition of the meaning of *knowing* which can be pragmatically and efficiently used in a didactical *problématique*.

A knowing is characterised as the state of dynamical equilibrium of an action/feedback loop between a subject and a milieu under proscriptive constraints of viability.

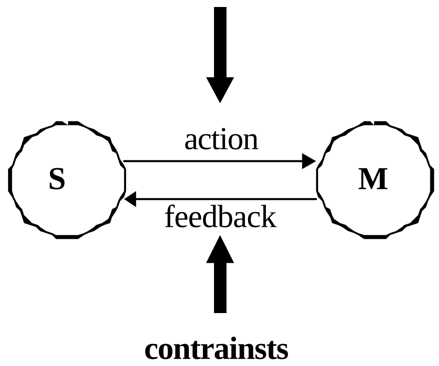


Figure 1. The action/feedback loop between a subject and a milieu under prospective constraints of viability

2.1.1.4 Problems and situations

The characterisation of knowing we have chosen has its roots in both the best understanding we have nowadays about knowledge construction: the intrinsic link between the subject and the *milieu* (basic to situated conceptualisation), the systemic and ecological nature of sense making (basic to constructivism), the key role of problems in the characterisation of both sense making and learning.

Such a characterisation raises to issues which should be here explicitly addressed. The recognition of a problem, i.e. of a perturbation of the system, is not straightforward as we suggested already. But even if there is the identification of a perturbation, it does not mean that its "meaning" is given. Remember the classical injunction: "don't you see the problem!" (for a typical illustration see (Balacheff and Gaudin 2002) in Annex IV). This issue is of a special importance in the case of teaching, since the essential role of the teacher is to make students aware of the existence and of the nature of a problem, which intends to "justify" and/or "encourage" learning. For this purpose, the teacher has to organise *situations* in which he or she manages the introduction of students to their environments so that the [S<->M] system can emerge.

Especially in a teaching/learning context, the environment is not only material (physical and chemical), but also social or even symbolic. This creates a specific difficulty for teachers and trainers. For example, in the case of argumentation on scientific facts, the situation must allow the construction of a milieu which favour the making of a difference between what is specific to the social and institutional setting and what is specific to the knowledge at stake. The issue is to distinguish between argumentation as a tool to convince somebody, from argumentation as a tool to establish the validity of a statement.

The socio-epistemological nature of teaching is a key issue that we recognize and that we will have to address, though *this will not be done within the current project* (it is studied by some of the members of the consortium, but in the framework of other 2).

2.1.2 THE cK¢ MODEL

2.1.2.1 A characterisation of conception

The word “conception” has been used for years in educational research, but most often as a common sense notion rather than explicitly defined. It functions as a tool but its definition remains implicit; it is not taken as a object of study as such (Artigue 1991, p.266). There is a need for a better grounded definition of conceptions, and for tools allowing to analyse their differences and commonalties. A need already noticed by Vinner (1983, 1987). This section aims at proposing a solution to this problem.

We expect a formalisation to be a way to clarify and understand better the complexity pointed out of the multiplicity of a subject’s conceptions which in some cases can even prove contradictory. In the context of this formalisation we propose definitions of the terms “conceptio
will specify the meaning of the words “knowing” and “concepts” in the model.

It should be emphasized that “conception”, “knowing” and “concept” as elements of the model, are abstract entities whose differences lie in their functions and relations in this model and not in their outside possible connotation. Indeed, we must then discuss how far this formalisation is adequate to the question of meaning in education and whether it helps to solve problems raised in our domain.

We call conception C a quadruplet (P, R, L, Σ) in which:

- P is a set of problems;
- R is a set of operators;
- L is a representation system;
- Σ is a control structure.

The three first components are the key features identified by Vergnaud (1991 p.145, this definition was in fact coined at the beginning of the 80's) in order to characterise a concept; the fourth one is introduced for reasons we explain hereafter.

The very first question of any researcher in education will be that of knowing how to relate this formal definition with the “reality” he or she is faced to. We will consider this point for each of the four elements of the definition.

The question of the concrete and pragmatic characterisation of P, the set of problems which enters in this definition, is complex. Two opposite solutions have been proposed:

- to consider all the problems for which the considered conception provides efficient tools to elaborate a solution. This is the option suggested by Vergnaud in the case of additive structures (1991 p.145). This option is not specific enough and fails to help if one considers very basic notions like natural numbers.
- to consider a finite set of problems with the idea that other problems will derive from them. This is the solution proposed by Brousseau following his postulate that “each item of knowledge can be characterized by a (or some) adidactical situation(s) which preserve(s) meaning” (1997, p.30). But this option leaves open the question of establishing that such a generative set of problems can be constructed for any conception.

Instead, we will here adopt a *pragmatic* position, deriving the description of P from the observation of students in situations to be characterised with reference to the related content, the criteria being provided by the specificity of these situations with respect to this content. This raises the issue of the dependency of the proposed model of learner conceptions to possible irrelevant characteristics of experimental situations used (what can be seen as the effects of an experimental contract. A tool to make progress on this question is the analysis of historical and actual uses of mathematics (e.g. Sierpinska 1989, Thurston 1994). A similar approach is that of ethno-mathematics (e.g. d'Ambrosio 1993, Lave 1988, Nuñes *et al.* 1983).

The question of the concrete and pragmatic characterisation of the set of operators and of the representation system is more classical. With respect to them, we follow the today shared idea that operators and representations available for working with a concept are constitutive parts of its meaning. We go here even beyond the line drawn by Sfard when she writes that "a profound insight into the processes underlying mathematical concepts, maybe even a certain degree of mastery in performing these processes, should sometimes be viewed as a *basis* for understanding such concepts rather than as its *outcome*." (Sfard 1991 p.10). We claim that the mastery of the processes related to a given concept—and hence the related representations and operators—is part of its understanding which cannot *and* must not be separated from other dimensions more often mentioned (like the capacity to formulate a definition, or to "recognize" it in a situation).

The last dimension, the control structure, is often left implicit although it is an essential element of the understanding of a concept. After Polya and a long tradition of research on meta-cognition, Schoenfeld (1985 pp. 97-143) has shown the crucial role of control in problem-solving. Robert (1993) emphasised the role of meta-knowledge, demonstrating the benefit from treating control structures as such. The control structure will have an hypothetical character, possibly more than the other dimensions, since it can be identified to the organised set of criteria which allow to decide whether an action is relevant or not, or that a problem is solved.

It is important to insist on the fact that the characterisation of a conception by the above quadruplet is not more related to the subject than to the milieu with which he or she interacts. On the contrary, it allows a characterisation of the subject/milieu system: the representation system allows the formulation and the use of the operators by the active sender (the subject) as well as the reactive receiver (the milieu). The control structure allows to express the means of the subject to decide of the adequacy and validity of an action, as well as the criteria of the milieu for selecting a feedback.

The proposed model allows not only to characterise conceptions and hence provides a framework to discuss their diagnostic, it has also as a first result the potential of helping to more precisely establish a link between conception.

Let us consider two conceptions: $C=(P, R, L, \Sigma)$ et $C'=(P', R', L', \Sigma')$

Generality

[C is more general than C'] if it exists a function of representation $f: L'\rightarrow L$, so that for all problems p from P' then $f(p)\in P$.
In other words, problems from the sphere of practice of C can be understood by C' and enter in its own sphere of practice.

Falsity

[C is false from the point of view of C'] if it exists a function of representation $f: L\rightarrow L'$, and it exists [$p\in P$, $r\in R$, $\sigma\in\Sigma$ and $\sigma'\in\Sigma'$] so that $\sigma(r(p))$=true and $\sigma'(f(r(p))$=false
In other words, p has a solution from the point of view of C which is false from the point of view of C'.

The notions of "generality" and of "falsity" of a conception get a more precise meaning in the model. First, they are not properties of a conception but relations between two conceptions. Second, they need the hypothesis of the existence of a kind of translation from one system of representation to an other.

The fact that one speaks of a misconception, or that one states that a conception is local, has it origin in the existence of the conception of the speaker who, more often implicitly, considers oneself as a privileged knower able to judge and evaluate others' knowings. This common reality lead us to propose to have available in the model a kind of conception of reference which we will here define first and then comment.

But before that, we need to tell more about problems, problem-solving and links between problems and conceptions.

2.1.2.2 Problems and conceptions

We have introduced problems as means to characterise a conception. Indeed, most problems are not specifically related to one single conception, but are related to a set (or several sets) of conceptions which contribute to its resolution.

We will provide a characterisation of the relation between problems and conceptions which goes beyond the description of a detailed solution, involving operators and controls.

Let p be a problem, and {C1,…, Cn} a family of conceptions. We will tell that {C1,…, Cn} solves p iff it exists a sequence of operators (r_{i1}, …, r_{im}) whose terms are taken in one of the Ri for i in {1,…,n} so that:

$p_1=r_{i1}(p)$

for k between 2 and m: $p_k=r_{ik}(p_{ik-1})$

it exists s from Σ_{im} so that $s(p_{im})$= solved.

Classically, one says that (r_{i1}, …, r_{im}) is a solution of p.

Elementary problem

A problem p is elementary for a conception C if {C} solves p

Problem specific to a conception

A problem p is specific to a conception C, if C is an element of any set of conception solving C.

An interesting characterisation of a conception C, although not minimal, would be to define P as the set of all the problems specific to C.

Equivalent conceptions on a set of problems

Two conceptions C and C' are told equivalent on a set E of problems iff, for all p from E, any substitution of C by C', and reciprocally, in a set of conceptions which solves p is still a set of conceptions which solves p.

Equivalent conceptions

If P=P'=E, we will just say that C and C' are equivalent.

Partial conception

A conception C is said partial with respect to a conception C' iff C and C' are equivalent

Conceptions of a different nature on E

Two conceptions C and C' are of a different nature on a set of problems E iff, for all p in E: it exists a set of conceptions which solves p and contains C, and one which solves p and contains C', but the substitution of C to C', and reciprocally, in a set of conceptions which solves p is no longer a set of conceptions which solves p.

Conceptions of a different nature

If P=P'=E, we will just say that C and C' are of a different nature.

Conceptual field

The notion of conceptual field introduced by Vergnaud (1991 p.146) can be expressed in the model: the conceptual field of a conception C is the set of problems which solutions involves C.

Object of a conception

Let us consider two different conceptions C and C', a third conception C_a more general than C and C' and functions of representation f: $L \rightarrow L_a$ and f': $L' \rightarrow L_a$. If for all p from P it exists p' from P' such that $f(p)=f'(p')$, and reciprocally, then C and C' will be said as having the same object with respect to C_a.

The fact that two conceptions have the same object is, eventually, the point of view of a third conception (let say that of an observer) and depends heavily on the existence of functions of representation — which may be seen as a translator or interpreter — and does not imply that these conceptions have another type of relationship (one being false with respect to the other, of more general, or partial, or else).

To have the same object with respect to a conception C_a is an equivalence relation, Ca is the reference conception for the object it allows to define.

2.1.2.2 Conception, Knowing, Concept

We call *knowing* a set of conceptions which have the same object with respect to a given conception C_μ — which is not necessarily an element of this knowing (it can be seen as the point of view of an observer). Hence, a knowing with respect to C_μ is a subset of the equivalence class it defines. From the characterisation we gave above, this does not tell more about the kind of relationships the conceptions constituting a knowing share.

We may speak of the domain of validity of a knowing (i.e. the union of the domain of validity of the related conceptions) but at the same time we can acknowledge its possible contradictory character if one of its conceptions is false from the point of view of another one.

Let $C\mu$ a conception defining a given µ-object, that is the more general conception of the class of equivalence of the conceptions which have the same object with respect to $C\mu$. All the subsets of this equivalence class are knowings; knowings of the same µ-object.

We call *concept* the set of the knowings of the same object. It should be emphasized that a concept is set of sets, but it can be seen also through the set of conceptions obtained by the union of the set of conceptions which constitute knowings.

In the case of function (Balacheff and Gaudin 2002), for example, investigating its meaning consists first in constructing a schema like the following and understanding the relationships at the level of the conception (is it possible for example to provoke the passage from one conception to another?), the knowing and questioning the concept level.

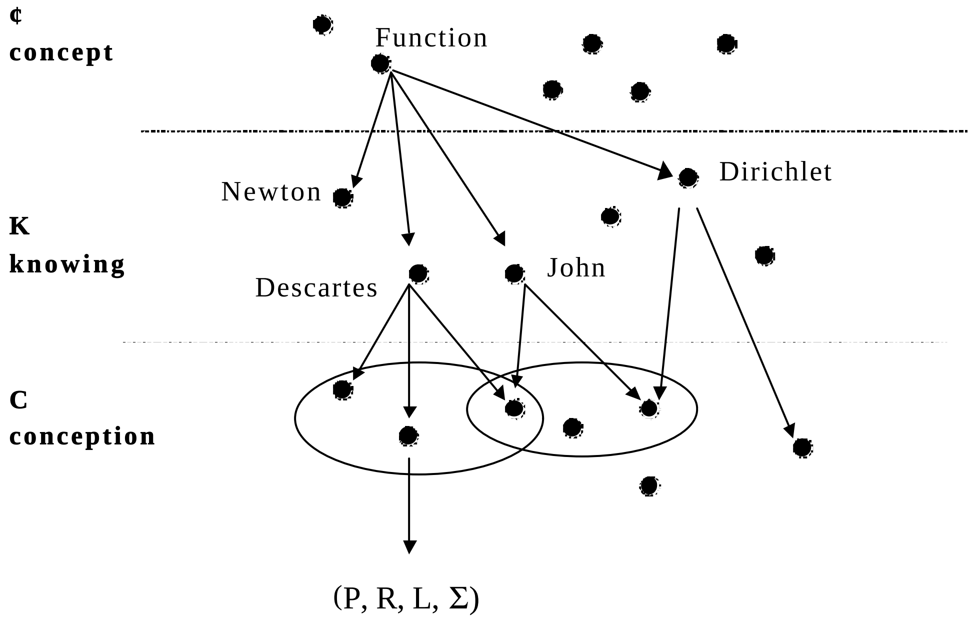


Figure 2. The three levels of the cK¢ model

To summarise: conception is the instantiation of the knowing ascribed to a subject by a situation (it characterises the *subject/milieu system in a situation*), or it could be considered as the instantiation of a concept by a the pair (subject/situation). From the relationships between conceptions induced by the definition adopted here, and from their properties, we can draw in a natural way properties and relationships between knowings, as well as between concepts. Finally, we come back to the knowing subject a set of conceptions which use is stimulated by a particular milieu.

We have considered here the milieu, as well as the interactions between the subject and the milieu from the unique point of view of adidactical situations, the teacher is not considered in this model. The teacher must be considered as soon as we consider the conditions for a given student to come to an interaction with a milieu which we consider likely to allow some learning.

2.1.3 Learning

Learning is a process which allows to pass from a conception to an other one.

In order to model learning within the cK¢ framework, we need to have the notion of domain in order to organise the set of the conceptions, and within which we can express the question of evolution.

2.1.3.1 Domain of reference and problématique

We call domain D of reference related to a set of concepts, a subset of the knowings defining these concepts. We attach to D a conception C_d=(P, R, £, Σ) which we claim to be more general than any of the conceptions which are elements of the knowings participating in its characterisation.

To simplify the discourse we will use the expression "concept of D" for a concept of which a knowing is in D.

Let p be a problem, we call fundamental statement of p following D, the representation £(p) of p with respect to C_d. For a given conception C of a concept related to D, it exists a function f: £→L which allows to have a representation of p with respect to C, that is f(£(p)).

Two problems from two different conceptions are isomorphic if they have the same fundamental statement.

Let ¢ be a concept from D, we call problématique of ¢ the set ∏¢ of the fundamental statements of the problems elements of the conceptions constituting the knowings defining ¢.

The problématique of a knowing K will be the set of the problems from ∏¢ which have a representation for at least one of the conception defining K.

In the following, we call problématique ∏ of a domain D, the union of the problématique of the elements of D.

2.1.3.2 Subject's states of knowing

For a while, we will focus on one of the constituent of the [S<->M] system: the subject S. Our aim is to consider the state of knowing of S, with reference to a domain D.

Let {¢1, … , ¢n} be a set of concepts from D, a state of knowing of S would be intuitively the set {K1, … , Km} of the knowings such that for any i∈{1, m} it exists j∈{1, n} so that Ki is a defining element of ¢j.

But if we take a pragmatic approach, the state of knowing of a subject will always be derived from a diagnosis of his or her conceptions related to some problem-solving competency. Hence, we prefer the following characterisation:

Let Qs be a subset of the conceptions, elements of the knowings which are defining concepts from {¢1, … , ¢n}, πs be a subset of the problématique ∏ of D such that any problem p from πs has a representation f(p) with respect to at least one of the elements of Qs, and for which it exists a family of conceptions elements of Qs which solves f(p). We call state of knowing of S with respect to {¢1, … , ¢n} the couple (Qs, πs).

2.1.3.3 State of knowing of a D-object

Let us consider the set of conceptions constituting the knowings defining a domain D. The equivalence relation "to have the same object with respect to Cd" allows a partition of D in what we will call D-objects.

Le M be a D-object we will say that an element of its class of equivalence is a conception of M. We can in a systematic way enumerate the state of knowing related to M:

Let π be the set of problems from ∏ which have at least one representation with respect to a conception of M. Let πe⊂π and Qe⊂Q: if all element of πe can be represented in Qe and has a solution in Qe, then (Qe, πe) is a state of knowing of M.

It should be noted that this modelling does not account of the psychological plausibility of such a definition. It may be the case that some of the states of knowing will not correspond to a "real" state of knowing as it might be experimentally diagnosed (that is, ascribed to a "real" learner). Understanding why is by itself interesting. In fact, we are exactly there in the position of the didactical a priori analysis to which the model provides a possible tool.

The model provides a natural way to organise the states of knowing with respect to a given D-object M:

Let two states of knowing of M be E1=(Qe1, πe1), E2=(Qe2, πe2). One will tell that E2 is a progress with respect to E1 if any of the following is satisfied:

- πe1⊂πe2, that is if E2 can participate in the solution of more problems than E1
- E2 is more efficient than E1 (following a cost function to be defined)
- If πe2=πe1 and E2 is as efficient as E1, the progress will be to have a more homogeneous state, that is: |Qe2|<|Qe1| (optimal forgetting)

2.1.3.4 Properties of a state of knowing

A state of knowing (Qe, πe) is incoherent if it contains two contradicting conceptions (that is, each one is false from the point of view of the other) with respect to at least one problem from πe.

A state of knowing is potentially conflictual is it exists a problem from ∏ which has a representation for at least one element of Qe, and such that (Qe, πe∪{*f*(p)}) is incoherent.

Just as it has been done for conceptions, we can here define states of knowing which are equivalent, of a different nature, partial or false with respect to an other state.

2.2 Didactical principles and model specifications

2.2.1 EVOLUTION OF THE STATES OF KNOWING

Learning, in the approach we present, is a process of reconstruction of an equilibrium of the subject/milieu system which has been lost following perturbations of the milieu, or perturbations of the constraints on the system, or even perturbations of the subject itself (modification of his or her intentions, or as a consequence of a brain disease, etc.). Indeed, overcoming perturbation, in other words solving problems, does not always lead to learning in a strict sense. It may simply reinforce existing conceptions. The situations to stimulate learning must be carefully designed, so that they provoke perturbations likely to call for significant readjustments or new constructions.

The didactical *problématique* considers these perturbations, provoked on purpose, with the intention to stimulate learning. The indicator of a perturbation is *the gap recognised by the subject* between the expected result of an action and the actual feedback from the milieu. This means on one hand that the subject is able to recognise the existence of a gap not acceptable with reference to her intention, and on the other hand that the milieu can provide an identifiable feedback.

The role of the teacher, with respect to a given content to be taught, is to organise the encounter between a subject and a milieu so that a knowing — which can be seen as acceptable with respect to the didactical intention — can emerge from their interaction. Such an encounter is not a trivial event. To be in an environment is not enough for the student to be able to "read" in it the milieu relevant to the teaching purpose. To select the relevant features of the environment, to identify the feedback and to understand it with respect to the intended target of action is not self-evident. The means for the teacher to succeed in this task is to construct a *situation*, which allows the devolution to the students of both the milieu and the relevant relationships (action/feedback) to this milieu. But the didactical intention of such a situation can act as a constraint; this is the case when the student believes in a teacher expectation, that could modify the nature of the subject/milieu system equilibrium and then the nature of the related knowing. This is the basic complexity of didactical systems.

In short, teaching/learning situations are under three main categories of constraints: time constraints, epistemological constraints (Arsac *et al.* 1992) and institutional constraints. The first are due to the way schooling is organised (duration of the school life, organisation of the school year, organisation of the lessons, etc.). Such constraints can be identified for institutional training situations as well. The second is due to the existence of a "knowledge of reference" which underlies any content to be taught and which *de facto* provides criteria to the acceptability of any learning outcome. The third is fundamental to didactical situations as opposed to natural learning situations.

The notion of *didactical contract* has been coined to clarify this complexity; it intends to describe the responsibility of all partners involved in a teaching/learning situation considered from the point of view of the content at stake. The didactical contract "states" (although it is essentially implicit) what each of the partners has the responsibility to do, or what he is responsible for.

- The teacher creates the conditions likely to allow the learning process to occur (devolution of a problem), and witness that the expected learning has "occurred" (institutionalisation).
- The learner has to understand the problem and to find a solution, that is: to learn.

Then there are two points of view, two games, at once: the game of the teacher in charge of organising and allowing to make sense of the game of the student within an environment, the game of the student with the milieu. We call the later adidactical game, or adidactical situation, to acknowledge that we consider that in this situation the student does not try to please the teacher but to satisfy requirement directly related to the nature of the learning content at stake. We capture it in a single formula:

[didactical situation] = [didactical contract]([adidactical situation]).

In the current project we are interested in the case where we have the following formula:

[didactical situation]=[didactical contract]([machine])

That is, when a machine, or in more general cases a system (including resources from the web), takes in charge the interactions to stimulate or guide a learning process in a situation which has been designed by or is the result of the willing of an institution whose representative is a teacher or a trainer.

The cK¢ model intends to cover the requirements for a formalization of an adidactical situation in which the machine provides both the relevant milieu to favour learning and feedback to stimulate or guide the learning process.

2.2.2 EVOLUTION OF THE STATES OF KNOWING

Let us come back to the notion of *state of knowing* (Qe, πe) of a D-object M. A state of knowing can evolve as a consequence of the evolution of the elements of Qe or of πe, or both. A modification of Qe can also come from the addition of a new element, or on the contrary from the elimination of an element, or even the substitution of a conception to a set of conceptions in Q. The set of problems π can be modified in the same way.

Indeed, an interesting evolution is an evolution which leads to the elimination of the potential contradictions in a state of knowing: *a modification of a state of knowing will be considered as learning if it is a progress*. An idea which in the end is not so easy to capture.

A didactical process can be modelled as a transition function on a network of states of knowing which allows (i) to ascribe a state of knowing to a subject based on the observation of the functioning of the [S<->M] system, (ii) to decide on a path within the network which can allow to pass from a state of knowing diagnosed to a chosen state of knowing (object to be taught).

This means that the object of a didactical process are not the conceptions alone, but the pairs (conception, problem), since in the end problems are the key of the diagnosis of a conception, and they are the main tool to stimulate and achieve the intended learning.

Let [C, p] be the current state of the diagnosis and C_t (P,R, L, Σ) the expected learning output. The most critical evolution is the one obtained when the initially diagnosed conception C is false with respect to the targeted conception C_t. For this purpose it is necessary to find a problem which has a representation for both conceptions and which could witness a contradiction: a solution should be conceivable with the diagnosed conception, but not possible to solve correctly with this conception. It may be the case that such a problem does not exist and that intermediary conceptions are necessary to allow the intended learning.

To define a didactical process on a set of conceptions (or states of knowing), requires the specification of a topology related to problem-solving. This problem has been identified but has not yet got a proper answer. A possible way to solved it, would be to use a measure of similarity which should hold the following basic properties:

let C1 and C2 be two conceptions; the key element to appreciate their “distance” should be essentially about the same concept, what requires an evaluation of the weight of the shared problems (that is, of $P_1 \cap P_2$). Then, on this intersection, an indicator of similarity must weight the operators shared on $P_1 \cap P_2$ and such that a solution from the point of view of one conception could be “understood” from the point of view of the other one, with the same evaluation. Here appears the key role played by the representation system in appreciating the similarity between two conceptions. An issue that we have not explored yet, but which necessitates further investigation. When the system of representation is a priori shared, what could be the case when a single learning environment is used, then the similarity could be evaluated only relying on the quantification of the relation between the set of problems.

Problems and conceptions are of a dual nature: on the one hand conceptions needs problems as constituents of there characterization, and on the other hand problems get their meaning from the conceptions which contribute to their solutions and, indeed, from the nature of this solution. As we have seen in the preceding sections, from this can be derived relations between conceptions being mediated by problems and between problems being mediated by conceptions. This is exactly the idea of the Vergnaud’s Conceptual field. The following schema illustrates that.

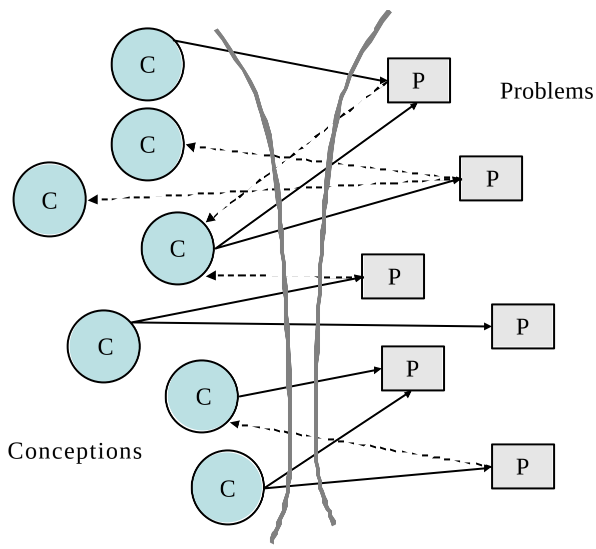


Figure 3. Duality of problems and conceptions

An other sight on this schema evidences that the didactical process consists of a navigation in a bipartite graph, the edges of which hold their semantic from the cK¢ model: conceptions are linked to problems they contribute to solve, and sometimes problems are specific of this conception (entering in its definition); problems are tools to diagnose or to reinforce a diagnosis, to question, to destabilise, or to reinforce conceptions.

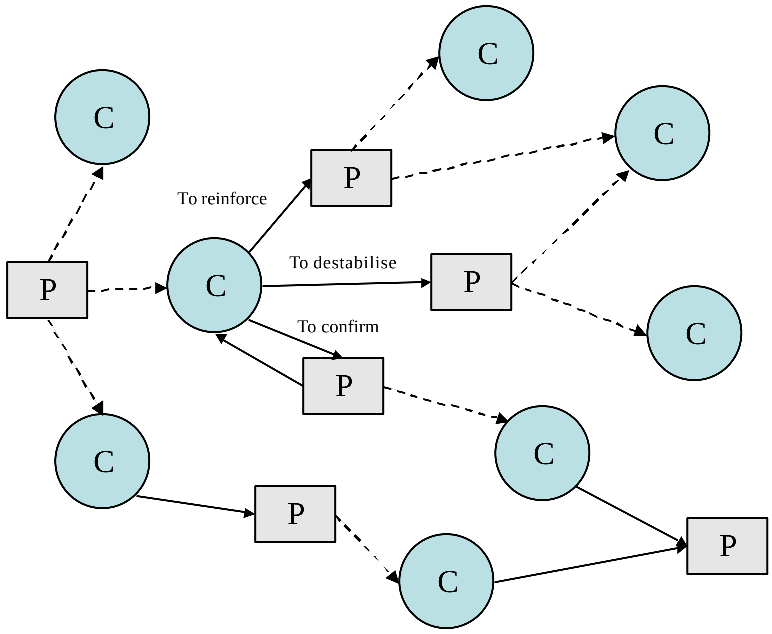


Figure 4. Bipartite graph of conceptions and problems

References

Ambrosio U. d' (1993) *Etnomatemática*. São Paulo: Editora Atica.

Aebli H. (1963) *Didactique psychologique.* Neuchâtel: Delachaux et Niestlé.

Arsac G., Balacheff N., Mante M. (1992) Teacher's role and reproducibility of didactical situations. *Educational Studies in Mathematics* 23 (5) 5-29.

Artigue M. (1991) Épistémologie et didactique. *Recherches en didactique des mathématiques.* 10(2/3) 241-285.

Artigue M. (1992) Functions from an algebraic and graphic point of view: cognitive difficulties and teaching practices. In: Dubinsky E., Harel G. (eds.) *The concept of Function*. (MAA Notes Vol. 25, 109-132). Mathematical Association of America.

Bachelard G. (1938) *La formation de l'esprit scientifique*. Paris: Vrin.

Balacheff N. (1987) Processus de preuves et situations de validation. *Educational Studies in Mathematics* 18(2) 147-176.

Balacheff N. (1995) Conception, connaissance et concept. In: Grenier D. (ed.) *Didactique et technologies cognitives en mathématiques, séminaires 1994-1995* (pp.219-244). Grenoble: Université Joseph Fourier.

Balacheff N. (1995a) Conception, propriété du système sujet/milieu. In: Noirfalise R., Perrin-Glorian M.-J. (eds.) *Actes de la VII° Ecole d'été de didactique des mathématiques* (pp.215-229). Clermont-Ferrand: IREM de Clermont-Ferrand.

Balacheff N. (1998) Construction of meaning and teacher control of learning. In: Tinsley D. J., Johnson D. C. (eds.) *Information and Communication Technologies in School Mathematics* (pp. 111-120). Chapman & Hall.

Balacheff N., Gaudin N., (2002) Students conceptions: an introduction to a formal characterization, *Les Cahiers du Laboratoire Leibniz n°65.* http://www-leibniz.imag.fr/LesCahiers/Cahiers2002.html. In Annex IV of this report.

Bourdieu P. (1990) *The logic of practice*. Stanford, CA: Stanford University Press [English translation of: *Le sens pratique*. Paris: Les éditions de Minuit. 1980]

Breidenbach D., Dubinsky, Hawks J., Nichols D. (1992) Development of the process conception of function. *Educational Studies in Mathematics* 23, 247-285.

Brousseau G. (1997) *Theory of didactical situations in mathematics.* Dordrecht: Kluwer Academic Publishers.

Castela C. (1995) Apprendre avec et contre ses connaissances antérieures. *Recherches en didactique des mathématiques.* 15(1) 7-47.

Confrey J. (1986) “Misconceptions” accross subject matters: charting the course from a constructivist perspective. *Annual meeting of the American Educational Research Association.* (unpublished manuscript).

Confrey J. (1990) A review of the research on students conceptions in mathematics, science, and programming. In: Courtney C. (ed.) *Review of research in education*. American Educational Research Association 16, pp. 3-56.

Dhombres J. (1988) Un Texte d'Euler sur les Fonctions Continues et les Fonctions Discontinues, Véritable Programme d'Organisation de l'Analyse au *18ième Siècle. Cahier du Séminaire d'Histoire des Mathématiques, Université Pierre et Marie Curie, Paris.*

Douady R. (1985) The interplay between different settings. Tool-object dialectic in the extension of mathematical ability. In: Streefland L. (ed.) *Proceedings of the ninth international conference for the psychology of mathematics education* (Vol. 2, pp. 33-52). Utrecht: State University of Utrecht.

Edwards C. H. Jr. (1979) *The historical development of calculus*. Berlin: Springer-Verlag.

Gaudin N. (2001) *Conceptions de fonction et registres de représentation, étude de cas au lycée* (submitted)

Glasersfeld E. von (1984) An introduction to radical constructivism. In: Watzlawick P. (ed.) *The invented reality* (pp. 17-40). New York: Norton.

Kleiner I. (1989) Evolution of the function concept: a brief survey. *The College Mathematics Journal* 20(4) 282-300.

Kline M. (1972) *Mathematical Thought from Ancient to Modern Times*. New York: Oxford University Press.

Lave J. (1988) *Cognition into practice.* Cambridge: Cambridge University Press.

Mesa V. M. (2001) Prototypical uses of function present in seventh- and eight- grade textbooks from fifteen countries. In: van den Heuvel-Panhuizen M. (ed.) *Proceedings of the 25th conference of the international group for the Psychology of Mathematics Education* (Vol. 3 pp. 367-374). Utrecht, state University of Utrecht.

Monna A. F. (1972), The Concept of Function in the 19th and 20th Centuries, in Particular with Regard to Discussions between Baire, Borel and Lebesgue. *Archive for History of Exact Sciences* 9 (1) 57-84

Nuñes T., Carraher D., Schliemann A. (1983) *Mathematics in streets and schools*. Cambridge: Cambridge University Press.

Pichot A. (1994) Pour une approche naturaliste de la connaissance. *Lekton* 4(2) 199-241.

Rabardel P. (1995) Qu’est-ce qu’un instrument ? *Les dossiers de l’Ingénierie éducative.* 19, 61-65.

Resnick L., Collins A. (1994) *Cognition and Learning*. Pre-print. Learning Research and Development Center. University of Pittsburgh.

Robert A. (1992) Problèmes méthodologiques en didactique des mathématiques. *Recherches en didactique des mathématiques.* 12(1) 33-58.

Robert A. (1993) Présentation du point de vue de la didactique des mathématiques sur les métaconnaissances. In: Baron M., Robert A. (eds.) *Métaconnaissances en IA, en EIAO et en didactique des mathématiques.* RR LAFORIA 93/18. (pp.5-18). Paris: Institut Blaise Pascal.

Salin M.-H. (1976) *Le rôle de l’erreur dans l’apprentissage des mathématiques de l’école primaire.* IREM de Bordeaux.

Sfard A. (1991) On the dual nature of mathematical conceptions: reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics* 22, 1-36.

Shoenfeld A. (1985) *Mathematical Problem Solving*. Orlando: Academic Press.

Shoenfeld A. (ed) (1987) *Cognitive Science and Mathematics Education*. Hillsdale: Lawrence Erlbaum Associates.

Sierpinska A. (1989) *On 15-17 years old students’ conceptions of functions, iteration of functions and attractive fixed points*. Institut de Mathématiques, preprint 454. Varsovie: Académie des Sciences de Pologne.

Smith D. E. (1958) *History of mathematics*. (Vol. II, esp. chap X). New York: Dover Publications Inc.

Stewart J. (1994) un système cognitif sans neurones: les capacités d’adaptation, d’apprentissage et de mémoire du système immunitaire. *Intellectika* 18, 15-43.

Tall D. (1996) Functions and calculus. In: Bishop A. *et al.* (eds.) *International Handbook of Mathematics Education* (pp. 289-326). Dordrecht: Kluwer Academic Publishers.

Thurston W. P. (1994) On proof and progress in mathematics. *Bulletin of the American Mathematical Society* 30(2) 161-177.

Vergnaud G. (1981) Quelques orientations théoriques et méthodologiques des recherches françaises en didactique des mathématiques. *Recherches en didactique des mathématiques.* 2(2) 215-231.

Vergnaud G. (1991) La théorie des champs conceptuels. *Recherches en didactique des mathématiques.* 10(2/3) 133-169.

Vinner S. (1983) Concept definition, concept image and the notion of function. *International Journal of Mathematical Education in Science and Technology* 14, 293-305

Vinner S. (1987) Continuous functions - images and reasoning in college students. In: Bergeron J. C., Herscovics N., Kieran C. (eds.) *Proceedings of the Eleventh International Conference for the Psychology of Mathematics Education* (Vol. 3 pp. 177-183). Montréal, Canada: Université de Montréal.

Vinner S. (1992) The function concept as a prototype for problems in mathematics education. In: Dubinsky E., Harel G. (eds.) *The concept of Function*. (MAA Notes Vol. 25, 195-213). Mathematical Association of America.

Vinner S., Dreyfus T. (1989) Images and definition for the concept of function. *Educational Studies in Mathematics* 20(4) 356-366.

Youschkevitch A. P.(1976), The Concept of Function up to the Middle of the19th Century. *Archives for History of Exact Sciences* 16 (1) 37-85.

3 Multi-agents and emergence, targeted architecture

Leader of the task: Jean-Pierre Muller

Authors: Jean-Pierre Muller, Michel Occello

The aim of this section 3 is to propose a target architecture rooted in the notion of emergence. Accordingly, we shall fully discuss the notion of emergence and this discussion shall result in an operational definition which will be concretised in a design methodology. This is done in Annex II of the report. Applying the obtained methodology to the notion of learning and conception described in task 1.1 (section 2 of the report), we will provide a target architecture whose project mock-up (task 2.1, section 4) is a partial instance to validate the approach.

3.1 Towards a definition of emergence in multi-agents systems

Most of the discussion about the notion of emergence come from the work of a special interest group called « COLLINE » which stands for « COLLective, Interaction, Emergence », part of the Multi-agent System chapter of AFIA « Association Française d'Intelligence Artificielle » (Artificial Intelligence French Association). The first aim was to study the phenomena occurring in collective entities, either natural (human groups, ants colony, etc.), or artificial (soccer robot teams, software agents, etc.). Studying these phenomena quickly came to the notion of emergence which is more and more widely used. Originally used by the philosophers, the biologists and the physicists to characterize that something comes from something else without being directly produced by a causal effect, most scientific domains invoke this notion. Referring to the emergent properties of nature contrasts with the usual reductionist approach. These emergent properties are commonly invoked by the saying "the whole is more than the sum of its parts". Such an evolution also gained the attention of computer science in particular in the domain of Distributed Artificial Intelligence and Multi-Agent Systems (MAS). In effect, it is generally postulated than in the MAS domain, the complex collective activity observed at the level of the whole system results from the interactions of simpler activities at the agent level. Emergence being potentially a way to articulate the activities at the "micro-level" with the activity at the "macro-level", it is becoming interesting for designing artificial societies.

Features of emergence. A definition.

As exposed in Annex II, the notion of emergence is multiple defined rather than ill-defined and covers two aspects: the static aspect (emergence as a result or observable) and the dynamic aspect (emergence as a process). It seems essential for the Baghera project to come up with a positive, temporal (where time appears explicitly) and constructive definition of emergence.

A definition initiated by Ch. Lenay (Lenay 96) proposes this positive and dynamical approach of emergence. To begin, it is important to distinguish some characteristics of emergent systems implying the subject and its environment and where emergence results from the interaction between both.

- The first essential feature of a multi-agent system is that no agent controls entirely the dynamics of the population. The agents are *limited* and there are differences of a global system they are unaware of. Therefore, there is an exteriority relative to each agent: an *environment*.

- The second feature is that, by definition, the agents act and therefore modify this environment. But, the agents can only perceive and act locally in this environment. Said otherwise, each agent *interprets* the environment given his limited means (using the distinctions it is able to make).
- The third feature is that the exterior of each agent contains other agents. There are several agents in a common environment (they are exterior to each other). The interpretation of the environment by the various agents can possibly be different. In the case of reactive agents, the environment contains the objects and the other agents. In the case of cognitive agents, the environment can also contain messages.

Therefore, the dynamics proceeds by iteration of interpretation of the local environment by the agents, action of the agents on this environment, new interpretation of the modified environment, new actions, etc.

When such a dynamics (or some of its components) stabilizes, we can talk of *emergence* of a structure or of a global functionality. Notice that at any moment, it is the environment possibly modified by all the other agents (and itself) that each agent submits to its interpretation. It is the condition for the global dynamics to by more than a simple sum of independent dynamics. If it was the case, we could only speak of emergence in a very weak sense and only for an external observer. But whenever, through the environment, the whole feedbacks on the parts, there is emergence in a strong sense, emergence *for the agents* if it is the global and stable emergent state which selects the individuals behaviours of each agent.

In this definition, the dynamics of interaction is postulated as a basic condition or emergence of phenomena, structures, etc.. Also notice the importance of the link whole-parts which characterises the various kinds of emergent phenomena. In the following, we will derive a more operational definition by characterising the whole and the parts and most importantly the feedback whole-parts. This definition is inspired by S. Forrest (Forest 1990) and M. Bunge (Bunge 1977) and postulates:

- A system of entities in interaction whose expression of the states and dynamics is made in a vocabulary or theory D;
- The production of a phenomenon which could be a process, a stable state, or an invariant which is necessarily global regarding the system of entities;
- The observation of this global phenomenon either by an observer or by the entities themselves.

This observation can only be done through an inscription of the phenomenon on one hand and on the other hand the interpretation of this inscription by the observer or by the entities themselves in a vocabulary or a theory D' distinct from D. A theory of emergence would be a theory D0 of inscription by a system of entities in interaction and its interpretation.

This definition distinguishes precisely two levels: a micro level (the one of the interacting entities) and a macro level (the set of entities). From this set of entities, one considers a global production (stable state) from this global set.

Another issue raised by this definition is the problem of the interpretation of the inscriptions which provides two possible meaning to emergent phenomena. A weak sense where the inscriptions refer to a same reality understood at two levels that it is useful, or even necessary to distinguish. The problem is summarised as "another way to talk about things" which would be simpler as the temperature regarding the kinetic energy of the molecules. Finally notice that this emergence can be linked to the ignorance of the observer or its inability to formalize an underlying compositionality. Without assuming a two radical dualism, we can consider emergence in the strong sense if the inscription which is globally produced does not have the same order of reality than the individual productions.

How to use these definitions? Or how to apprehend emergence in Multi-agent Systems? The natural approach would be to situate it with respect to emergence observed elsewhere. In effect, the classical approach of modelling and explaining emergence goes from the observation of natural phenomena to the reproduction with artificial systems. In certain cases, modelling is restrained to an *interpretative* attitude of the observed phenomena without being able to validate the proposed hypotheses. It is the case for some collective behaviours observed in ants societies (altruism,…) or in human societies. The models in Artificial Life attempt to reproduce these hypotheses and their conditions of emergence of some phenomena in simulation, that allows to get a better understanding of these phenomena (but uses strong presuppositions on the sense of the term emergence). The extension of the obtained results allows then to adopt a predictive attitude by creating the conditions of emergence of new artificial phenomena by simulation or by experimenting (where there are other technological constraints). It is the case of collective robotics where the specification and the combination (sometimes random) of basic behaviours of a set of robots, can produce emergent global behaviours. It is also the case in programming when it is the only way to solve the problem of trivial compositionality: all what is made in computer science is necessarily the combination of the execution of elementary instructions. But talking of computer science, is to go further than physical processes. The emergentist approach provides a way to go further. The relationship between emergence elsewhere and emergence in computer science goes through inscriptions.

Finally, it is important to notice that the justification of a global production (i.e. the behaviour of the system) as emergent is not only its adequacy to the definition, but also its subordination to the existence of an emergence in the strong sense.

3.2 Designing an emergentist multi-agent system for Baghera

The goal of this part is to use the design methodology proposed in Annex II in order to derive the target architecture of a multi-agent system producing learning by emergence. We will first shortly review the principles of Baghera. We will then proposed a design methodology and finally develop the interpretation process and the generation process.

Two complementary principles in Baghera

Let us remind the principles underlying the Baghera project:

- Human knowing will be interpreted as consisting of a diversity of conceptions whose basic criteria of relevance is not their conformity to a reference knowledge but their efficiency in specific spheres of practice;
- The learning process is an emergent phenomena from the interactions between artificial and human agents.

The last principle states that we have to design a multi-agent system made of artificial and human agents (including learners and teachers) whose interaction will produce learning as an emergent phenomenon. It means that we have to design the interactions between the agents (the micro-level) such that the trace of what is going on globally can be interpreted by an external observer (or an agent within the system) as being learning. Therefore we have to find a theory D at the micro-level producing by emergence a theory D' of learning. This last theory is exactly what the cK¢ theory describes and is related to the interpretation of what a problem solver does as the invocation of one or more conceptions. It is the first principle underlying the Baghera project.

3.2.1 METHODOLOGY

Applying the methodology of emergentist multi-agent systems (cf. Annex II), the design of a emergentist SMA for learning must go through the following steps:

- Define the set of possible global processes (macro-level)
- Derive the interacting entities producing these processes (micro-level), the actual process will depend on the interaction with the environment/user
- Validate the micro-level against the macro-level expectations

In the learning context:

- The macro-level is the learning process itself;
- The micro-level is the interaction of a student (S) with the computer (M) which is itself a set of agents, some of them being avatars of other human agents (other students or teachers);
- The relationship is the interpretation of the trace produced by the micro-level as the trace of a learning process.

It raises a number of questions to be, at least partially, answered in the following sections:

- What is the trace? Which marks in the environment shared by both the student and the computer will constitute the basis of an interpretation process?
- How to interpret the trace as a learning process, building the observer macro-level?
- How to produce the trace? This amounts to find out the set of agents whose interactions with the student will produce the trace.
- How to use the interpretation of the past trace to change the future trace, producing strong emergence?

The first set of questions requires the exhaustive specification of the interactions to occur between the student and the machine because it is this interaction that will be interpreted as the occurrence of learning.

The second question is much more difficult in the case of learning than in any of the other cases we have described earlier. In effect, what has been built in the head of the student is inaccessible and can only be assessed by its performance. In fact, it is nothing but one of the main question of didactical sciences. In our case, we will rely on the cK¢ theory.

The third question will allow us to specify the target architecture of the multi-agent system. In fact the interaction is by definition a reciprocal interaction, from the student to the machine and from the machine to the student. Only the last part has to be taken into account because we do not have (and cannot) program the student!

Finally most of the didactical guidance in the learning process is based on the interpretation of what is going on between the student and the teaching system in the interaction process. Therefore, it is clear that strong emergence makes sense in this interpretation (again in contrast with earlier examples where only weak emergence was necessary).

After introducing what can count as a trace, we will first discuss the interpretation of the trace before discussing the production of the trace, hence the target architecture, including strong emergence.

3.2.2 THE TRACE

Basically the trace of the interaction between the student and the machine is a sequence of events $\{e_1,\ldots,e_n\}$ where e_i is either:

- A problem which is given to the student and therefore is generated by the machine M;
- A problem solving step or control which is proposed by the student S;
- An advice which can be a problem solving step or a set of these proposed by M;

- An answer which determines the end of a problem solving process from S;
- A judgement which can be an answer to the propositions of the student generated by M at any moment of the problem solving process either to suggest that the student is the right or wrong direction or to assess the answer;
- Other events can freely be added to this list.

In order to cover the trace of a complete learning process, we have to consider a sequence in which a set of problems are proposed to the student until the acquisition of given reference conceptions is assessed.

3.2.3 THE INTERPRETATION OF THE TRACE

For interpreting the trace of the interaction between the student and the machine, we will use the following definitions:

- Learning is defined as the acquisition by the student of a *reference knowledge*, which is assessed by the use of one or several *reference conceptions* related to the given knowledge;
- Using a *conception* is itself assessed by the invocation by the student of *problem solving steps* and *controls* for solving *problems* which are provably representative of the given conception.

The problem solving steps, the controls and the proposed problems being the only observables of the system, the goal is to deduce which conceptions are invoked by the student and consequently whether its successively invoked conceptions are going toward the required conceptions. We will interpret learning as going from a conception (or a set of conceptions) to a conception (or a set of conceptions) of reference related to a knowing.

Therefore, we have to distinguish two levels of interpretation:

- the one going from the trace to the invocation of conceptions
- the one going from the trace of the successive invocation of conceptions to the description as a learning process.

This interpretation can be used by an external observer (either a teacher or a didactician) in order to assess the functioning of the multi-agent system. It would be a case of weak emergence. Or it can be fed back into the agents in order to guide the interaction with the student in which case we would have a case of strong emergence because the interpretation of what is going on is retro-acting on the process generation itself either trough human agents or artificial agents.

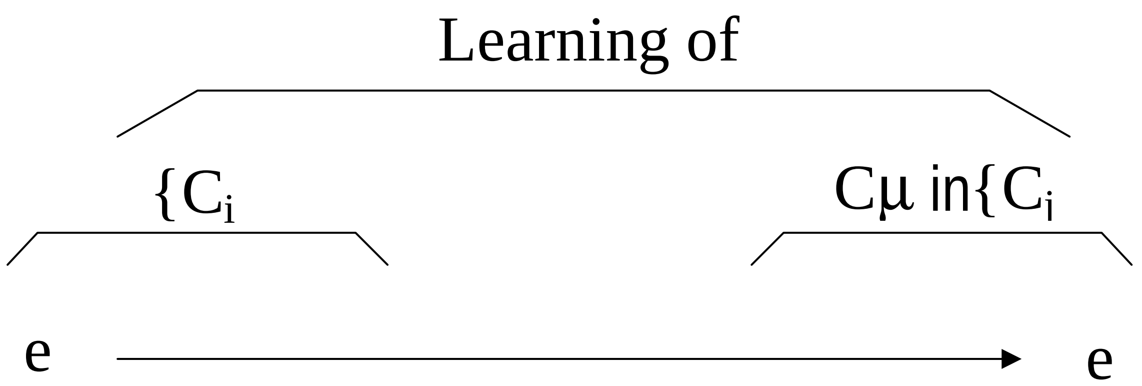


Figure 5. Trace and its interpretation.

Recognizing conceptions in the trace

The recognition of conceptions in the interaction trace can be made in two ways:

- From a catalogue of pre-defined conception
  - A conception is a structure <P,R,L,C> where P is a set of problems, R a set of operators, L the language used and C a set of controls;
  - Each instance of elements of P, R, L or C suggests the use of a given conception;
  - Given the observed trace, one has to arbitrate between a set possible conceptions (ranking, voting);
- From an intrinsic definition of conception
  - One possibility is to characterize P,R,L,C such that there is some logical coherence between rules and with controls, or that chains of inference can be obtained by the rules;
  - Logical coherence criteria could possibly use the services of ATINF.

The advantage of the first solution is that there exist a catalogue of known conceptions in the didactic literature from which such pred-defined conceptions could be (and have been in this project) extracted. But there are two important disadvantages:

- Building an exhaustive is a very heavy if not impossible task and this work would have to be done in each area of knowledge;
- Students use a huge diversity of conceptions depending on their culture and within a culture of the kind of intuitive metaphors they are using to help them conduct the solving process.

Consequently, using an intrinsic definition of a conception would be quite useful but at this stage only weak characterisations like logical coherence actually exist. One solution consists in letting the system propose possible conceptions and to have them validated and completed by the didacticians while running the system. This method would also help capitalize the known conceptions in the catalogue in the course of using the system.

In this project the interpretation of the trace as an invocation of conceptions has been itself implemented as an emergent computation where the interpretation emerges from the interaction between candidate conceptions.

Recognizing learning

From the preceding section, the interpretation process in terms of conceptions is actually levering the trace from a sequence of low level events to successive or concomitant invocation of conceptions. The problem becomes to interpret such a sequence of conceptions as a learning process. Intuitively, we would like to say that the student is making progress or is blocked in a set of conceptions or even is kidding (doing anything).

The suggestion is to use the definition of topologies on conceptions based on various criteria:

- Similarity: when two conceptions relate to the same concept (it would be an equivalence relation);
- Specificity/generalization: when a conception solves subsets/supersets of problems;
- Dependency: when a conception needs the acquisition of another one to be mastered;

This last relation of dependency can be derived in different one, from simple compositionally to acquisition paths.

These topologies organize the conceptions in a space in which the performance of the student can be interpreted as trajectories which can themselves be qualified regarding the conception of reference.

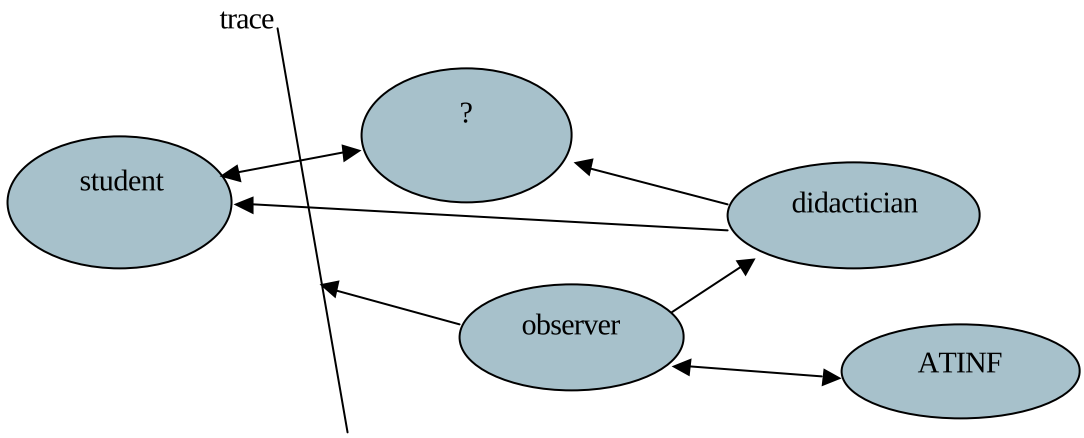


Figure 6.

3.2.4 THE PRODUCTION OF THE TRACE

As a reminder, the trace of the interaction between the student and the machine is a sequence of events $\{e_1,\ldots,e_n\}$ where e_i is either :

- A problem which is given to the student and therefore is generated by the machine M;
- A problem solving step or control which is proposed by the student S;
- An advice which can be a problem solving step or a set of these proposed by M;
- An answer which determines the end of a problem solving process from S;
- A judgement which can be an answer to the propositions of the student generated by M at any moment of the problem solving process either to suggest that the student is the right or wrong direction or to assess the answer;
- Other events can freely be added to this list.

The goal of this part is to identify and to design the agents to order to ensure the production of the events generated by M in such a way that a learning process occurs.

In a first approximation, we can have a bunch of agents responsible of producing the various kinds of event:

- The problem agents to propose to the student new problems to solve;
- The advice agents to give the advices and hints;
- The judgement or diagnosis agents to assess what is going on and formulate it to the student.

Running all these agents in parallel raises the problem of arbitration between the agents which can be done in a centralized way by a didactical agent (or even a set of didactical agents) or in a decentralized way by negotiation between the agents. We will specify these agents in turn. For each agent, we have to specify the goal (otherwise it would not be an agent), the perception and the possible actions. The communication related to the arbitration among the agents will be described on the part on arbitration.

The problem agents

- The goal of a problem agent is to propose the most suitable problem given the context. In order to do that, the agent has to perceive the context which consists of: (i) the invoked conceptions, (ii) the targeted conceptions, (iii) the diagnosis.

The invoked conceptions are provided by the agents in charge of the interpretation of the interaction trace. The target conception can be set by the teacher or a didactical agent has a strategy to guide the student towards other target conceptions. The diagnoses are produced by the diagnoses agents and describe the student trajectory.

The open problems are:

- How to characterize the suitability of a problem in a given context? It is clearly related to didactic decisions. It amounts to define a ranking between different possible actions (i.e. problem propositions).
- How to define the granularity of an agent? An agent could be associated to each problem in which case the competition is directly between the problems. An agent could be associated to a conception in which case the competition is between the conceptions and the suitability is related internally to the best next problem (most specific) and externally to the next best conception to teach to the student. An agent could be associated to a concept putting the problem at the conceptual level (in which concept space to move).

The advice agents

The goal of an advice agent is to propose the most suitable advice given the context. In order to do that, the agent has to perceive the context which consists of:

- Elapsed time when nothing happens for some time;
- Rules coherence, for example of a chain of inference;
- Trajectory in the conception topology (being near something he already knows);
- Diagnoses.

The coherence could be checked by the ATINF agent. The trajectory necessitates an agent in charge of managing the conception topology and representing the history of the student in such a topology (also useful for the diagnoses).

Unlike the previous agent, an open question is the set of possible actions, which could be:

- Counter-examples
- New rules

The diagnosis agents

The goal of a diagnosis agent is to propose the diagnosis in given the context. In order to do that, the agent has to perceive the context which consists of:

- Completeness of a subtrace
- Conception
- Rules

The completeness of a subtrace could be checked by the ATINF agent. Coherence between conceptions can be used to assess the student problem solving process. The rules could also individually checked for validity using ATINF. The possible actions could be:

- Signalling mistakes
- Signalling misconceptions
- Assessing positively or negatively a problem solving step or the final answer to a problem.

These actions raise the problem of characterizing mistakes and misconception which is a research debate per se.

Arbitration

Executing all the above mentioned agents in parallel raises the problem of arbitration among the agents because, problems, advices and judgements cannot be sent to the student simultaneously. This raises the problem of controlling the emergence of a coherent interaction as a whole without the didactic process being stated in advance for better adaptation to the student capabilities.

An extreme would be to have a centralized control where each agent proposes its most suitable action from his point of view to a central didactical agent. The final decision could be based on the trajectory of the student in a global conception/concept topology.

Another extreme would be to let each agent being fed by the interpretation process and negotiating between themselves on the best decision to take.

The open problem is again to find the suitable granularity of the agents. An important guide line is to reflect on the structure of the domain to be teached and in particular its hierarchical structure on which the hierarchy of agents could be mapped. The general problem of designing multi-agent system is to articulate the micro-level level to the macro-level and where the relationships is highly structured a number of intermediate levels can be designed in order to master the complexity. Spontaneous grouping of agents into substructures could also be foreseen.

Finally, one has to take into account the choice of having both human and artificial agents (mediated by software agents). Consequently, the distribution of decision between artificial agents and humans has to be carefully thought out given the synthesis capability of humans, artificial agents can hardly implement.

3.2.5 Resulting architecture

Given the above discussion, the targeted architecture can be the one illustrated in the following figure:

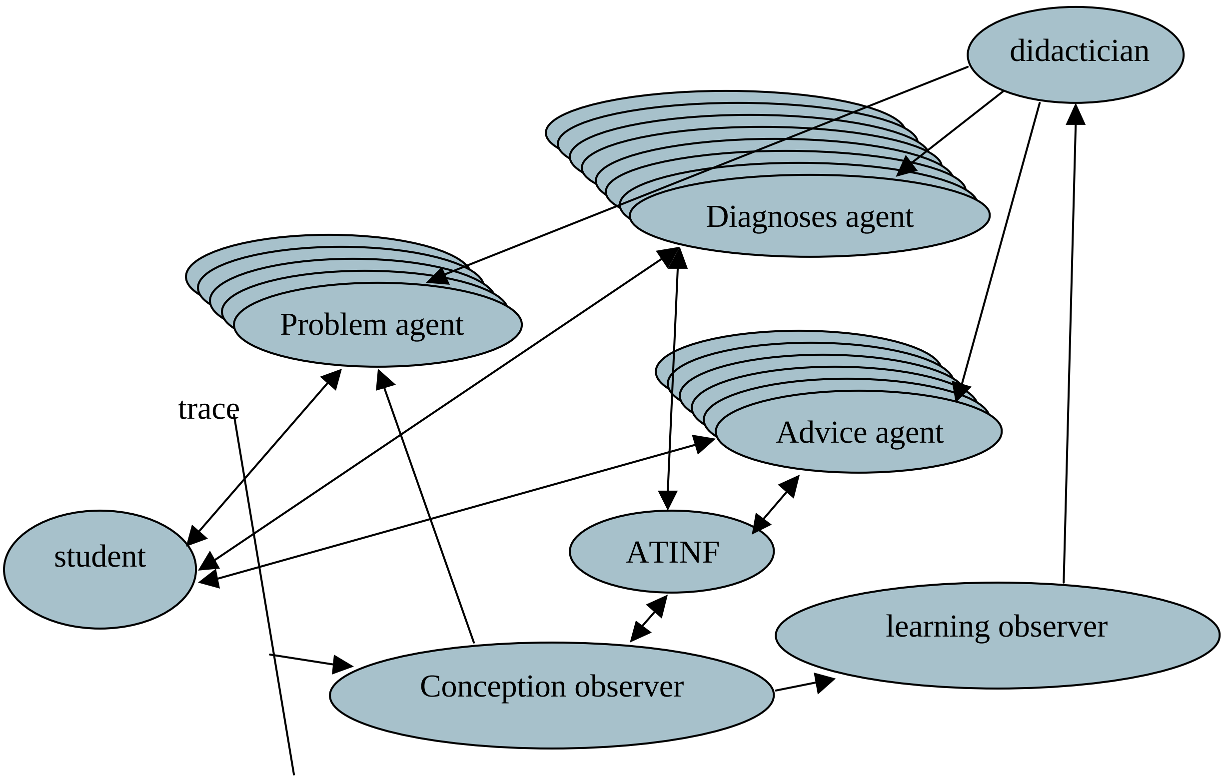


Figure 7. Targeted architecture.

References

BUNGE M., (1977), “Emergence and the mind”, in Neuroscience, 2:501-509.

FORREST S., (1990), “emergent computation: Self-organizing, Collective, and Cooperative Phenomena in Natural and Artificial Computing Networks”, introduction to the proceedings of the ninth annual CNLS Conference, in “emergent computation”, MIT Press, Cambridge, MA, 1990, pp 1-11.

FORREST S., MILLER J.H., (1990), “Emergent behavior in classifier systems”, in “emergent computatio Cambridge, MA.

LENAY C., (1996), “Coopération et intentionalité”, journées françaises sur les systèmes multi-agents, JFIADSMA, Avril 1996, Sète, France, pages 265-272.

4 Design methodology and mock-up

Leader of the task: Sylvie Pesty

Authors: Ricardo Caferra, Nicolas Peltier, Sylvie Pesty, Jana Trgalova, Carine Webber

4.1 Baghera 1.0 Platform

Baghera is a distance learning environment based on a two-level multi-agent architecture. Two-levels are distinguished since the application is composed by two multi-agent systems (MAS). The higher-level MAS is composed of cognitive agents, which provide the main functions of the system. The lower-level MAS is composed of a large number of reactive agents responsible for diagnosing students' conceptions. The two levels communicate through tutor agents, whose educational decisions are based on emergent results coming from the lower level. Besides the multi-agent approach, our work is founded on the emergent theory and employs a mechanism of voting for capturing group decision. During BAP project we have developed the lower-level MAS whereas the higher-level MAS was implemented during a previous project (French National and Regional Projects). This section presents methodological and theoretical aspects of our platform and its general architecture. The section 4.2 presents in detail the lower-level MAS specific to BAP.

4.1.1 INTRODUCTION

Multi-agent methodology has appeared as an alternative to conceive AI-based educational systems. The traditional architectures have proved to be too monolithic to deal with the new expectations of systems that should be able to provide "learning anytime and anywhere". Aspects such as data persistence and mobility become extremely important in the design of this new class of educational systems. Besides, researchers in the educational field have shown that it is not possible to find a general strategy of teaching if we take into account human differences but it is rather probable to think that learning is an emergent result of rich and coherent interactions occurred during time (Balacheff 2000a). The multi-agent methodology can certainly bring several advantages to the development of educational applications since it deals well with applications where such crucial issues (distance, cooperation among different entities and integration of different components of software) are found. As a result, multi-agent approach together with technologies of networking and telecommunications, bring powerful resources to develop educational systems.

Several projects implement learning systems based on multi-agents architectures. Some of them work on a generic platform of agents (Capuano *et al.* 2000, Machado *et al.*1999, Silveira *et al.* 2000, Vassileva *et al.* 2001). For example, JTS is a web-based environment for learning Java language (Zapata-Rivera and Greer, 2001) based on a CORBA platform and using Microsoft agents. In this environment, students have access to their student models and they are able to change it, in the case they do not agree with the information represented. Another example is I-Help (Vassileva *et al.* 1999), a web-based application that allows students to locate human peers and artificial resources available in the environment to get help during learning activities. I-Help is an example of a large-scale multi-agent learning environment (Vassileva *et al.* 2001). Moreover, interesting results have been achieved by pedagogical agents (Johnson *et al.* 2000) regarding the student motivation and companion agents (Chan 1996) acting sometimes as mediators (Conati and Clave 2000) of the learning process. Finally, tutor agents (Ritter 1997) are usually related to student modelling and didactic decision taking (Ritter and Koedinger 1996).

Next section introduces our methodological approach to design MAS; it is followed by two sections describing higher and lower levels of MAS. Last section illustrates Baghera platform.

4.1.2. AEIO METHODOLOGY

We have employed the AEIO methodology for a multi-agent-oriented analysis and design of systems (Demazeau 1995). This methodology considers the problem to be modelled as composed by four elements: Agents, Environment, Interactions and Organization. The first step of the methodology consists in defining the four elements. According to the specificities of each problem, one or more of the followings approaches may orient the system designer:

1. Agent-oriented: when it is possible to define the number of agents, their roles, their type (cognitive, reactive, hybrid) and their behaviours to guide problem solving;
2. Environment-oriented: when a well-represented world is the most important part of the system (e.g., mobile robots moving in a room);
3. Interaction-Organisation-oriented: when powerful mechanisms of coordination and cooperation among agents are needed; when interactions among agents are central to the application; and finally, when agents need to form coalitions to solve problems.

Once the four components are identified, the methodology proposes how to specify them. Of course the best tool to specify the component's behaviour depends on its function, role and type. Usually agents' behaviours are expressed by finite automates (for reactive agents) or more complex models as knowledge-based systems (for cognitive agents). Environment modelling is dependent on the application domain and very often spatial models are employed. Interaction languages can be based on models of force or allow a higher level of communication based on speech act theory. Finally, organisation model may be inspired by behaviours observed by researchers on biology or sociology.

In the case of the target application, the approach employed has privileged interactions between artificial agents and humans (students and teachers). The interaction-organisation-oriented approach was chosen and we started by defining the agents without an extensive description of their roles. Next, natural-language-based scenarios were created to study interactions among users and the application (the agents) for each use case. This process was repeated many times until a coherent set of agents, interactions and behaviours has came out. Through this process a so-called higher-level MAS was specified and it is presented in the next section. We distinguish the main MAS, whose behaviour provides the main functions of the application, from the secondary MAS, or lower-level MAS. Reactive agents compose the secondary MAS and they are responsible for diagnosing student's conceptions.

4.1.3. HIGHER-LEVEL MAS

Students and teachers interact with different kinds of agents. Persistent data are kept in students' schoolbags and teachers' electronic folders and they are in the model represented by objects belonging to the environment. They are personal repositories of data and teachers have access rights to schoolbags of students belonging to their classes. Each student counts on three artificial agents:

Companion - Student's Personal Interface Agent

It is an agent associated with the student's interface with a wide range of goals. Mainly it monitors the student's actions, notifying other agents when needed and giving access to system resources. This agent controls the access to the student's schoolbag and brings to the user information about the whole learning environment.

Tutor Agents

Tutor agents propose the most suitable problem/situation to the student, regarding educational goals and the context of learning. Furthermore, their didactical decisions are based on students' conceptions. To accomplish their goals they are able to launch the lower-level MAS whenever a diagnosis is needed (e.g. a student has finished an exercise, so the tutor has to decide what to propose as next activity) and, once diagnosis phase is over, they plan interactions with other agents and users.

Mediator Agent

The aim of this agent is to choose an appropriate problem solver to send the student's solutions. As it is shown later for the case of geometry proof learning, this agent is connected to an automatic theorem prover, being able to perform proof verification, propose alternative proofs and build counter-examples. Besides these functions, this agent implements techniques to analyse and present proofs.

Similarly, each teacher counts on the two following artificial agents:

Teacher's Personal Interface Agent

It is an agent associated with the teacher's interface. This agent controls the access to the teacher's electronic folder and brings to the user information about the whole learning environment. This agent mediates interface functions related to: communication with other human and artificial agents, edition of new activities to the students, distribution of such activities to students, and supervision of work done by students.

Assistant Agent

An assistant agent is also a kind of personal agent whose goals include assisting the teacher with the creation and distribution of new activities, which are kept in the teacher's electronic folder. This agent controls the access to the teacher's electronic folder and, when demanded, it hands the activities out to students.

As an open MAS, the number of agents in the society increases or decreases depending on the number of users logged in. For instance, in a specific moment, given a number n of students and m of teachers logged in, the number of active artificial agents is 3n+2m. We consider this an important remark since the number of connections is not limited and the number of agents is not fixed in the society.

The general architecture of the agents is shown in the Figure 8 and it was inspired by the BDI model (Rao and Georgeff 1995).

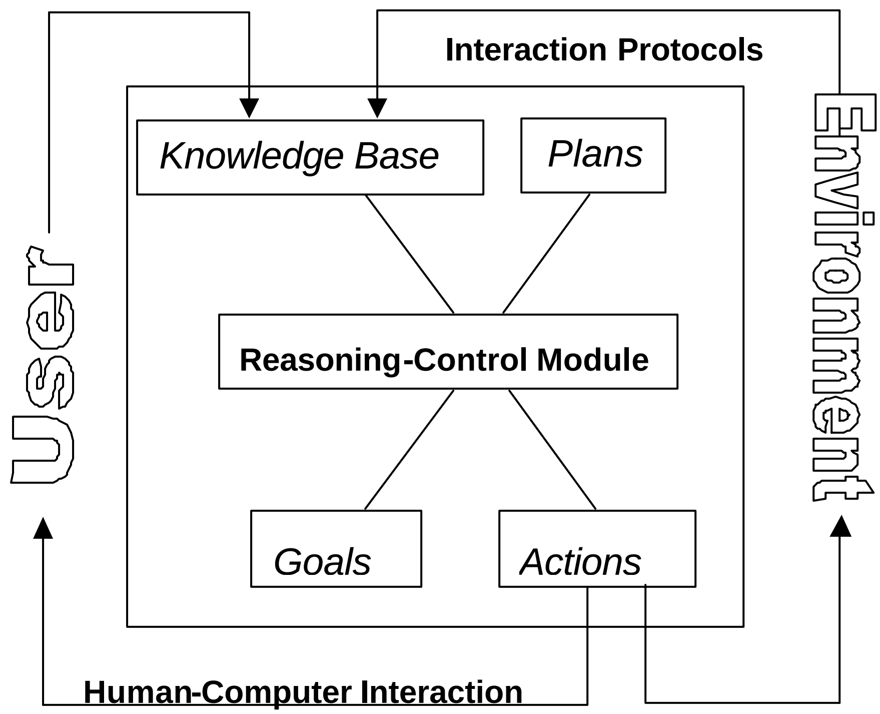


Figure 8. General agent architecture.

The knowledge base represents internal states of the agent itself and the others. Agents communicate by interaction protocols. In addition, an agent may have plans, goals and actions to run. Actions are expressed by one or more messages destined to other artificial agents or users. The reasoning and control module is responsible for assuring the execution of protocols and determining the sequence of actions.

4.1.4. LOWER-LEVEL MAS

The lower-level MAS has the goal of diagnosing student's conceptions. This is one of the bottlenecks of research on learning environments. Our approach was conceived taking diagnosis as the emergent result of collective actions of reactive agents. Here we present briefly the general aspects of the lower-level MAS. All the details will be given at section 4.2.

Conceptions are characterised by sets of agents. The society of lower-level agents is composed of four categories: problems, operators, language and control. Each element from the quadruplet C (P, R, L, Σ) is the core of one reactive agent. So, each agent taking part in the society belongs to one category and has a unique behaviour inside the society. An agent, in a given time slot, can be either active, because it is satisfied, or inactive and unsatisfied. This state can vary according to changes it perceives in its environment, in the objects and in other agents' states.

The general role of any agent is to check whether the element it represents is present in the environment. In the presence of the element, the agent becomes satisfied. Once satisfied, the agent is able to influence the satisfaction of other agents by voting. Voting mechanism is introduced later in this report. In the case of the absence of the element represented, the agent looses the right to vote but still may have some influence over the satisfaction of other agents. A description of the role of each category of agents is given later at section 4.2.

When agents from the lower-level MAS become stable, since no significant changes are perceived in agent's states and voting preferences, diagnosis is considered to be over. At this moment, the tutor agent (from the higher-level MAS) uses the results from voting as an input to decide how to respond to the student. For instance, once a correct conception is diagnosed, tutor agent may want validate the diagnoses by proposing a problem where correct procedures of resolution are tested and possibly reinforced. On the other hand, if an incorrect conception is diagnosed, the tutor agent may propose problems where the procedure of resolution, previously employed by the student, fails or it may propose problems where the correct procedures of resolution appear, making the student aware of them. Besides the interactions with the lower-level society of agents, other sources like problem solvers — HOARD-ATINF discussed later — and other tutor agents may be useful providing more inputs for tutors' didactical decision making.

4.1.5. BAGHERA LEARNING ENVIRONMENT

Baghera provides individualised support for problem solving in the domain of geometry proof. The current version of the platform was developed using JatLite[1], a package of programs for creating software agents. Each agent was extended by an interaction module. It provides support for creating protocols and coordinating interactions and the execution of protocols (Huget *et al.* 2000). Agents have the ability to communicate with other agents and take decisions. Communication among agents is based on the speech act theory in accordance with FIPA-ACL standards[2].

Students and teachers have access to Baghera through our website (http://www-baghera.imag.fr). A user identification and password are required by the applet, which gives access to the application. Baghera was implemented using Java language and Swing libraries. Figure 9, Figure 10, and Figure 11 show student's interfaces.

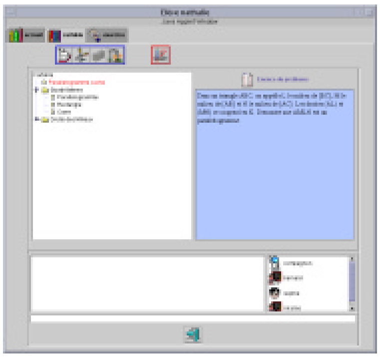


Figure 9. Student's electronic schoolbag.

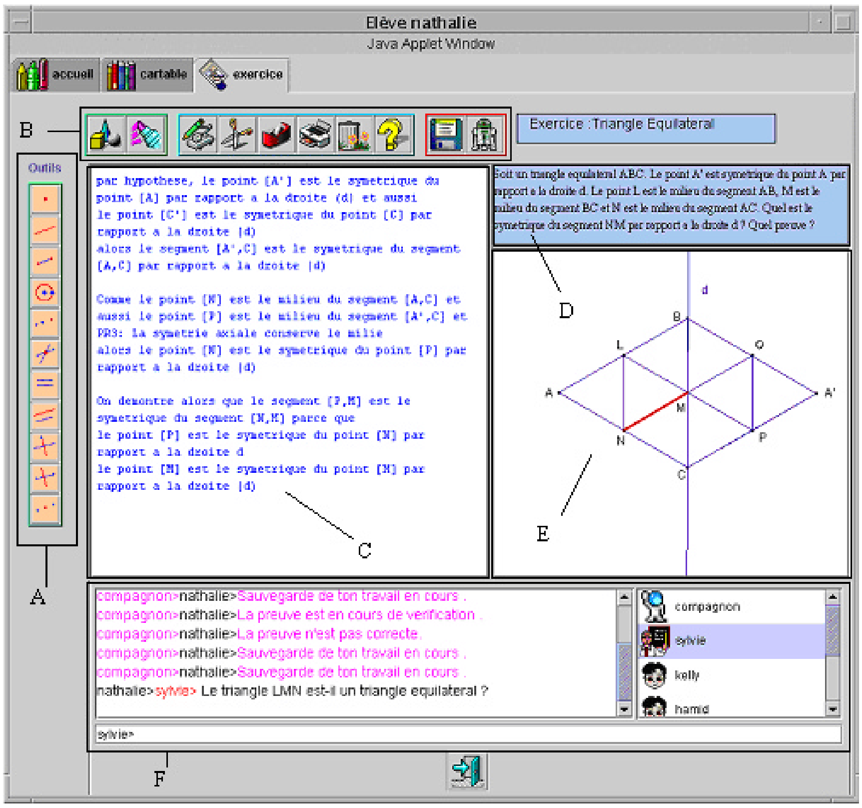


Figure 10. Problem solving interface. A and B are tools to edit a proof. C is a free text. D is the statement of a problem and E is a dynamic geometrical figure.

[1] JATLite - Java Agent Template, Lite. Available at http://java.stanford.edu/.
[2] Foundation for Intelligent Physical Agents. Available at http://www.fipa.org/.

Figure 9 illustrate the student's schoolbag, where problems sent by teachers are kept and organised. Once a problem is chosen, the student goes to problem solving interface, shown on Figure 10. On this interface students find tools to edit a proof (A and B) as a free text (C). The problem proposed is composed by a statement (D) and a dynamic figure (E) constructed using Cabri-Java[3]. The dynamic figure can be manipulated in some ways that the student may verify the geometrical properties of it and new problem hypothesis may come out.

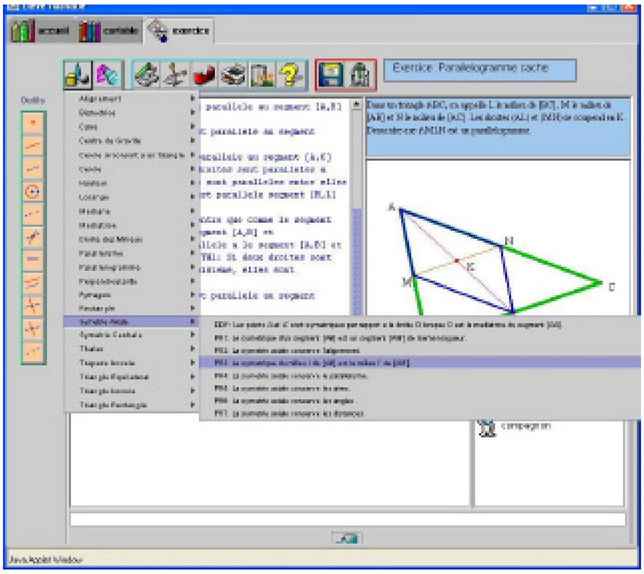


Figure 11. Library of theorems in geometry.

Following, Figure 11 illustrates the library of theorems and properties from geometry that students can consult and insert into their text. The proof can be verified whenever wanted by the student. And at this moment, proof is translated and sent to ATINF, an automatic theorem prover instantiated for the domain of geometry (Caferra *et al.* 2000). The companion agent informs the student about the state of his/her proof.

Teachers' interfaces have extra functionalities, as partially shown on Figure 12, Figure 13 and Figure 14.

Figure 12. Creating new problems.

Figure 13. Virtual class.

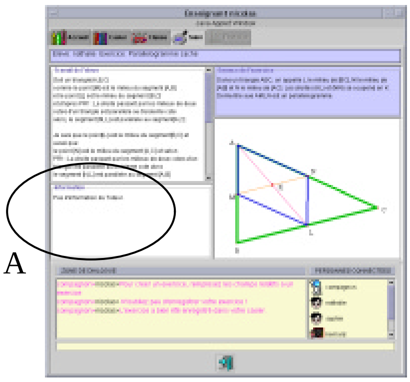


Figure 14. Supervision of student's work, including the diagnosis from lower-level MAS (A).

[3] Cabri Java. Available at http://www.cabri.net/cabrijava/.

The first capture (Figure 12) shows the interface for creation and diffusion of problems. The second one (Figure 13) presents the interface that allows the teacher to control student's connections and access to schoolbags. Finally, the last capture (Figure 14) allows the teacher to supervise one student's work by the observation of proofs constructed regarding particular problems. Also through this interface, the diagnosis from lower-level MAS is presented to the teacher (A).

4.1.6. CONCLUSIONS

Baghera is a hybrid multi-agent architecture of a learning environment. Cognitive agents form the higher-level and their interactions determine the behaviour of the whole system. On the other hand, reactive agents are launched to collectively diagnose student's conceptions. Diagnosis is taken as the emergent result of the interactions of reactive agents. It is intended to constitute an educational community of artificial and human agents, which cooperation helps students to learn. In addition, it is important to remark that our main goal is not only toward the development of a new distributed architecture for educational systems. We intend to go further and search for models based on emergent theory, which could overcome the absence of a general pedagogical model of teaching by allowing the dynamic construction of strategies based on local educational solutions.

4.2 Emergent Diagnosis via Coalition Formation

4.2.1 INTRODUCTION

According to Maxion, diagnosis is a form of high-level pattern recognition of symptoms or symbols and it may be one emergent property of certain complex systems (Maxion 1990). Usually diagnosis is defined as a process of identifying a situation or a system condition from its intrinsic characteristics. Additionally, a diagnosis may allow a system to adapt itself to the immediate constraints of the environment, to resources reallocation, and to the different categories of users.

Essentially, we assume that a diagnosis is a process where microscopic observable findings are recognised by macroscopic entities and they may determine global system's behaviour. For instance, in the case of user-adapted applications, a limited set of elements observed from user interactions may permit a system to diagnose the user's level of expertise and adjust its behaviour according to it. In this case, the level of expertise ascribed to a user could be the result of a diagnosis process.

Systems having multiple interacting components and a behaviour that cannot be simply inferred from the behaviour of the components are qualified as complex (Schweitzer and Zimmermann 2001; Holland 2000). Diagnosis, as a product, may be an emergent property of some complex systems. Multi-agent approach brings some advantages for modelling complex systems since: (1) its application is not dependent on the number of agents (contrarily to certain approaches where a large number of elements is necessary); (2) agents can have heterogeneous behaviours; (3) interactions of different levels of complexity are allowed; (4) it is applicable to several domains (social sciences simulations, computational economy, ecology, physics, and so on).

The idea that systems constituted of several agents having simple behaviour (a behaviour described by a few rules, for instance) can show a dynamic global behaviour having properties not easily predictable (even if the external condition are known) is exploited in our work. A few examples have demonstrated such emergent and unpredictable behaviour. For instance, simulations in the domain of voting theory have shown how parties emerge from voters' choices (Schreiber 2000; Stadler 1999). Also, experiments in computational economics have illustrated the emergence of markets based on the behaviour of agents representing costumers and vendors (Vriend 1995) as well as the emergence of social classes and social norms (Axtell *et al.* 2000). In a similar fashion, the two domains apply coalition formation mechanisms to simulate and study social behaviours. But beyond simulation purposes, coalitions are as well applied to problem solving. Task allocation has been one of the most applicable examples of coalition formation (Shehory and Kraus 1995). Most recently, electronic marketplace has shown to enclose enough dynamic aspects to constitute an excellent testbed for mechanisms of coalition formation.

The work we describe here considers diagnosis as a problem-solving task and we propose to solve it by coalition formation. In addition, we assume that user modelling, more specifically student modelling, is a process of diagnosing the 'state of conceptions' hold by an student in interaction with a learning environment (Webber *et al.* 2002). The diagnosed conceptions are ascribed to the student and kept in his/her student model to guide didactical decisions (the choice of problems, advices, etc.) of Baghera, a distance learning environment. The framework we propose for student modelling is based on the model of conceptions (Balacheff and Gaudin 2002; Balacheff 2000b).

Sections describing the diagnosis are organised as follows. Next section briefly describes the theoretical framework of our approach; it is followed by one section which introduces the MAS for diagnosing conceptions and a second one which describes the elements composing the system and mechanisms for coalition formation. Finally, some experiments realised are described.

4.2.2 Theoretical Framework

Emergence

An emergent system is characterised by having a behaviour that cannot be predicted from a centralised and complete description of the component units of the system (Forrest 1990; Sawyer 2000). In emergent systems, the overall behaviour is the result of a great number of interactions of agents obeying very simple laws. The overall behaviour cannot be anticipated by simple reduction to individual behaviours, following a logico-deductive model, but it is rather conditioned by the immediate surroundings, like other agents and objects in the environment.

Very often the definition of emergence is attached to the notion of levels and detection (Bonabeau *et al.* 1995). For this reason diagnosis can be effectively seen as an emergent property of certain complex systems, since in a diagnosis process lower-level symptoms or symbols are recognised by higher-level entities (Maxion 1990). Note that emergent objects have a representation distributed over many different elements. Each of these elements may take part of many different objects simultaneously. This may be observed in the classifier systems proposed by Forrest and Miller (1990), in the system for diagnosis of communication networks (Maxion 1990), and in the emergence of conceptions presented here.

Conception Theory

We review here some elements from the conception theory (details are given in section 2 of the report) in order to present the model we propose.

In this model a conception is characterised by a quadruplet C (P, R, L, Σ) where:

- P represents a set of problems;
- R represents a set of operators involved in the solutions of problems from P;
- L is a representation system allowing the representation of P and R;
- Σ is a control structure.

An element from any set can contribute to the characterisation of several different conceptions; for example two conceptions may share problems in their domain of validity or may have common operators. A deep presentation of the conception theory was already given in this report. Despite of this, we propose the examples presented in Figure 15 and Figure 16.

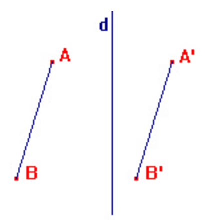


Figure 15. Parallelism

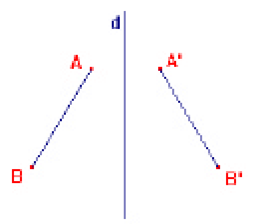


Figure 16. Reflection

Figure 15 presents a construction made by a student holding a misconception stating that "if two line segments are symmetrical then they are parallel". Figure 16 was constructed by a student holding the correct conception of reflection.

In order to illustrate how diagnoses occur, we examine the problem described on Figure 17.

Problem statement: Let ABC be an equilateral triangle. Point A' is the symmetric of A with respect to the line d. L is the midpoint of segment [AB], M is the midpoint of segment [BC], N is the midpoint of segment [AC]. P is the intersection point of the line (LM) and the line (CA') and O is the intersection point of the line (MN) and the line (BA').
What is the symmetric of the segment [MN] with respect to the line d? How can you prove it?

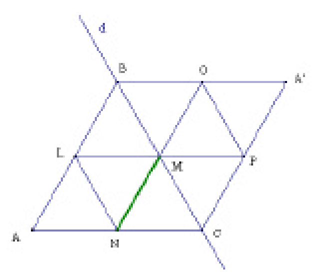


Figure 17. A problem in the domain of reflection

In this problem students are asked to prove, using geometrical properties of reflection, that the line segment [NM] has a symmetrical object with respect to the axis d. Let's consider one strategy to solve it, which consists on proving each step from table 1. In the case of step 6, we consider four alternatives (6.a, 6.b, 6.c and 6.d) in order to exemplify how students holding different conceptions would express the solution. First, consider the proof composed by steps 1-2-3-4-5-6.a where the student has proven that [OM] is the symmetrical line segment of [NM]. This solution, by the operators used to construct it (6.a1-6.a7), characterizes the so-called misconception of 'central symmetry'.

1		A'BC is an equilateral triangle
2		ABA'C is a lozenge ([AB]//[CA'] and [BA']//[AC])
3		[AB]//[CA']; [AB]//[NO]; [NO]//[CA']
4		O is the middle point of [A'B]
5		P is the middle point of [A'C]
6.a	1	M is its own symmetrical point with respect to d
	2	As [AC]//[BA'] and
	3	N is the middle point of [AC] and
	4	O is the middle point of [A'B] and
	5	Line segments [NM] and [OM] have the same size
	6	O is the symmetrical point of N
	7	So, [OM] is the symmetrical line segment of [NM] with respect to point M
6.b	1	M is its own symmetrical point with respect to d
	2	Line segments [NM] and [PM] have the same size
	3	[NP] is perpendicular to axis d
	4	P is the symmetrical point of N
	5	So, [PM] is the symmetrical line segment of [NM] with respect to d
6.c	1	As [NO] // [CA'] and [NM]//[PA']
	2	As [NM] and [PA'] are parallels and have the same size, they are symmetrical.
	3	P is the symmetrical point of N
	4	A' is the symmetrical point of M
	5	So, [PA'] is the symmetrical line segment of [NM] with respect to d
6.d	1	As [NO] // [CA'] and [NM]//[CP]
	2	As [NM] and [CP] are parallels and have the same size, they are symmetrical.
	3	P is the symmetrical point of N
	4	A' is the symmetrical point of M
	5	So, [CP] is the symmetrical line segment of [NM] with respect to d

Table 1. Possible strategies to solve the problem

The second alternative (1-2-3-4-5-6.b) gives the correct answer ([PM] is the symmetrical segment of [NM]) and its attached to the conception of reflection. Third (1-2-3-4-5-6.c) and forth (1-2-3-4-5-6.d) alternatives, even though they give different answers, they characterize the same misconception of 'parallelism'. In these cases (6.c and 6.d), students state that two line segments are symmetrical if they are parallel and have the same size and it is possibly an inversion of the correct operator, which states that two parallel line segments having the same size are symmetrical with respect to a parallel axis.

It is important to note that a diagnosis is based on a sequence of problems solved by the student. Different problems in a well-oriented sequence permit the construction of a student model having enough information to characterize student's conceptions in a specific domain of knowledge.

Voting Theory

Voting models are widely used in social sciences and have their roots in Game Theory. Social sciences research about voting has been investigating new approaches to studying voting schemes, voter behaviour, and the influences of manipulation of votes and insincere voting. Studies based on simulation of elections have led to models providing explanations to voter behaviour, so as explanations to group decisions and coalition formation.

In the domain of MAS, voting theory has been used as a technique for reaching consensus in a negotiation process and group-decision making (Sandholm 1999; Weiss 1999). In the simulation of coalitions formation, agents have been used to demonstrate how it occurs from individual voter preferences (Schreiber 2000).

Furthermore, it has been shown that emergent structures can be resultant of a voting process. Schreiber (2000) has demonstrated through multi-agent simulations that elites and parties are emergent consequences of the behaviour and preferences of voters.

In essence our interest in voting theory relies on the possibility of capturing group decision as well as modelling the influence of an agent preference over the preferences of the rest of agents.

Our approach is based on spatial models of simulation of voting behaviour. Spatial voting theory has its origins in the field of Political Science (Stadler 1999). This model assumes that political issues can be quantified and therefore voters and candidates can be represented by points in a so-called issue space. Similarly, each candidate is described by a platform position in the issue space.

Usually this space is viewed as the Euclidian vector space RI having I dimensions. Each voter is represented by a vector in the issue space, called ideal point that represents its opinion on each issue. In these models, voters form coalitions with other voters close to them in the issue space. Coalitions start with a small number of voters and possibly form coalitions with other coalitions to increase their potential. A hierarchy of coalitions is built until a coalition is created with the majority of voters. The coalition with the majority rules and the competing coalitions adapt platforms to gain greater support. Each voter may represent a single voter or a team of voting people. Agents may be voters or coalitions and, in the last case, they may represent an aggregation of voters but having no votes by themselves. The action in this model takes place only in the issue space.

4.2.4 MULTI-AGENT ARCHITECTURE

We propose a MAS where, at a microscopic level, agents behave and interact and, at a macroscopic level, coalitions emerge. The emergence of coalitions is interpreted as the emergence of a diagnosis of student’s conceptions.

The conception theory allows to model student’s conceptions. However, conceptions are not elements possible to be directly observed. Observable elements are operators used by student, the problem solved, the language used to express them, and theoretical control structures. For this reason, the micro-level is composed by elements (from the quadruplet) defining a conception. At the macro-level, conceptions can be seen as sets of agents of four categories: problems, operators, language and control. Each element from the quadruplet is the core of one particular agent.

Agents

An agent, in a given time slot, can be either active or inactive. This state can vary according to changes it perceives in its environment. The first action of any agent is to check whether the element it represents is present in the environment. In the presence of the element, the agent becomes satisfied. Once satisfied, the agent takes part in the issue space and it is able to form coalitions. Notice that an agent knows previously to which conceptions the element it represents belongs. A description of the role of each category of agents is given below.

Problem Agents. A problem agent becomes satisfied when the category of problems it represents is present in the environment. In the domain of reflection, a category of problems is described by four didactical variables named: line of symmetry orientation, segment orientation, angle formed between line of symmetry and line segment and intersection formed between the line of symmetry and line segment. The combination of the different values that these didactical variables could take, leads to problems of different complexity, allowing to focus on different aspects of the learning of reflection and most important, allowing the expression of different conceptions.

Operator Agents. An operator agent becomes satisfied when the element r of R it represents, is present in the solution constructed by the student. An operator transforms a problem in a new problem. A sequence of operators leads to the problem solution. An example of an operator is as follows: if two symmetrical objects have one point in common, then this point belongs to the axis of symmetry.

Language Agents. A language agent becomes satisfied when the element l of L it represents, is present in the solution constructed by the student. It can be a grammar, a graphical representation, or an alternative way of expression allowing the description of the problem and the solution. Given that our problems ask for the construction of a proof, the language is based on a grammar (for reasons of brevity it is not presented here).

Control Agents. A control agent becomes satisfied when the element s of Σ it represents, is present in the solution constructed by the student. During problem solving, learners choose operators, validate actions and validate the final result. Each of these three decisions is guided by control structures. Control elements are perceptive when attached to the fact that the learner makes assertions based on something "seen" on the screen and uses this information to take and validate decisions. On the other hand, control structures are theoretical when a learner bases decisions and validations on knowledge previously acquired. Reflection involves many visual elements of control; for instance, a learner holding the conception of parallelism may accept that a problem is correctly solved when the image line segment "looks" parallel to the original line segment. In the case of our system, we consider only theoretical controls and some perceptive controls that can be expressed by means of a proof.

Environment

The environment represents the external world that agents have to deal with and it co-evolves with the agents. In the environment there is a representation of the problem solved by the student, the proof corresponding to the student's solution and the issue space where coalition formation takes place.

Interactions

Agents interact with the environment and through these interactions they transform the environment (issue space). Such transformation generates changes in agents' behaviours. The cycle of interactions continues indefinitely until no more coalitions can be formed or merged.

Organisations

We apply a dynamic approach to conceive organisations. We consider that agents form dynamically coalitions when they are needed to solve a problem. Our approach of coalitions formation is based on emergent approach. Moreover, it is considered that once the problem is solved, coalitions are not applicable anymore for a later processing.

In the next section we proceed with the formalisation of the most relevant components of the model.

4.2.4 FORMAL DESCRIPTION

Environment

The environment is described by a set Env = {PR, SP, IS} where PR represents the problem that the student has solved, SP represents the proof constructed by the student as a solution to the problem PR and IS is the issue space. The space of votes $þ^I$ is the Euclidian space having I dimensions. The number of conceptions to be diagnosed determines the number of space dimensions. Voters are represented by a position in the space corresponding to their 'opinions' about candidate conceptions. A position is represented by a vector $v \in þ^I$.

Diagnosis Problem

Given a set of conceptions $C = \{c_1,\ldots,c_i\}$, a set of n agents $A = \{a_1,\ldots,a_n\}$ and a set representing the state of the environment Env = {PR, SP, IS}, the problem we propose to solve consists in assigning one or more (possibly concurrent) groups G of agents ($G \subset A$) representing the state of student's conceptions reflected in the environment.

Agent

Let C be a set of conceptions $\{c_1, c_2,\ldots,c_n\}$. Consider that any conception c_i in C is defined by a quadruplet (P_i,R_i,L_i, c_i) where :

- P_i is a set of problems $\{p_1, p_2,\ldots,p_n\}$ of c_i ;
- R_i is a set of operators $\{r_1,r_2,\ldots,r_n\}$ of c_i ;
- L_i is a grammar for the expression of P_i and R_i of c_i ;
- Σ_i is a set of control structures $\{ó_1, ó_2,\ldots,ó_n\}$ of c_i.

Let A be a set of agents $\{a_1,\ldots,a_n\}$. Let K_i be a set of n candidate conceptions $\{k_{i1},k_{i2},\ldots,k_{in}\}$ for an agent a_i where $K_i \subset C$. Let E be a set of elements $\{e_1,\ldots,e_n\}$ from the conceptions formalisation and assume that e_i is the element in the core of an agent a_i. About any element e_i it is known that $e_i \in P_i$ or $e_i \in R_i$ or $e_i \in L_i$ or $e_i \in \Sigma_i$. Let Q_i be the set of acquaintances of an agent a_i (acquaintances are detailed later in this report). Finally, V is the set of votes $\{v_{1k},\ldots,v_{ik}\}$ given by the agent to preferred candidate conception from K_i .An agent a_i is defined by: an identifier N_i, an internal state $S_i \in$ {satisfied, unsatisfied}, a set of acquaintances Q_i, a set of candidates conceptions K_i, an element e_i, a satisfaction function $f_i(e_i, Env)$ and a vector V_i representing its starting position in the Euclidian space $þ^I$.

Agent Behaviour

Agents are created in an unsatisfied state and the satisfaction function may change its state to a satisfied one. When an agent becomes satisfied, it creates its vector to be added to the issue space. Since it is situated in the issue space, its acquaintances are set. Agents start forming coalitions with each member of its acquaintances list. Any agent may take part in any number of coalitions. Once an agent takes part in proposed coalitions, it can accept or refuse it. Besides, an agent tries to merge coalitions in which it takes part into. And finally, agents are free to enter and leave coalitions at any time. When it is not possible anymore for an agent to execute any of these actions, then it stops running. The major steps of the algorithm defining the behaviour of an agent are as follows: (1) initialise the data structure; (2) calculate agent's power of voting; (3) set agent's acquaintances; (4) while (list of proposed coalitions is not empty): (4.1) propose coalitions; (4.2) accept coalitions; (4.3) refuse coalitions; (4.4) calculate coalition's utility; (4.5) merge coalitions; and (4.6) abandon a coalition.

Finding Acquaintances

The most important feature of an agent is its voting vector, representing its choice of candidate conceptions. Its acquaintances represent a list of agents that are spatially located close to it in the issue space.

Let a_α and a_β be two different agents. Let V_α and V_β be vectors representing respectively the positions of a_α and a_β in the space $\flat^I$. Let Q_α and Q_β be the set of acquaintances (initially empty) of respectively a_α and a_β. We assume that a_α and a_β are acquaintances if they satisfy the neighbourhood condition.

Neighbourhood condition is calculated by the formula of Euclidian distance between the two vectors V_α and V_β. The two agents satisfy the condition if the distance is a value under a specific threshold and in this case, $a_\beta \in Q_\alpha$ and $a_\alpha \in Q_\beta$. Otherwise, a_α and a_β are not acquaintances to each other.

Coalition Formation

A coalition is a nonempty subset Co of A and it has as utility value the sum of utilities of all agents belonging to it. We follow the traditional approach of coalition formation in the domain of MAS (Sandholm 1999).

The initial number of coalitions is reduced since the initial coalitions are formed between any two agents situated spatially close in the issue space. When agents form a coalition it has a status of proposed coalition and when it is accepted by all of its members it becomes an accepted coalition.

4.2.5 First Experiments and Some Final Considerations

For the first experiments, we have run a diagnosis considering the proof presented at section 2.2 (steps 1-2-3-4-5-6c). The MAS is composed of 101 agents, distributed as follows: 60 operator agents, 30 problem agents and 11 control agents. As explained before, agents become active if the element represented is found in the proof. For this experiment, 13 agents (1 problem agent, 10 operator agents and 2 control agents) have become active. The issue space has 4 dimensions, representing 4 conceptions on reflection (central symmetry, oblique symmetry, parallelism and reflection). Vectors of agents' opinions have 1 in each dimension representing a good candidate and 0 otherwise. The threshold for calculating acquaintances was 1.

In the end of this experiment two coalitions have emerged. The greatest number of coalitions, reached at interaction number 150, was of 134 coalitions. Coalition formation was stopped when this number was reduced to 2 coalitions and no relevant changes in the system were observed. The coalition having the greatest utility is considered the winner. Among the 13 agents involved in the diagnosis process, 10 of them took part of the winner coalition. In this experiment, the winner represents the misconception of parallelism; the second coalition represents the conception of reflection (the correct one). The result is satisfactory and it indicates that the student possibly holds the misconception of parallelism. We interpret the fact that a weak coalition (representing reflection) has appeared as a result of correct operators and properties present in the student's proof; they certainly must have a weight in the diagnosis process. However, incorrect operators have appeared as well and they have induced the diagnosis of parallelism by forming the strongest coalition.

As long as the diagnosis process is over, macro-agents (belonging to Baghera main MAS) take didactical decisions based on it. For instance, tutor agents may propose new problems to confirm a diagnosis or propose a problem where incorrect procedures of resolution fail, according to didactical strategies.

Following to these first experiments, we have integrated the diagnosis system (lower-level MAS) to the Baghera platform. We have run the system over 30 student's solutions and the result of this "artificial diagnoses" were compared to the "natural diagnoses" (made by the researchers from mathematics education working in the project). The analysis of the results can be checked at section 6 of this report.

To conclude, the main challenge of this task has been to define and implement a computational framework to model student's conception supported by the conception theory developed in the domain of mathematics education.

4.3 The logical diagnosis

As it has been shown previously (section 4.2), the learner's problem solving activity is analysed by the system from the point of view of conceptions. The learner is thus ascribed a set of conceptions that describe the present state of the learner's knowings related to reflection. This point of view does not take into account the coherence and validity of proofs provided by the learner. This information is needed in order to suggest and support didactic decisions aiming at evolving the learner's knowings. An automatic theorem prover, HOARD-ATINF, is being developed for checking proof correctness and coherence.

4.3.1 HOARD-ATINF – A GENERIC LOGICAL THEOREM PROVER

In this document, we present the theorem prover HOARD-ATINF we developed in the context of the Baghera project. HOARD-ATINF is a generic (logical) theorem prover especially devoted to solve problems in geometry. The word "theorem-prover" is to be understood in a broad sense here: of course, the system is able to construct proof, but it also has a lot of useful functionalities that are usually absent from existing systems. It allows proof analysis (verification of proofs in a broad sense), it detects analogies between a new formula and previously proven one (and is able to use this analogy to guide the search for a proof of the new formula), and constructs, in a purely automatic way, models (or counter-example) of sets of formulae (for this particular application domain, models are geometric figures).

Though HOARD-ATINF may be used independently, it is mainly intended to be used in connexion with the Baghera plate-form (this explains why no user-interface has been developed for communicating directly with the prover). When reasoning capabilities are needed, the agents send requests to the prover (on the form of computer-generated input files) and the output files are automatically sent to them.

For the moment, only the proof verification abilities are effectively integrated to the plate-form. However, in the future, the use of more advanced capabilities will be investigated, in particular of help diagnosis.

4.3.1.1 The language

The language we use to encode geometric assertions and theorems is a subset of first-order logic with equality. We consider terms and atoms built on a set of predicate symbols (*parallel*, *perpendicular*, etc.) and on a set of constructors (line(A,B), segment(A,B), etc.).

Theorems and definitions are expressed as rules (*operators*) for instance:

parallele(d1,d2), parallele(d2,d3) ® parallele(d1,d3)

The reader should note that the language used to express proofs and to communicate with the plate-form is *completely independent* from the one used by HOARD-ATINF to make inferences and *generate* proofs. The problems sent to the prover are translated before being processed into a language used in an internal way by HOARD-ATINF, and the result is translated back from this internal language before being sent to the plate-form. Indeed, despite being very close to each other, the two languages must obey different constraints: in particular, the choice of the formulation may be crucial for the efficiency of the theorem prover (in particular for reducing redundancy). Clearly, these constraints are not relevant for the agents in the Baghera plate-form hence should not interfere with the description of the "external" language. This enables to change the language used by the plate-forms without having to modify the theorem prover itself (and conversely).

The interested reader may find the complete description of the language in Peltier (2001) (list of the geometrical constructors, definitions and theorems currently used by the plate-form).

Adding of new operators

Many operators are built-in the prover. They corresponds to well-known properties of geometrical objects that are both sound and well-established (most of them are mentioned as theorems in textbooks). The name of these properties are shared between the prover and the agents in the Baghera platform. Adding new operators in this list is rather straightforward, but obviously requires to modify the theorem prover.

However, it is also possible to work with "ad-hoc" operators. For instance, some rules may be identified when analysing the students' productions. Of course the operators do not necessarily correspond to "official" geometrical theorems or definitions, and can even be incorrect or incomplete w.r.t. standard geometry, but nevertheless it may be necessary to be able to reason with these operators in order to analyse students'proofs. For example this enables to check that a proof is correct w.r.t. a set of operators, to determinate whether a set of operators – though incorrect from a mathematical point of view – is *consistent* and operative on a set of given problems etc.

Therefore, HOARD-ATINF offers the possibility of defining new operators in the input file. These user-defined operators may be used by the theorem-prover exactly as built-in operators.

Notice that the theorem prover does not check that the considered property is correct. Thus uncontrolled adding of new operators may lead to incorrect or even inconsistent answers.

**Non degeneracy conditions*

A number of geometrical properties that can commonly be accepted as "valid" actually only hold w.r.t. some implicit additional hypothesis, insuring that the geometrical objects are non degenerated (for example a quadrilateral may be flat or non convex, a line may be reduced to a single point etc.). For example, let us consider the following operator: "Let (A,B,C,D) be a parallelogram. If E is the midpoint of [BD] then E is the midpoint of [AC]". Clearly, this assertion is wrong if (A,B,C,D) are on the same line (which corresponds to a "degenerate" parallelogram).

The “correct” version is the following operator: “Let (A,B,C,D) be a parallelogram. If E is the midpoint of [BD] then E is on [AC]”. Then, using the unicity of the intersection, in case (AC) and (BD) are distinct we deduce that E is the midpoint of [AC].

These conditions are called “non-degeneracy conditions”.

They are necessary to insure that the considered properties are valid, but are often not explicitly checked during the proof (i.e. the operator are applied without checking that the geometrical object satisfy the desired properties).

HOARD-ATINF offers 3 distinct way of handle such conditions.

The first possibility is simple to ignore these conditions and apply the operator as they were always satisfied. This is what is done for example by the system Tigre (Mentoniezh project (Py 1996)). The technique is fast and efficient but the obtained proofs may reveal incorrect.

The system may try to systematically check all the non-degeneracy conditions. This technique ensures that proofs are correct, but may be very costly. In particular, all the hypotheses and axioms that are necessary for proving these conditions must be provided (which may be a problem because they are not always available).

A third possibility consists in trying to use a figure to check whether these conditions are valid or not. The idea is to construct a model of the hypotheses, and check whether the non-degeneracy conditions hold or not. This technique is used by the system Geometry Expert . It is very elegant and discard most incorrect inferences, which ensures in most cases (but obviously not always) that the proofs are correct. A drawback of course is that the system must build his own figure (notice that making Cabri-geometry’s figure accessible to the prover would obviously solve this problem).

4.3.1.2 Proof syntax

A **proof** may be seen as a tree (more precisely a DAG, **D**irected **A**cyclic **G**raph) labelled by formulae (which correspond either to geometrical assertions or to operators). The root of the tree is the *conclusion* of the proof, whereas the leaves are either the hypotheses of the proof or the theorems and definitions used during the proof. The children of a given node corresponds to the set of assertions, theorems and definitions that are used to deduce the considered formula.

A proof is denoted by a list of proof steps, delimitated by “$”. Each proof step is specified by a list of formulae starting by the conclusion of the proof step followed by its premises (including the name of the theorems and definitions that are used in the proof step, i.e. that justify its correctness).

For instance:

```
parallele(segment(E,F),segment(I,J))
parallele(segment(A,B),segment(E,F))
parallele(segment(A,B),segment(I,J))
transitivite_parallelisme
```

denotes the following proof step: “(EF)//(IJ) since (AB)//(EF) and (AB)//(IJ) (by the transitivity of the parallelism relation).”

**Implicit knowledge*

The standard definition of the notion of proof stipulates that the conclusion of each proof step must be deduced from the premises by applying one of the inference rules (operators). Clearly, the proofs constructed by human beings seldom fulfil this requirement. Indeed, explicitly specifying all the elementary reasoning steps is often time-consuming and useless. In order to make proof shorter and more readable, some parts of the proof are often omitted, most of the cases because they are (rightly or wrongly) supposed to be easy and self-evident for all readers.

In order to take this into account, we allow proof steps that correspond to a sequence of applications of *several* operators. Some of these operators must be specified in the premises as explained above, but others are considered as *implicit* thus do not even need to be specified. We assume that these properties are well-known and simple, so that there is no need to mention them explicitly.

The number of implicit operators allowed may of course depend on the level of knowledge of the student. Therefore, the list of operators that can be omitted must be provided to the theorem prover in the input file.

Some of the most basic properties (such as for example the commutativity of the constructor line: *line(a,b) = line(b,a)*) are not expressed as operators, but have been directly encoded into the prover (for example in the unification algorithms). This allows to improve the efficiency of the system. Indeed, it is well-known that using purely generic logical provers for solving geometrical problems turns out to be inefficient, mainly because of the high number of symmetries in the formulae, that greatly increases the number of redundant inferences.

**Constructive axioms*

The axiomatization also contains constructive operators, i.e. operators that do not merely assert properties of existing objects, but that introduces (i.e. assert the existence of) *new* geometrical objects. For example the operator

$$secant(d1,d2) \Rightarrow (\exists x)\ (x \in d1,\ x \in d2)$$

states that any pair of secant lines have a intersection point.

From a logical point of view, these operators may be translated into clausal logic simply by replacing the variable x by a new function symbol. For instance one can introduce a function *i(d1,d2)* mapping each pair of secant lines to its intersection point. Then the above axiom becomes:

$$secant(d_1,d_2) \Rightarrow (i(d_1,d_2) \in d_1,\ i(d_1,d_2) \in d_2)$$

This technique can be made systematic and allows to eliminate all existential quantifiers (with the cost of having to introduce new function symbols). This very well-known process is called "skolemization" (see for example (Fitting 1990)). This allows to handle constructive operators exactly in the same way as the other ones, without having to introduce any particular mechanism for them (as it is done for example by the theorem prover described in Chou *et al.* (2000)).

However, the new function symbols introduced during this process should not occur in the proof. They are only used in an internal way by the prover but are simply meaningless to the agent in the Baghera environment since they are not part of the built-in constructors. Therefore, the newly added function symbols are deleted from the proof afterwards and replaced by the explicit adding of new points.

To this purpose, the keyword `il_existe` is used to introduce new geometrical objects (for instance a point, a line etc.) on the figure during the proof. Two parameters are needed: the name of the object and the property it must satisfy. For instance the following proof step:

```
il_existe(X,milieu(X,segment(A,B)))
$
```

adds a new point X defined as the midpoint of the segment [A,B] (it is equivalent to the assertion "let X be the midpoint of [A,B]").

The constructor "et" (and) allows to introduce complex properties combining several assertions, for example:

"`il_existe(X,et(alignes(X,A,B),alignes(X,C,D)))`" defines a point X as the intersection of the lines (A,B) and (C,D).

We give below an example of application of this technique.

Example 1 *We consider the following problem. "Let (AB) and (CD) be two parallel lines. Let M,N the midpoints of [AC] and [BD] respectively. Let E be the intersection point of (MN) and (BC). Prove that M is the midpoint of [BC]."*

The proof relies on the use of a new point not defined in the hypotheses: the midpoint of the segment [AD]. Indeed, using this point allows to prove – by use of the midpoint theorem – that the lines (MN) and (CD) are parallel.

HOARD-ATINF automatically constructs the following proof:

```
alignes(E,B,C)
$
alignes(E,N,M)
$
egal(droite(E,N),droite(N,M))
alignes(E,N,M)
appartenance
$
egal(droite(E,N),droite(E,M))
alignes(E,N,M)
appartenance
$
egal(droite(E,M),droite(N,M))
egal(droite(E,N),droite(N,M))
egal(droite(E,N),droite(E,M))
egalite
$
point(D)
$
point(A)
$
il_existe(p7,milieu(p7,segment(D,A)))
point(D)
point(A)
construit_milieu
$
milieu(N,segment(B,D))
$
parallele(segment(p7,N),segment(B,A))
milieu(p7,segment(D,A))
milieu(N,segment(B,D))
milieux
$
parallele(segment(B,A),segment(C,D))
$
parallele(segment(p7,N),segment(C,D))
parallele(segment(p7,N),segment(B,A))
parallele(segment(B,A),segment(C,D))
transitivite_parallelisme
$
milieu(M,segment(C,A))
$
parallele(segment(p7,M),segment(C,D))
milieu(p7,segment(D,A))
milieu(M,segment(C,A))
milieux
$
parallele(segment(p7,N),segment(p7,M))
parallele(segment(p7,N),segment(C,D))
parallele(segment(p7,M),segment(C,D))
transitivite_parallelisme
$
alignes(p7,N,M)
parallele(segment(p7,N),segment(p7,M))
appartenance
$
egal(droite(p7,N),droite(p7,M))
```

```
alignes(p7,N,M)
appartenance
$
egal(droite(p7,N),droite(N,M))
alignes(p7,N,M)
appartenance
$
egal(droite(p7,M),droite(N,M))
egal(droite(p7,N),droite(N,M))
egal(droite(p7,N),droite(p7,M))
egalite
$
parallele(segment(N,M),segment(B,A))
parallele(segment(p7,N),segment(B,A))
egal(droite(p7,N),droite(p7,M))
egal(droite(p7,M),droite(N,M))
egalite
$
milieu(E,segment(B,C))
alignes(E,B,C)
egal(droite(E,M),droite(N,M))
parallele(segment(N,M),segment(B,A))
milieu(M,segment(C,A))
milieux
$
```

**Analogy*

Geometric proofs often contain parts that are symmetric, due to the presence of symmetries in the problem itself.

Therefore, for the sake of conciseness, it is possible to omit some parts of the proof if they are similar (“analogous”) to previous proof steps. This save times and makes the proofs more readable.

The keyword `analogie` is used for this purpose. If this keyword occurs in a proof step, this indicates that the conclusion of the step can be proven in a similar way as the assertion following the keyword (called the *source*). This means that the proof should be obtained by using *exactly* the same sequence of operators. Of course the hypotheses may differ. Note that the source must have been proven before the occurrence of the keyword `analogie`.

Here is an example of a proof using this technique:

```
milieu(J,segment(A,C))
$
milieu(I,segment(B,C))
$
milieu(E,segment(A,K))
$
milieu(F,segment(B,K))
$
parallele(segment(A,B),segment(I,J))
milieu(J,segment(A,C))
milieu(I,segment(B,C))
milieux
$
parallele(segment(A,B),segment(E,F))
parallele(segment(A,B),segment(I,J))
analogie%specifies that the assertion AB//EF may be proven
          % as the assertion AB//IJ
$
```

4.3.1.3 Proof construction

Two strategies are possible: forward or backward proof search.

However, most of the experiments so far have been done with forward proof search, thus in this document we limit ourselves to a very brief informal description of this strategy. The interest reader can consult (Caferra *et al.* 2000) for more details: formal description of the proof procedure and soundness and completeness proofs.

Rather than using the standard resolution calculus, we prefer to use a new inference rule, called *E*-hyper-resolution, that combines hyper-resolution with equality reasoning and rewriting techniques. *E*-hyper-resolution can be seen as a *macro inference rule* combining in a single inference step, several applications of the positive ordered paramodulation (i.e. replacing equals by equals) and positive resolution rules (a generalization of modus ponens, as defined for example in Robinson (1965)) in an effective manner.

E-hyper-resolution is especially useful when the number of equational literals is not very important w.r.t. the size of the clause set, which is the case in most of the problems we have to deal with.

The use of this rule has many advantages w.r.t. standard approaches:

- First, only *ground* clauses are deduced. Non ground "intermediate" clauses that are usually generated by standard paramodulation calculus are not explicitly generated here. This greatly simplifies the algorithm for storing the deduced clauses and for detecting redundancy. In particular, the complexity of subsumption tests are reduced if the clauses are ground.
- Second, some inferences may be avoided which reduces the search space, as shown in Caferra *et al.* (2000).

Given a set of hypotheses `hyp1, hyp2,...`, a conclusion C and a set of theorems and definitions (including implicit axioms) the theorem prover tries to prove that the formula C is a logical consequence of the hypotheses `hyp1, hyp2,...` and of the given axioms. Some parameters may control the search process, for example by specifying limits on the maximal time available, on the depth/length of the considered proofs, on the considered strategy (forward or backward proof search), etc.

If it succeeds, the output file contains the flag `[preuve]` followed by the constructed proof. Otherwise, the file contains a keyword:

```
[non_deductible] | [incorrect] | [non_trouve]
```

The flag "`non_deductible`" states that the conclusion cannot be proved from the provided hypotheses (either because it is not a consequence of the hypothesis or because the axiomatization is not complete, for instance, if some theorems and definitions are missing). "`incorrect`" means that conclusion is not a consequence of the premises (in this case there exists a least a counter-example, i.e. a model of the premises that does not satisfy the conclusion). "`non_trouve`" means that the theorem prover did not find a proof and that the search has been stopped, due to limits in the available time, memory, or proof depth or even due to a bug in the system.

Example 2 *As an example, we consider the following problem: "Let (A,B,C) be a triangle rectangle in A. M is the midpoint of [BC] and D is the midpoint of [AB]. H is the point on (BC) such that (AH) is perpendicular to (BC). G is the circle of diameter [AM]. Show that D is on G.".*

This problem corresponds to the following input file:

```
[traitement]
preuve
[hypotheses]
triangle_rectangle(C,A,B)
```

```
milieu(M,segment(B,C))
milieu(D,segment(A,B))
diametre(segment(A,M),cercle(O,R))
[theoremes]
tous
[conclusion]
est_sur_cercle(D,cercle(O,R))
```

We obtained the following output file:

```
[correct]
milieu(D,segment(B,A))
$
alignes(D,B,A)
milieu(D,segment(B,A))
def_milieu
$
parallele(segment(D,A),segment(B,A))
alignes(D,B,A)
appartenance
$
milieu(M,segment(B,C))
$
parallele(segment(D,M),segment(A,C))
milieu(D,segment(B,A))
milieu(M,segment(B,C))
milieux
$
triangle_rectangle(C,A,B)
$
perpendiculaire(segment(B,A),segment(A,C))
triangle_rectangle(C,A,B)
triangle_rectangle
$
perpendiculaire(segment(D,M),segment(B,A))
parallele(segment(D,M),segment(A,C))
perpendiculaire(segment(B,A),segment(A,C))
perp_para
$
perpendiculaire(segment(D,M),segment(D,A))
parallele(segment(D,A),segment(B,A))
perpendiculaire(segment(D,M),segment(B,A))
perp_para
$
diametre(segment(M,A),cercle(O,R))
$
est_sur_cercle(D,cercle(O,R))
perpendiculaire(segment(D,M),segment(D,A))
diametre(segment(M,A),cercle(O,R))
th\_cercle
$
```

4.3.1.4 Proof verification and analysis

A proof is considered as “correct” if for each proof step, the conclusion can be deduced from the premises by using the properties mentioned in the proof step and the set of implicit operators. This test is done by calling the theorem HOARD-ATINF.

A proof can fulfil two different kinds of properties.

- First, it can be logically and “locally” correct, i.e. each proof step must be valid w.r.t. to the considered axiomatisation and each assertion that is used in the proof must have been proved (without any “circular” argument) . A proof satisfying these two properties is said to be “coherent” which is indicated by the flag [preuve_coherente].

- Second, it must be correct relatively to the considered problem. This means that the hypotheses of the proof must belong to (or must be “easily” deducible from) the set of hypotheses (data) of the considered problem, and that the conclusion of the proof must be equivalent to the one that is expected.

A proof satisfying both properties is said to be correct, which is indicated by the flag [preuve_correcte].

In all cases, the output file contains an “annotated proof” i.e. a proof in which each reasoning step is followed by an annotation indicating whether the step is correct or not. If the proof step is not correct, additional analysis are performed in order to extract as much information as possible from the proof. The system tries to compute why the proof is incorrect: for example HOARD-ATINF uses abduction algorithms for detecting missing hypotheses, checks non-degeneracy conditions, tries to identify missing axioms etc. All these information are sent to the agents in order to help diagnosis (note that no didactic decisions are made at this step: it is up to the Baghera plate-form to figure out whether the proof can be accepted, refused, or if an interaction must be started with the student in order to clarify things).

The systems also detects redundant and useless steps, and identifies analogies in the subpart of the proof (which may help to “structure” the proofs by introducing lemmas).

The interested reader can consult the HOARD-ATINF manual (Peltier 2001) for more details about the different functionalities.

Example 3 *Here is an example of output file:*

```
milieu(I,segment(A,B))
[correct]
milieu(J,segment(A,C))
[correct]
parallele(segment(B,C),segment(I,J))
milieu(I,segment(A,B))
milieu(J,segment(A,C))
milieux
[correct]
milieu(K,segment(B,C))
[correct]
parallele(segment(A,C),segment(I,K))
milieu(I,segment(A,B))
milieu(K,segment(B,C))
milieux
[analogue]
parallele(segment(B,C),segment(I,J))
% this indicates that this step is analogous to the
% previous one
[correct]
parallelogramme(I,J,C,K)
arallele(segment(I,J),segment(B,C))
parallele(segment(I,K),segment(A,C))
parallelogramme
% here the system detects that some further information
% are necessary for carrying out this inference
[hypotheses_omises]
milieu(K,segment(C,B))
milieu(J,segment(C,A))
$
% since this hypotheses hold, the proof is nevertheless accepted
```

```
% though not correct from a logical point of view
[conclusion_prouvee]
[preuve_coherente]
[preuve_correcte]
```

4.3.1.5 Analogy detection

The importance of analogy has long since been highlighted in all kinds of reasoning. However, analogy-related studies in the field of automated reasoning are surprisingly not numerous. The ignorance of analogy by theorem provers was early identified as a major drawback for their performances (Bledsoe, 1977) (Polya, 1973) and good use of analogy is still considered as a challenge in automated deduction (Wos, 1988).

In the present work, we are not really interested in using analogy for improving the efficiency of the prover, but we are rather interested in the *intrinsic* value of analogical reasoning, for a better understanding of proofs. The main goals are to be able to verify the analogies proposed by a student and to suggest analogies in a student's proof. Analogy detection is especially important for the presentation and structuring of proofs: detecting analogies into the proof is a way of introducing lemmas, thus making them shorter and improving their readability.

The analogy reasoning performed by HOARD-ATINF is mainly based on the method developed in for detecting and using analogies in resolution proofs. In this section, we briefly recall the basis of this approach, and we particularly emphasizes how it has been adapted to geometric reasoning.

The proposed method relies on generalization techniques. It can be divided into two steps:

1 **Generalization step**. It occurs just after proving a (new) formula (i.e. a theorem). The formula is transformed into a *more general formula schema*. Predicate and function symbols are replaced by higher-order variables, and specific generalization rules are applied, in order to transform the considered formula into a more general one, while preserving its proof. The idea is to try to find a formula which is *more general* than the original one but that admit a "similar" (i.e. with the same structure) proof. In some sense, the generalization algorithm infers, from a given proof, information useful for a larger class of problems.

2 **Matching step**. Then the system tries to compare the problem at hand to the previously generalized formulae in order to detect potential similarities. If such similarities do not exist then the system tries to infer *lemmas* that have to be proven in order to complete the analogy. The calculus for solving such matching problems has been presented in Défourneaux et Peltier (1997a, 1997b). It is based on higher-order unification techniques.

We refer to (Défourneaux *et al.* 1998) (Défourneaux et Peltier 1997a, 1997b) for a detailed presentation of this approach. Here, we only emphasize the main modifications that have been introduced into the algorithms in order to deal more easily with geometry and we give an example of use of the presented techniques.

- First, it is clear that some of the function and predicate symbols are part of the language, hence should not be generalized at all (i.e. should not be replaced by higher-order variables). For example, it would not make any sense, from a geometric point of view, to detect analogies by replacing the predicate "*parallel*" by the predicate "*perpendicular*" in a formula. Therefore, the application of the generalization rules in (Défourneaux *et al.* 1998) should be carefully controlled. Only function symbols without any intended semantic should be generalized. Though it reduces the number of potential similarities, it also has the big advantage to strengthen the constraints on the matching problem on step 2, thus making the matching process much more easy.
- Second, the matching algorithm should be performed modulo the particular theories of geometry predicates (commutativity, circularity and parallelism).

The output file can contain three distinct answers, with the following meanings:

- `correct`. The analogy is correct. The proof (obtained by instanciating the generalized proof generated from the source proof) is given afterward.
- `partielle`. The analogy is only partial. The proof is provided, together with the list of missing hypotheses (i.e. the assertions that could not be proven).
- `non_deductible`. Indicate that the theorem could not be proven.

Other features are provided in order to combine reasoning by analogy with standard deduction. In this cases, HOARD-ATINF tries to reconstruct the proof (in case the analogy is partial) by checking whether the lemmas are provable from the considered hypothesis (see the user manual for details).

Example 4
```
[traitement]
analogie
[hypotheses]
triangle(A,B,C)
symetrie_centrale(P1,P,A)
symetrie_centrale(P2,P1,B)
symetrie_centrale(P3,P2,C)
milieu(I,segment(P,P3))
tous
[conclusion]
parallelogramme(A,B,C,I)
[preuve]
milieu(I,segment(A,B))
$
milieu(J,segment(B,C))
$
milieu(K,segment(C,D))
$
milieu(L,segment(A,D))
$
parallele(segment(I,J),segment(A,C))
milieu(I,segment(A,B))
milieu(J,segment(B,C))
milieux
$
parallele(segment(L,K),segment(A,C))
milieu(L,segment(A,D))
milieu(K,segment(C,D))
milieux
$
parallele(segment(I,J),segment(L,K))
parallele(segment(I,J),segment(A,C))
parallele(segment(L,K),segment(A,C))
transitivite_parallelisme
$
parallele(segment(L,I),segment(B,D))
milieu(I,segment(A,B))
milieu(L,segment(A,D))
milieux
$
parallele(segment(J,K),segment(B,D))
milieu(J,segment(B,C))
milieu(K,segment(C,D))
milieux
$
```

```
parallele(segment(I,L),segment(J,K))
parallele(segment(I,L),segment(B,D))
parallele(segment(J,K),segment(B,D))
transitivite_parallelisme
$
parallelogramme(I,J,K,L)
parallele(segment(I,J),segment(L,K))
parallele(segment(L,I),segment(J,K))
parallelogramme
$
```

In this example, the analogy cannot be established without using the definition of the symmetry in order to prove that A,B,C are the respective midpoints of [PP1], [P2P1], [P3P2].

```
[correct]
parallelogramme(A,B,C,I)
parallele(segment(A,B),segment(I,C))
parallele(segment(I,A),segment(B,C))
$
parallele(segment(A,B),segment(I,C))
parallele(segment(A,B),segment(P,P2))
parallele(segment(I,C),segment(P,P2))
$
parallele(segment(A,B),segment(P,P2))
milieu(A,segment(P,P1))
milieu(B,segment(P1,P2))
$
symetrie_centrale(P,P1,A)
$
milieu(A,segment(P,P1))
symetrie_centrale(P,P1,A)
symetrie_centrale
$
symetrie_centrale(P2,P1,B)
$
milieu(B,segment(P2,P1))
symetrie_centrale(P2,P1,B)
symetrie_centrale
$
parallele(segment(I,C),segment(P,P2))
milieu(I,segment(P,P3))
milieu(C,segment(P2,P3))
$
milieu(I,segment(P,P3))
$
symetrie_centrale(P3,P2,C)
$
milieu(C,segment(P3,P2))
symetrie_centrale(P3,P2,C)
symetrie_centrale
$
parallele(segment(I,A),segment(B,C))
parallele(segment(A,I),segment(P1,P3))
parallele(segment(B,C),segment(P1,P3))
$
parallele(segment(A,I),segment(P1,P3))
milieu(A,segment(P,P1))
milieu(I,segment(P,P3))
$
parallele(segment(B,C),segment(P1,P3))
milieu(B,segment(P1,P2))
milieu(C,segment(P2,P3))
```

4.3.1.6 Automated Model Building

This functionality allows to construct automatically a geometrical figure satisfying a given set of geometrical assertions. The system may return 3 different answers:

- `[figure]`. Indicate that a figure has been built. The figure is provided by specifying the coordinate of all the points contained in it (of course more complex objects such as lines may be specified by giving a set of points belonging to them).
- `[non_trouvee]`. Means that the system did not find any figure (due to time limitation or any other reason).
- `[incorrect]`. Means that there is no figure satisfying the desired property (because the properties are contradictory).

4.3.2 Integration of HOARD-ATINF to the Baghera platform – a case study

In this section, the way HOARD-ATINF is actually integrated to the Baghera platform will be presented, followed by a few perspectives for further investigations of the prover capabilities. For the sake of clarity, we will analyse in detail one case to illustrate the most important issues arising from this integration.

4.3.2.1 Translation of a learner's proof into HOARD-ATINF language

Let us consider the following problem:

Let ABC be an equilateral triangle. Point A' is the symmetric of A with respect to the line d. L is the midpoint of segment [AB], M is the midpoint of segment [BC], N is the midpoint of segment [AC]. P is the intersection point of the line (LM) and the line (CA') and O is the intersection point of the line (MN) and the line (BA') (see Figure 18).
What is the symmetric of the segment [MN] with respect to the line d? How can you prove it?

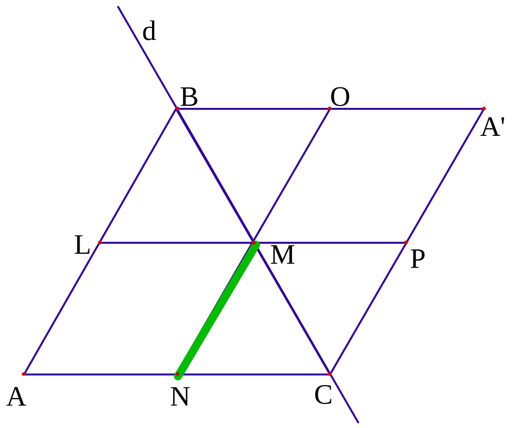


Figure 18. Geometrical figure corresponding to the problem.

Follows one student's solution of the problem:

1. The symmetric of [NM] with respect to d is [MO].
2. N is the midpoint of [AC].
3. O is on the same line as N but at the opposite of the figure and on the other side of the line (d).
4. M is on the line (d).
5. Therefore, M is its own symmetric with respect to (d).
6. O is the opposite of N or the symmetric of this point.
7. Therefore, [MN] is the symmetric of [MO].

In order to make this proof treatable by HOARD-ATINF, some necessary modifications, mainly in the proof structure, must have been done. The statements 2-7 have first been grouped into steps as required by the syntax of the prover. For each step, the inference rule (operator) used either explicitly or implicitly by the student, has been identified. The resulting proof and its translation into HOARD-ATINF language are shown in the table below.

Student's proof steps	*Translation into HOARD-ATINF language*

2. N is the midpoint of [AC].
3. O is on the same line as N but at the opposite of the figure and on the other side of the line (d).
6. O is the opposite of N or the symmetric of this point.
	`symetrie_axiale(O, N, droite(B,C))`
	`milieu(N, segment(A,C))`
	`est_intersection(M,droite(O,N),droite(B,C))`
	`sym_oppose`

4. M is on the line (d).
5. Therefore, M is its own symmetric with respect to (d).
	`symetrie_axiale(M,M,droite(B,C))`
	`est_sur_droite(M,droite(B,C))`
	`symetrie_axiale`

7. Therefore, [MN] is the symmetric of [MO].	`symetrie_axiale(segment(M,N),segment(M,O),droite(B,C))`
	`symetrie_axiale(M,M,droite(B,C))`
	`symetrie_axiale(O,N,droite(B,C))`
	`sym_ax_align`

Table 2. Student's proof and its translation into HOARD-ATINF language.

According to the proof syntax required by HOARD-ATINF, each step of a proof must include an operator which is an assertion connecting the conclusion with the list of hypotheses (see the section 4.3.1). Such an assertion may be either a definition (e.g., *symetrie_axiale* designates the definition of a reflection), a theorem or a property (e.g., *sym_ax_align* is the name of the property stating that a reflection preserves colinearity). Many operators corresponding to definitions and well-known properties of geometric objects are built-in the prover, but one does not need to restrict himself to this list. It is possible to define and to work with new operators, in particular those used by the students. Such operators do not need to be valid. In our example, the operator *sym_oppose*, that is obviously incorrect in a general case, has been added.

As it was mentioned previously (section 4.3.1), operators are not checked for their validity. However, in the proof verification, we need to consider only valid operators, otherwise any statement could be proven and no information of interest about the correctness and coherence of the proof could be obtained. Therefore, when defining a new operator, its domain of validity (i.e. the conditions under which the operator is logically valid) must be determined. In the above mentioned example, the operator *sym_oppose* is valid when M is the midpoint of the segment [ON] and [ON] is perpendicular to the line (BC) (see figure 2). These conditions are included to the definition of the operator.

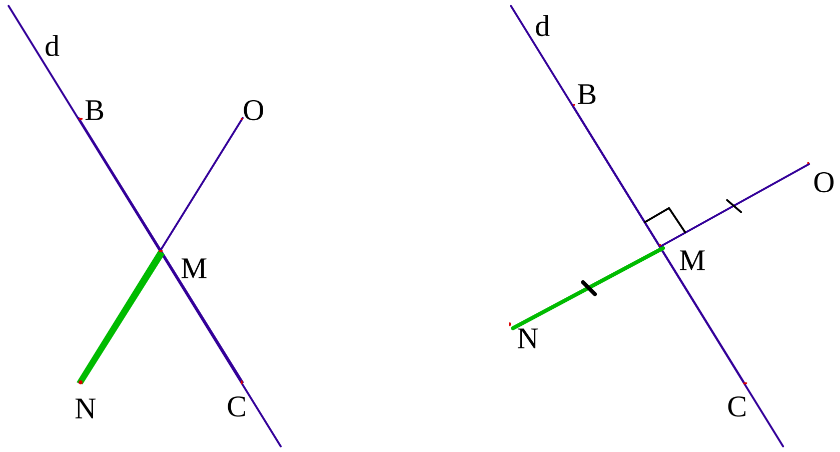


Figure 19. The *sym_oppose* operator is valid only if [ON] is perpendicular to d, and if NM=MO.

4.3.2.2 HOARD-ATINF feedback

The automatic proof verification can be called from the Baghera platform once the learner has finished his problem solving activity. An input file containing the givens of the problem (hypotheses), the learner's answer (the conclusion to be proven), operators used in the proof and the proof itself, is sent to HOARD-ATINF. The verification consists in checking whether the proof is correct or not. The output file contains the proof accompanied with local annotations related to each particular step independently of the rest of the proof, and global annotations indicating whether the whole proof is correct (presented in section 4.3.1).

In the above mentioned example, the output file obtained from HOARD-ATINF is shown in the second column of the Table 3.

Student's proof	*Automatic treatment by HOARD-ATINF*

2. N is the midpoint of [AC].
 	
3. O is on the same line as N but at the opposite of the figure and on the other side of the line (d).
 	
6. O is the opposite of N or the symmetric of this point.
	`milieu(N,segment(A,C))`
	[**correct**]
	
	`est_intersection(M,droite(O,N),droite(B,C))`
	[**correct**]
	
	`symetrie_axiale(O,N,droite(B,C))`
	`est_intersection(M,droite(O,N),droite(B,C))`
	`sym_oppose`
	[**hypotheses_manquantes**]
	`milieu(M,segment(O,N))`
	`perpendiculaire(segment(O,N),segment(B,C))`

4. M is on the line (d).
 	
5. Therefore, M is its own symmetric with respect to (d).
	`est_sur_droite(M,droite(B,C))`
	[**correct**]
	`symetrie_axiale(M,M,droite(B,C))`
	`est_sur_droite(M,droite(B,C))`
	`symetrie_axiale`
	[**correct**]

7. Therefore, [MN] is the symmetric of [MO].	`symetrie_axiale(segment(M,N),segment(M,O),droite(B,C))`
	`symetrie_axiale(M,M,droite(B,C))`
	`symetrie_axiale(O,N,droite(B,C))`
	`sym_ax_align`
	[**correct**]

	[**conclusion_prouvee**] [**assertions_non_utilisees**] `milieu(N,segment(A,C))` [**erreur_detectee**]

Table 3. Student's proof and its automatic verification by HOARD-ATINF.

In the first step of the proof, an error was detected; it is indicated by the flag [hypotheses_manquantes], meaning that although the operator *sym_oppose* allows deducing the conclusion, it cannot be used because some hypotheses necessary for the operator to be valid are not satisfied. The list of these hypotheses is given after the flag.

The global annotations indicate that the conclusion mentioned in the input file has been proven, but there has been an error detected. Moreover, an information about statements that were cited as hypotheses but were not used in the proof [assertions_non_utilisees] is also provided.

The output file with annotations is sent back to the Baghera platform to be treated by the system. The extracted information may be used in the diagnosis process, as well as in the didactic decisions making.

4.3.2.3 HOARD-ATINF feedback information treatment

Although the HOARD-ATINF feedback treatment is not implemented yet in the Baghera platform, we will present here a few issues showing the significance of the obtained information. For the sake of clarity, we will restrict ourselves to the analysis of the above mentioned example.

As it was mentioned above, the flag `[hypotheses_manquantes]` means that some hypotheses necessary for an operator to be valid are missing in the proof. This does not necessarily mean that the learner did not take these hypotheses into account. Some of these might have been considered but not stated explicitly, especially those that can be read from the figure accompanying the problem. Such hypotheses can be considered as implicit. This is the case of the hypothesis assuming that M is the midpoint of the segment [ON]. This hypothesis is true and it might have been considered by the student as obvious, so that s/he found it useless to make it explicit. On the other hand, a hypothesis that is identified as missing and that turns out to be wrong, indicates an erroneous conception in the student because in this case, s/he did actually use an operator outside its domain of validity. This is the case of the hypothesis assuming that the segment [ON] is perpendicular to the line d. Therefore, it is necessary first to verify whether the hypotheses listed as missing ones are satisfied or not in the given problem.

Furthermore, the information about the nature of the hypotheses identified as missing is significant and useful for providing an adequate counter-example or suggesting the characteristics of a problem to be proposed to the student to help her/him become aware of the limits of her/his conception.

4.3.2.4 Further investigations (perspectives)

Controls

In the approach presented above, we have investigated what kind of information can be obtained from the HOARD-ATINF feedback. It turns out that considering the *operators* used by a learner gives the possibility for hypothesizing an erroneous conception held by the learner. In order to deepen the insight in the nature of this conception, a further analysis of the learner's work is needed, mainly by considering the *controls* s/he might have used or not in the proof. In what follows, the mains ideas of the approach based on controls are illustrated on the above mentioned example.

When solving a problem, a learner is using several controls, in an explicit or an implicit way. These controls are used to take decisions, make choices, and judge on the use of an operator or on the state of a problem. They can be considered as properties ascribed by the learner to mathematical objects that are not necessarily true in general. The fact that a learner uses controls that are not valid, or s/he does not use controls that are valid, is an indicator of an erroneous conception, and the nature of these controls can sometimes help in the characterization of the conception.

At the implementation level, a list of controls that can be possibly used by learners can be established and the learner's work can be checked by HOARD-ATINF in order to identify controls that might have been used by the learner and those that are absent.

Let us go back to the previous example. In the table below (Table 4) the analysis of the student's work is in terms of controls.

Student's proof steps	*Controls used explicitly or implicitly*
2. N is the midpoint of [AC]. 3. O is on the same line as N but at the opposite of the figure and on the other side of the line (d). 6. O is the opposite of N or the symmetric of this point.	C1: The image of a point by a reflection over a line is a point. C2: The image of an object by a reflection over a line is on the other side of the axis.
4. M is on the line (d). 5. Therefore, M is its own symmetric with respect to (d).	C1 C3: The image by a reflection over a line of a point lying on the axis is the point itself.
7. Therefore, [MN] is the image of [MO] by the reflection.	C4: The image of a segment by a reflection over a line is a segment. C5: A segment and its image by a reflection over a line lie on the same line.

Table 4. Controls identified in the student's work.

In the Table 5, a non-exhaustive list of controls is given and the results of the verification (simulated "manually") are presented. The use of an incorrect control, or the absence of a correct control, indicates one or more conceptions the learner possibly holds about reflection.

	Controls	*Used?*	*Conception*[4]
TRUE	The symmetric of a point with respect to a line is a point.	Yes	
	The symmetric of a segment with respect to a line is a segment.	Yes	
	A reflection preserves colinearity.	Yes	
	The symmetric of an object with respect to a line is on the other side of the axis.	Yes	
	The symmetric of a point on the axis is the point itself.	Yes	
	The line joining a point and its symmetric is perpendicular to the axis.	**No**	**Reflection or Oblique Symmetry or Parallelism**
FALSE	**A segment and its symmetric with respect to a line are parallel.**	**Yes**	**Reflection or Parallelism**
	The line joining a point and its symmetric is horizontal.	No	
	A segment and its symmetric with respect to a line lie on the same line.	**Yes**	**Central symmetry or Oblique Symmetry**

Table 5. Identification, based on controls, of conceptions about reflection possibly held by the student.

New conceptions

We have presented the way HOARD-ATINF can be used to make a local analysis, i.e. an analysis of a given proof in terms of operators and controls used by a learner. On the other hand, we can also imagine HOARD-ATINF making a global analysis in terms of manipulation of conceptions. The

[4] Four conceptions about reflection identified by mathematics education researchers are considered here: Central symmetry, Oblique symmetry, Parallelism, Reflection, see §6.2 of the report.

fact that the prover does not check for a validity of user-defined operators and is thus able to work with any operators, valid or not, we can define a set of operators used by a learner and ask the prover to find proofs for a given set of problems. In this way, a logical coherence of this set of operators can be verified. This kind of analysis may perhaps reveal new conceptions. Further investigations are planned to be carried out in these directions.

References

Axtell, R., Epstein, J., Young, H.P. (2000) The emergence of classes in a multi-agent bargaining model. Center on Social and Economics Dynamics, Working paper no.9, February 2000.

Balacheff, N. (2000a) Teaching, an emergent property of eLearning environments. In: Conférence IST 2000. Nice, France. Available at http://www-didactique.imag.fr/Balacheff/TextesDivers/IST2000.html.

Balacheff, N. (2000b) A modelling challenge: untangling learners' knowing. In: Journées Internationales d'Orsay sur les Sciences Cognitives: L'apprentissage, JIOSC2000, Paris.

Balacheff N., Gaudin N., (2002) Modelling students conceptions - the case of functions. Mathematical Thinking and Learning – an International Journal. (to appear).

Bledsoe W. W. (1977) Non-resolution theorem proving. *Artificial Intelligence*, 9:1—35.

Bonabeau, E.; Dessalles, J.; Grumbach, A. (1995) Characterizing Emergent Phenomena (1): A critical review. In: Revue Internationale de Systémique, vol.9, nº3, pp. 327-346.

Caferra R., Peltier N., and Puitg F. (2000) Emphazing human techniques in geometry automated theorem proving: a practical realization. In J. Richter-Gebert and D. Wang, editors, *Workshop on Automated Deduction in Geometry. Zurich, Switzlerland*. Springer, LNAI 2061, September 2000, pp. 38-59.

Capuano, N.; Marsella, M.; Salerno, S. (2000) ABITS: An Agent Based Intelligent Tutoring System for Distance Learning. In : Proceedings of the International Workshop in Adaptative and Intelligent Web-based Educational Systems. Available at http://virtcampus.cl-ki.uni-osnabrueck.de/its-2000/.

Chan, T.W. (1996) Learning Companion Systems, Social Learning Systems, and Intelligent Virtual Classroom. Invited Talk, World Conference on Artificial Intelligence in Education, Washington, DC, USA, Journal of Artificial Intelligence in Education, Vol. 7, No. 2, pp. 125-159.

Chou H., Gao X. S., and Zhang J.-Z. (2000) A deductive database approach to automated geometry theorem proving and discovering. *Journal of Automated Reasoning*, 25(3):219—246.

Conati, C. Klawe, M. (2000) Socially Intelligent Agents to Improve the Effectiveness of Educational Games. In: Proceedings of AAAI Fall Symposium on Socially Intelligent Agents - The human in the loop. 2000. Available at http://www.cs.ubc.ca/~conati/publications.html.

Défourneaux G., Bourely C., and Peltier N. (1998) Semantic generalizations for proving and disproving conjectures by analogy. *Journal of Automated Reasoning*, 20(1 & 2):27—45.

Défourneaux G. and Peltier N. (1997a) Analogy and abduction in automated reasoning. In M.~E. Pollack, editor, *Proceedings of IJCAI'97, Nagoya, Japan,* August 23—29 1997.

Défourneaux G. and Peltier N. (1997b) Partial matching for analogy discovery in proofs and counter-examples. In W. McCune, editor, *Proceedings of CADE 14*. Springer LNAI 1249, July 1997.

Demazeau, Y. (1995) From Interactions to Collective Behaviour in Agent-Based Systems. In: Proceedings of the 1st. European Conference on Cognitive Science. Saint-Malo, France, 1995.

Fitting M. (1990) *First-Order Logic and Automated Theorem Proving*. Texts and Monographs in Computer Science. Springer-Verlag.

Forrest, S. (1990) Emergent Computation. Physica D 42. Elsevier Science Publishers, North-Holland.

Forrest, S.; Miller, J. (1990) Emergent Behavior in Classifier Systems. In: Physica D 42, pp. 213-227. Elsevier Science Publishers, North-Holland.

Holland, J.H. (2000) Emergence from chaos to order. UK: Oxford University Press, pp. 258.

Huget, M-P; Koning, J-L.; Bergia, L. (2000) Une Plate-forme pour l'ingénierie des protocoles d'interaction. Application au projet de télé-enseignement Baghera. In: Systèmes multi-agents: méthodologie, technologie et expérience, 8emes Journées Francophones Intelligence Artificielle Distribuée et Systèmes Multi-Agents, JFIADSMA'00, S Pesty & C. Sayettat-Fau (eds.), pp 297-301.

Johnson, W. L.; Rickel, J.W.; Lester, J.C. (2000) Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. In: The International Journal of Artificial Intelligence in Education, 11, pp.47-78.

Machado, I.; Martins, A.; Paiva, A. (1999) One for All and All for One A leamer modelling server in a multi-agent platform. In: Proceedings of the Seventh International Conference on User Modelling, pp. 211-221.

Maxion, R.A. (1990) Toward diagnosis as an emergent behavior in a network ecosystem. In: Physica D 42, pp. 66-84. Elsevier Science Publishers, North-Holland.

Peltier, N. (2001) *Hoard-atinf — manuel d'utilisation.* Research Report. Available on http://www-leibniz.imag.fr/ATINF/Nicolas.Peltier/

Polya, G. (1973) *How to Solve It, a New Aspect of Mathematical Method.* Princeton University Press. Second Edition.

Py, D. (1996) Aide à la démonstration en géométrie : le tuteur Mentoniezh. *Sciences et Techniques Educatives*, volume 3 numéro 2, Hermès.

Rao, A.; Georgeff, M. (1995) BDI Agents: From Theory to Practice. Proceedings of the *First International Conference in Multi-Agents Systems*, San Francisco, USA.

Ritter, S. (1997) Communication, Cooperation and Competition among Multiple Tutor Agents. In: *Proceedings of the 8th World Conference of Artificial Intelligence in Education.* Du Boulay,B.; Mizoguchi, R. (Eds). IOS Press.

Ritter, S.; Koedinger, K.R. (1996) An architecture for plug-in tutor agents. *Journal of Artificial Intelligence in Education*, Vol.7, pp. 315-347.

Robinson J. (1965) Automatic deduction with hyperresolution. Intern. Journal of Computer Math., 1:227—234.

Sandholm, T.W. (1999) Distributed Rational Decision Making. In: Multiagent Systems: A Modern Introduction to Distributed Artificial Intelligence, -258.

Sawyer, R.K. (2000) Simulating Emergence and Downward Causation in Small Groups. In: S. Moss and P.Davidsson (Eds.) Multi-Agent-based Simulation. Second International Workshop, MABS 2000, LNAI 1979, pp.49-67. Springer-Verlag Berlin Heidelberg.

Schreiber, D. (2000) The emergence of parties: an agent-based simulation. In: The Annual Meeting of the Midwestern Political Science Association, Chicago, Illinois.

Schweitzer, F.; Zimmermann, J. (2001) Communication and Self-Organisation in Complex Systems: A Basic Approach. In: Knowledge, Complexity and Innovation Systems (Eds.M.M.Fischer, J.Frohlich) Advances in Spatial Sciences. Springer, Berlin, pp.275-296.

Shehory, O.; Kraus, S. (1995) Task allocation via coalition formation among autonomous agents. In: Proceedings of IJCAI'95, Montreal. pp.655-661.

Silveira, R.A.; Bica, F.; Viccari, R.M. (2000) JADE - Java Agents for Distance Education Framework. In: Proceedings of the Third Iberoamerican Workshop on Distributed Artificial Intelligence and Multi-Agent Systems. Alvares L.O.and Bazan A.L. (Eds.), Brazil, 2000. pp.112-122.

Stadler, B.M.R. (1999) Adaptative Platform Dynamics in Multi-Party Spatial Voting. Adv.Complex Systems vol.2, 101-116.

Vassileva, J.; Deters, R.; Greer, J. McCalla, G.; Bull, S.; Kettel, L. (2001) Lessons from Deploying I-Help. In: Proceedings of the Workshop Multi-Agent Architectures for Distributed Learning Environments, AI-ED'2001, San Antonio, TX, USA. pp. 3-11.

Vassileva, J.; Greer, J.; McCalla, G.; Deters, R. (1999) A Multi-Agent Approach to the Design of Peer-Help Environments. In: Proceedings of AIED'99, Le Mans, France, pp. 38-45.

Vriend, N.J. (1995) Self-Organization of Markets: An Example of a Computational Approach. Computational Economics, Vol. 8, No. 3, pp. 205-231.

Webber, C.; Pesty, S.; Balacheff, N. (2002) A multi-agent and emergent approach to learner modelling. In: ECAI 2002 - Proceedings of the 15th European Conference on Artificial Intelligence. F. van Harmelen (ed.), IOS Press, Amsterdam, pp.98-102.

Weiss, G. (Ed.) (1999). Multiagent Systems - A Modern Approach to Distributed Artificial Intelligence. The MIT-Press.

Wos L. (1988) Automated Reasoning, 33 Basic Research Problems. Prentice Hall.

Zapata-Rivera, J.D.; Greer, J. (2001) SMODEL Server: Student Modelling in Distributed Multi-Agent Tutoring Systems. In: Artificial Intelligence in Education. J.D.Moore, C.L.Redfield and W.L.Johnson (Eds.). IOS Press, pp.446-455.

5 Evaluation methodology

Leader of the task: Maria-Alessandra Mariotti

Authors: Michele Cerulli, Mirko Maracci, Maria-Alessandra Mariotti, Federica Olivero

Sophie Soury-Lavergne, Rosamund Sutherland

Baghera, as an intelligent tutoring system, is supposed to be used both for diagnoses of the system user/machine in terms of the cK¢ model, and for teaching interventions (see description of the project). In this phase of the project we cannot focus on teaching intervention because such aspects have not yet been implemented since it required to have available robust student modelling and diagnostic procedures. So, we focus first on diagnosis issues, with the aim of tackling the former aspect in the near future.

Starting from an interaction user/machine, where the user is involved in the solution of a mathematical problem, Baghera is expected to identify the set of possible conceptions (see cK¢, section 2.1) emerging from the user/machine interaction. The aim of the evaluation is to check whether this is the case and the efficiency of the emergence process. The main idea of the evaluation process is to compare the diagnosis produced by Baghera with the diagnoses produced by mathematics education researchers. In order to do that, we decided to carry out two types of analysis, i.e. a *Didactical analysis* carried out by didacticians (mathematics education researchers) and a *Baghera 1 analysis* carried out by the machine. Both analyses produce a diagnosis in terms of conceptions.

This section describes in detail the methodology which was implemented for the evaluation of Baghera. The process of evaluation is represented in Figure 20.

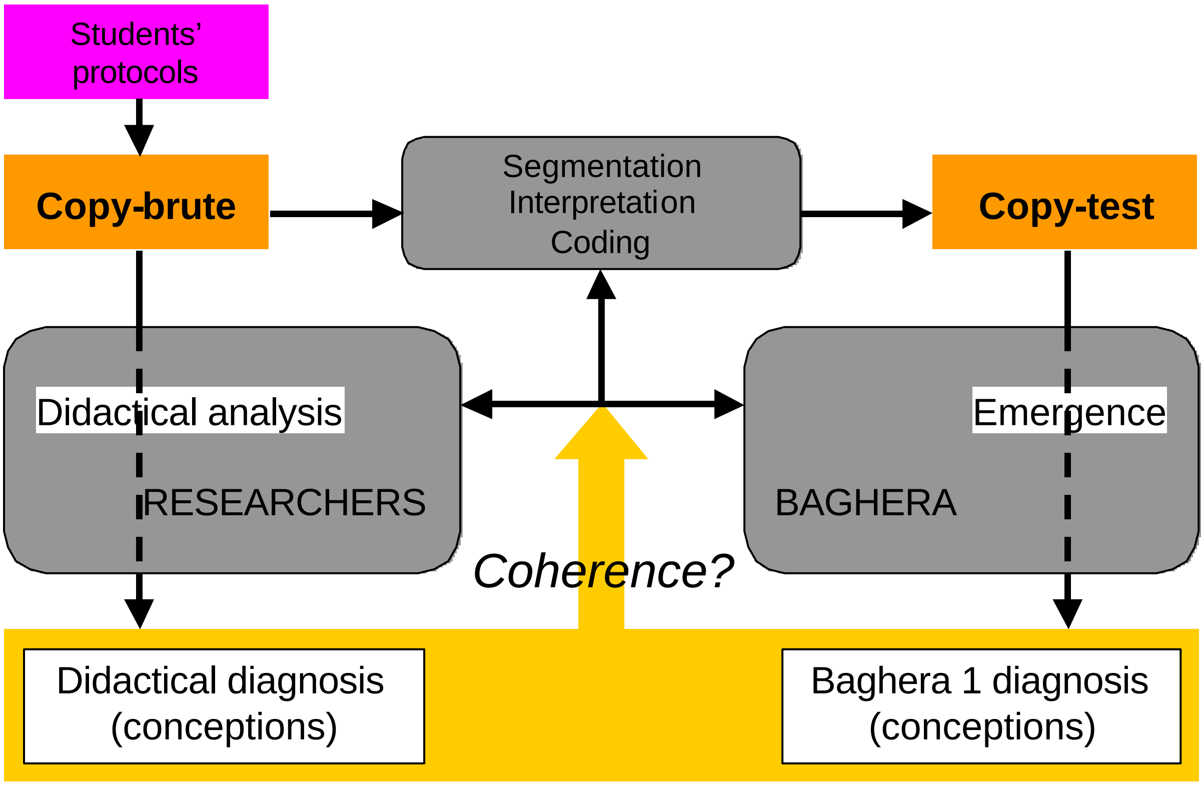


Figure 20. The evaluation methodology

In order to take into account time constraints and at the same time to have data treatable both by humans experts and by Baghera platform, we decided to collect written protocols produced by students. The main point was to be able to collect a large amount of protocols, representing diversity of students productions.

Given a set of students' written protocols two parallel analyses were carried out:

- the first analysis was carried out by the Baghera platform mainly based on the theory of emergence (from now on referred to as "*Baghera 1 analysis*");
- the second analysis was carried out by mathematics education researchers (from now on referred to as "*Didactical analysis"*).

The final objective was the comparison of the outputs of the two analyses, in terms of coherence in diagnosing conceptions.

5.1 Baghera 1 analysis

The *Baghera 1 analysis* consisted of a procedure that can be schematised into 3 steps (Figure 21). Starting from the written students' protocols (referred to as *"copy-brute"*) a process of "translation" led to a form of texts (named *"copy-test"*) that could be given as an input to Baghera. This step is done by humans at this stage, it will later consist of the users input. The resulting *copy-test* were analysed by Baghera which then produced a diagnosis according to the cK¢ model which is implemented and based on the theory of emergence. The product of this analysis, called the

Baghera 1 diagnosis, is the identification of one, or more, conceptions associated with each of the protocols analysed.

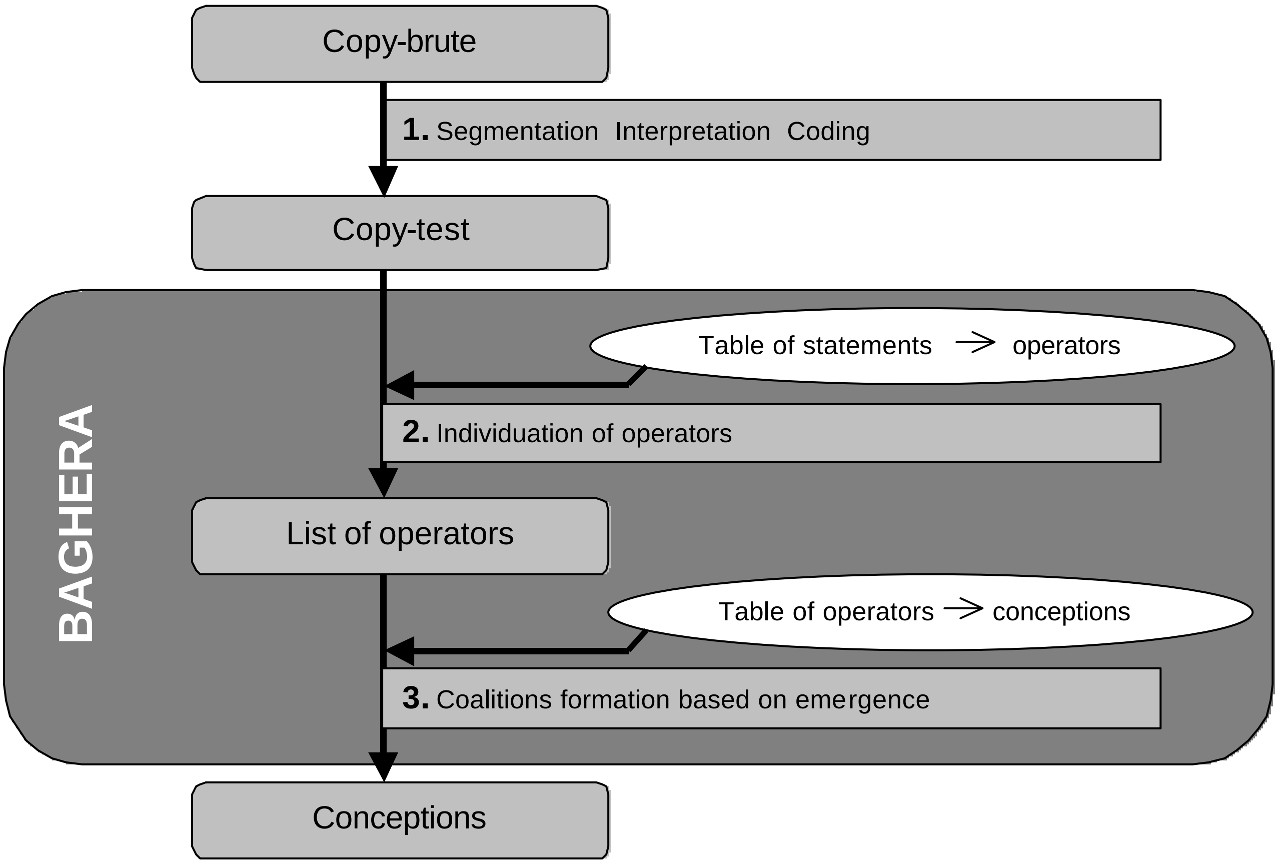


Figure 21. The *Baghera 1 analysis*

Step 1: segmentation, classification and translation of copy-brute into copy-test

Given a *copy-brute*, a *copy-test* is produced by segmentation into meaningful statements, classification of statements and translation into a specific code. This process, done by humans, consisted of:

- segmentation of the protocols into statements according to their syntactical structure;
- reorganisation of the statements into steps of proof, according to their status (i.e. hypothesis, conclusion or inference rule);
- coding of statements into the Baghera language, taking into account the constraints on expression of the Baghera interface.

Step 2: identification of operators and controls

The obtained *copy-test* is processed by Baghera which extracts a set of *operators & controls* to be associated with the *copy-test*. This process is based on a table of associations (statements/operators; statements/controls) that has been previously produced by humans.

Step 3: identification of conceptions

Starting from the list of operators and controls (obtained in step 2), and from a table of associations (operators & controls / conceptions), Baghera identifies one or more conceptions. At this stage, and for the purpose of the evaluation, the table of association is produced and inserted into the system by humans. The identification of conceptions is done entirely by the machine through the theory of emergence.

5.2 Didactical analysis

The human analysis was carried out on the *copy-brute*; the aim of this analysis was to identify one or more conceptions associated with each protocol. This analysis was carried out independently by the mathematics education researchers. The researchers were divided into three groups according to the country (city) they come from (Grenoble, Pisa and Bristol) and the three groups analysed the protocols independently. Even though we were all working within the cK¢ framework, different methods were used in the analyses of the protocols which could lead to different results. Therefore we decided to work independently in order to maximise the opportunities of potential disagreement amongst the groups and to provide internal validation of the analysis.

The results obtained by the three groups where then compared and this comparison was taken as an element of evaluation. This comparison produced a *didactical diagnosis* (which consists of one, or more, conceptions associated with a given protocol in the cases when all the didactical analyses agree).

On the contrary to *Baghera 1 analysis*, the *didactical analysis* performed by the three research groups did not follow (by choice) a predefined procedure.

5.3 The process of comparison between Baghera 1 analysis and Didactical analysis

The *didactical diagnosis* was compared with the *Baghera 1 diagnosis* focusing on consistency and inconsistency between the results. A good functioning should give consistency. And this is what was expected. However, we were ready to deal with cases of inconsistency and to analyse why they occurred. In the case of inconsistency we developed a methodology aimed to compare the two processes of analysis (carried out by the humans and by the machine) in order to understand possible factors influencing the inconsistency, i.e. each single step of the *Baghera 1 analysis* is analysed in order to identify where inconsistency may have been originated.

Step 1.

This step is done by humans, and produces a structured collection of statements expressed in the language of Baghera and classified as hypothesis, conclusions or inference rule. Such abstraction of a given *copy-brute* may not be univocally determined due to linguistic ambiguities and to possible different interpretations of the content. This analysis is similar to the one which must be carried out during the *Didactical analysis*. As a consequence it may happen that the produced *copy-test* has a structure that does not match the structure of the given *copy-brute*. This may be seen as a possible cause for inconsistency, but it is also a first information about the constraints of expression that Baghera interface presents to its users.

Step 2.

For each given *copy-test*, the Baghera platform associates to each statement one or more operators and controls. The association is determined by a table of correspondences that is previously built and inserted into the machine by humans. It may happen that the operators (and controls) identified are not compatible with the interpretation of the *copy-brute* given by the researchers in the *Didactical analysis* and that some statements are associated with different operators or controls. Again this may be a cause for inconsistency, which depends on the a priori content of the table of associations statements/operators and statements/controls.

Step 3.

In this phase the input is the list of operators & controls previously obtained, that at this point we may assume to be compatible with the researcher's interpretation of the *copy-brute* (see steps 1 and 2). Given such a list, and given a table of associations operators&controls/conceptions, a mechanism, based on the emergence theory and implemented in Baghera, indicates one or more conceptions to be associated with each *copy-test*. More precisely, following the table each of the given operators or controls is associated with a set of conceptions, then, starting from this associations, the theory of emergence is used to choose the conceptions to be associated with the *copy-test*.

Thus the two following events may occur:

a. a single operator or control is associated with a set of conceptions and such association is not compatible with *Didactical analysis*, leading to inconsistency with such analysis
b. each single operator and each control is associated with sets of conceptions that are compatible with *Didactical analysis*, but conceptions identified by the emergence mechanism are nevertheless different from those identified by the *Didactical analysis*. In this case the cause of inconsistency may be attributed to the emergence mechanism.

Possible conclusions of the evaluation process

In conclusion, we may observe that the first two steps of the Baghera 1 analysis do not depend on the implementation of the system: step 1 is done by humans; step 2 is done by Baghera but consists of applying a table of associations that is produced by humans. In other words, when inconsistency with the *Didactical analysis* is identified in the first two steps, then the implementation of the cK¢ model in Baghera through the emergence theory is not invalidated.

For what concerns step 3, in the case of event *a*, then again inconsistency may be attributed to humans' work, and in principle it can be overcome by revising and changing the table of associations operators & controls / conceptions. On the contrary, if event *b* occurs then we may say that the emergence mechanism is the cause of inconsistency.

6 Evaluation process

Leader of the task: Rosamund Sutherland

Authors: Michele Cerulli, Mirko Maracci, Maria-Alessandra Mariotti, Federica Olivero
Sophie Soury-Lavergne, Rosamund Sutherland

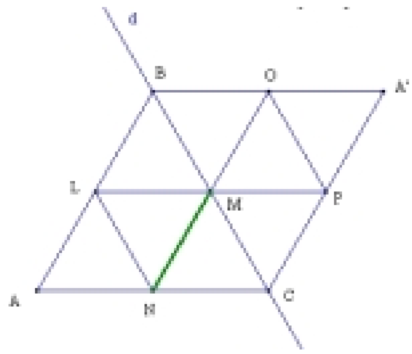
6.1. Data for the evaluation

After establishing the evaluation criterion in terms of evaluating the Baghera system through the comparison of its diagnosis with the diagnoses produced by researchers, we had to find data which could be processed both by the humans and by Baghera. A set of 28 protocols were taken from a collection of 173 protocols that had been produced within other projects, with aims relevant to Baghera ones. The considered protocols are students' written solutions of geometry proof problems, concerning 2D symmetries, solved in a paper and pencil environment (see Figure 26).

The problems used for the analysis are:

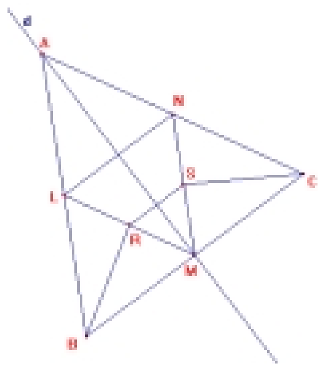
Problem 1

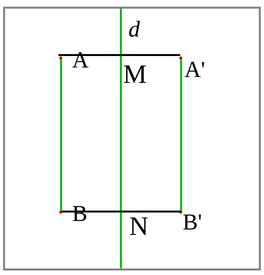
Let ABC be an equilateral triangle. Point A’ is the symmetric of A with respect to the line d. L is the midpoint of segment [AB], M is the midpoint of segment [BC], N is the midpoint of segment [AC]. P is the intersection point of the line (LM) and the line (CA’) and O is the intersection point of the line (MN) and the line (BA’).
What is the symmetric of the segment [MN] with respect to the line d? How can you prove it?



Problem 2

Let ABC be a triangle. The vertex C is the symmetric of point B with respect to the line d. The point L is the midpoint of segment [AB] and N is its symmetric with respect to the line d. Moreover, the point R is the midpoint of segment [LM] and the point S is the midpoint of segment [NM]. What is the nature of triangle RSM? Prove it.





Problem 3

Let segment [AB] be parallel to the line d. Let segment [A’B’] be the symmetric of [AB] with respect to the line d. Neither A, nor B belong to line d. What is the nature of quadrilateral ABB’A’? Prove it.

6.2 Didactical analysis

This section describes the part of the evaluation process consisting of the *Didactical analysis*. This is made of two steps: first analysis of the *copy-brute* carried out independently by three teams (Pisa, Bristol & Grenoble); then comparison of the results of the analyses and production of a *didactical diagnosis* which can be qualified by the level of agreement reached by the three team.

The didactical analysis of the *copy-brute* aimed to identify conceptions. From the literature in mathematics education, four possible conceptions were already identified (Grenier 1988; Tahri 1993).

Conception Reflection

It is the correct conception, which consist of taking into account the two proprieties of reflection: a segment formed by a point and its symmetric is perpendicular to the axis, a point and its symmetric are at equal distance to the axis (cf. Figure 22).

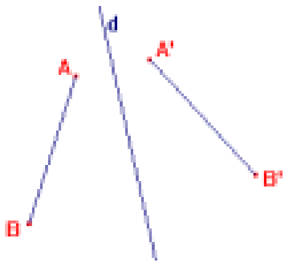


Figure 22. Reflection

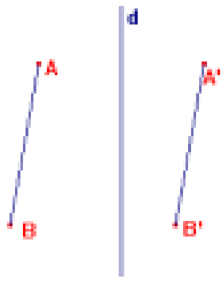


Figure 23. Parallelism

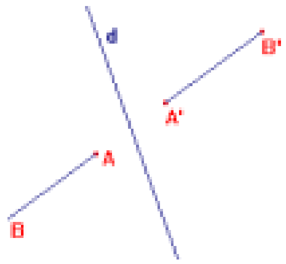


Figure 24. Central symmetry

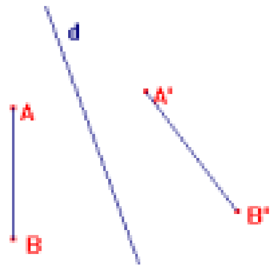


Figure 25. Oblique symmetry

Conception Parallelism

This conception consists in assuming that a segment and its symmetric are parallel and have the same length. As a consequence, a point and its symmetric are not necessarily at the same distance to the axis (cf. Figure 23).

Conception Central Symmetry

This conception deals with reflection as if it was a central symmetry. The procedure consist in choosing one or many points on the axis which play the role of the centre of the symmetry. As a consequence, the direction of the segment formed by a point and its symmetric is not necessarily perpendicular to the axis (cf. Figure 24).

Conception Oblique Symmetry

This conception consists in choosing a direction, not necessarily perpendicular to the axis, to support a point and its symmetric. The distances along this direction between a point and the axis and between the symmetrical point and the axis are equal. But the length of a segment and the length of its symmetric are not the same (cf. Figure 25).

The analysis was based on the cK¢ model. However, the identification of operators and controls was not strictly adhered to the formal definition, rather attention was paid more generally to the 'sense' linked to each conception.

The three teams followed a similar methodology of analysis which consisted of:

- analysis of the answer to the problem, referring to an a priori analysis of the problem which states the relation between student's answer, possible procedure of problem solving and theoretical explanation;
- analysis of the arguments given by the students to justify the answer. In particular, analysis of the proof structure: identification of proof steps, hypotheses, theorems and conclusions and suggestion of propositions not explicitly stated by students;
- identification of figures and configurations in the figure used by the student. In some cases possible elements not explicitly mentioned in the students' answer which could however be identified from the context (e.g. considering information related to the figure) were also taken into account;
- diagnosis of one or more conceptions recognizable in the solution of the problem and production of an explanation for that. Conception could be selected among those already identified or could be new ones.

6.2.1 EXAMPLE OF ANALYSIS OF A *COPY-BRUTE*

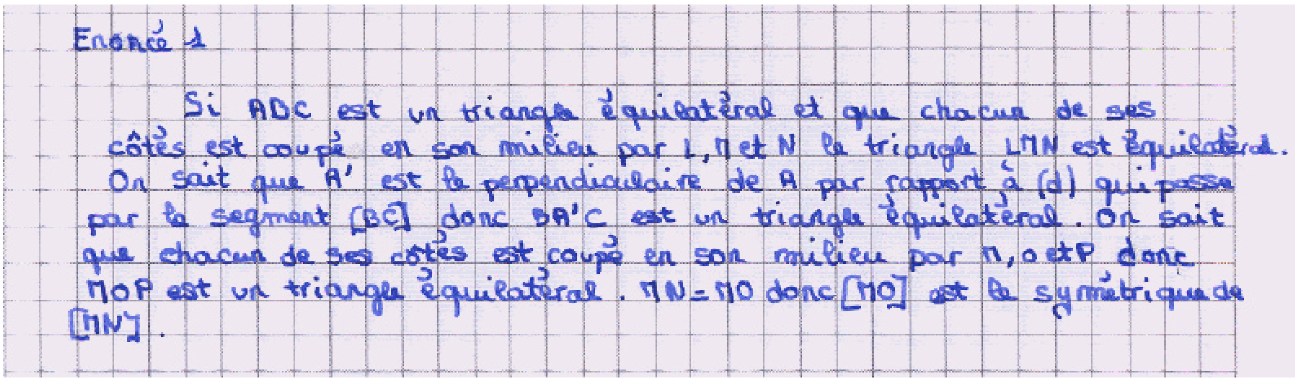

Enoncé 1
Si ABC est un triangle équilatéral et que chacun de ses côtés est coupé en son milieu par L, M et N le triangle LMN est équilatéral. On sait que A' est le perpendiculaire de A par rapport à (d) qui passe par le segment [BC] donc BA'C est un triangle équilatéral. On sait que chacun de ses côtés est coupé en son milieu par M, O et P donc MOP est un triangle équilatéral. MN = MO donc [MO] est le symétrique de [MN].

Figure 26. Copy-brute n°119 about problem 1.

Grenoble diagnosis. Central Symmetry. Answer of the student: segment [MO]. Explanation: The last step seems to be independent from the rest of the proof. This last step uses only the equality of the length of the two segments to state that they are symmetric. The three third steps may be a try to confirm global forms that are perceptively seen (equilateral triangles). Then, LMN and MOP being equilateral triangles inside to equal triangles, the conclusion of the student may be that the lengths MN and MO are the same, as it is stated in the last sentence.

It is a global solving. The proof is incomplete, presents useless steps and never states inference rules.

Pisa diagnosis. Central Symmetry. Explanation: *A'= s(A) so A'BC equilateral*. It is coherent with all conceptions. - *MN=MO so MO= s(MN)*

Bristol diagnosis. Central Symmetry. Explanation: *MN=MO therefore MO is the symmetric of MN*. Even if in the justification there are elements corresponding to Reflection, the main point to determine MO as an answer corresponds to Central Symmetry.

6.2.2 DIDACTICAL DIAGNOSIS RESULTING FROM THE THREE ANALYSIS

All the analyses were compared. The comparison suggests a classification into three different categories, according to the different extent of agreement among the teams.

Total agreement

This category includes the cases in which there is total agreement among all teams, i.e. the conceptions identified are the same for all researchers.

Copy	Problem	Student's Answer	Conceptions: Grenoble	Pisa	Bristol
8, 24, 80, 103, 110, 112	1	MP	Reflection		
82, 100, 104, 105, 113, 119, 163	1	MO	Central Symmetry		
93	1	PC	Parallelism		
148, 150, 152	4	AA'B'B rectangle	Reflection		
124	2	RSM equilateral	Reflection		
163	2	RSM isosceles	Reflection		

Table 6. 19 total agreements in the human diagnosis

In case of problem 1 (14 of the 19 total agreements), some student's answers to the question about the symmetrical segment are strongly related to one conception of the symmetry in the following way:

- segment [MP], which is the right answer, is an indication of the conception Reflection; it can't be proposed by a student using another conception;
- segment [MO] is mainly related to the conception Central Symmetry;
- segment [PC] is mainly related to the conception Parallelism.

So, about this problem specifically, the relations between the type of answer and the conception gives a strong indication of the conception used. The analysis of the *copy-brute* can gives clues to confirm this first diagnosis. In just a few cases, there is indication in the copy-brute of the use of another conception. This cases are detailed in the next section.

The 5 other total agreements concerns two other problems. For this two problems, the answer of the student about the nature of a geometrical figure doesn't allow a direct conclusion about the conception. But in case of a correct answer, the correct conception Reflection can be assign and gathers the teams.

Partial agreement

This category includes the cases in which there is one (or more) conception(s) identified by *all* teams (in grey in Table 7). But some of the teams also identify other conceptions. This means that the conceptions (the common one and the others) are identified with more or less certainty. Such uncertainty may be due to the identification of two incompatible or compatible conceptions. For example, in the case of copy 160, for Pisa and Bristol, the first part of the solution is associated to conception Parallelism and the second part to Central Symmetry. But for Pisa, the two conceptions are believed to be incompatible.

			Conceptions		
Copy	**Problem**	**Student's Answer**	**Grenoble**	**Pisa**	**Bristol**
7	1	MO	Parallelism? Central Sym.? Oblique Sym.?	Central Sym.	Central Sym.
121	1	MO CP	Parallelism? Central Sym.? Oblique Sym.?	Central Sym.? Parallelism?	Central Sym. Parallelism
160	1	MO	Central Sym.	Central Sym.? Parallelism?	Parallelism Central Sym.
149	4	AA'B'B rectangle	Parallelism	Reflection? Parallelism?	Parallelism Reflection

Table 7. 4 partial agreements among diagnosis. A conception is common to the three diagnosis (in grey) but other conceptions also appear.

Uncertainty of the diagnosis

Cases of *No agreement* never happened, i.e. it never happened that a team identifies a conception rejected by another team. However, it is interesting to notice that there are five cases of agreement on the uncertainty about the attribution of conceptions.

			Conceptions		
Copy	**Pb**	**Student's Answer**	**Grenoble**	**Pisa**	**Bristol**
6	4	AA'B'B rectangle	Non Parallelism Oblique Sym.?	Reflection? Oblique Sym.?	Reflection
112	4	ABB'A' parallelogram		Reflection? Oblique Sym.?	Oblique Sym.?
113	4	AA'B'B rectangle	Non Parallelism	Reflection? Oblique Sym.?	Reflection?
123	4	AA'B'B square			Parallelism?
160	2	RSM isosceles	Reflection		Reflection

Table 8. 5 cases of agreement about the uncertainty of the diagnosis

Conclusion about the didactical analysis

The didactical analysis were carried out independently by three research teams from three different countries. Every team was left free to carry out the analysis according to their own principles. The researchers' background, cultural context and educational experience are different and certainly influence the process of analysis. However, a high level of concordance between the three diagnosis has been observed: among 28 copies, 19 agreements about the diagnosis and 5 agreements about the uncertainty of the diagnosis.

This concordance can be explain by the characteristics of problem 1, on which we had the maximum of agreement. That problem allows clear-cut diagnosis of conception and that facilitates the possibility of agreement. Moreover, the correct answer leads more easily to the diagnosis of the correct conception and then rally the advices of the teams (11 of the 28 copies).

In the case of the other problems, according to the a priori analysis, different degrees of confidence were observed. A more global interpretation of the solution of the problem was carried out, leading to the impossibility of identifying a single conception.

6.3. Baghera treatment

The second part of the evaluation process consisted of submitting the *copy-test* to Baghera and carrying out the *Baghera 1 analysis*, in order to be able to compare the results produced by Baghera with the results produced by the human analyses presented in the previous section.

The schematisation of the two processes is illustrated in Figure 27.

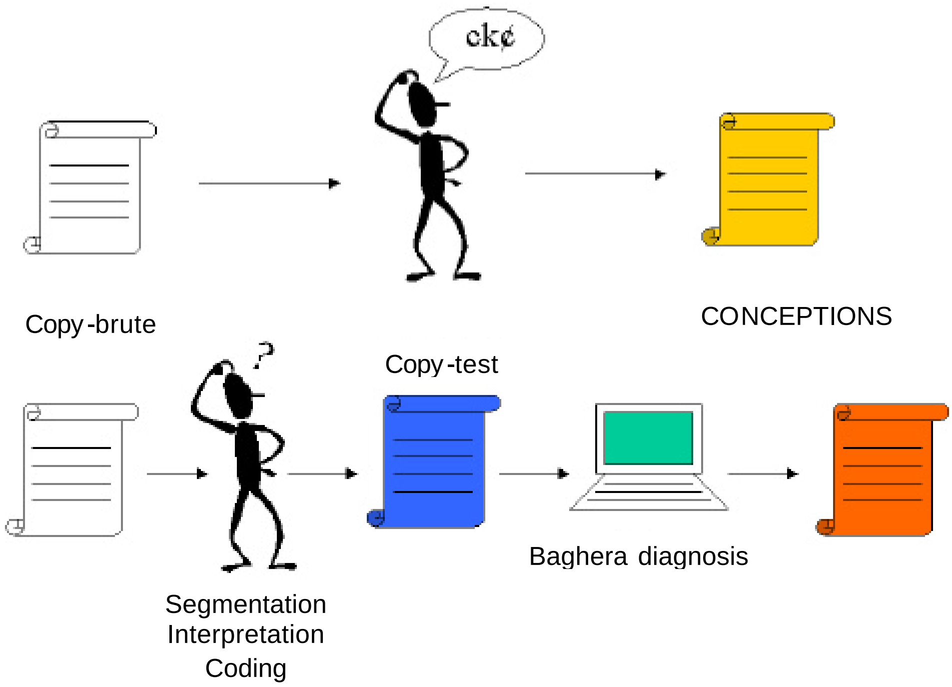


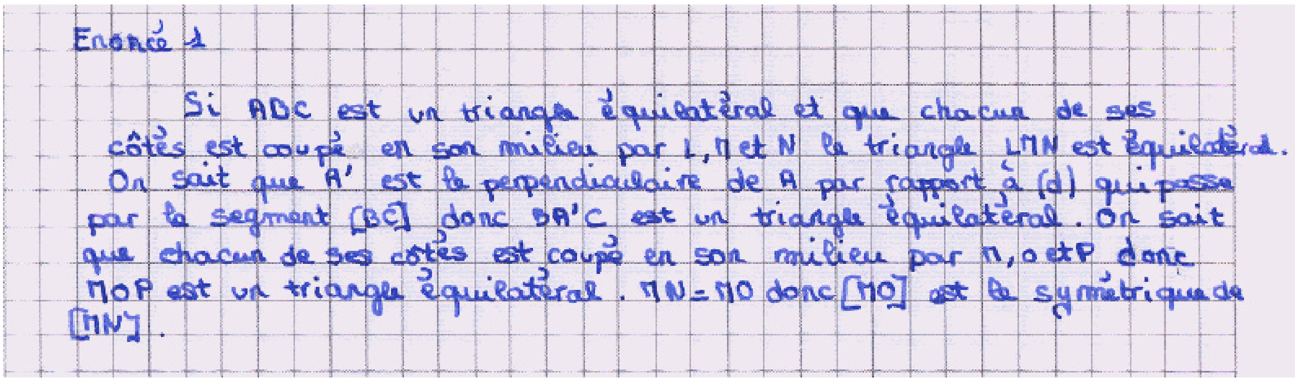
Figure 27. The didactical analysis and the Baghera 1 analysis.

While the *Didactical analysis* concerned directly the *copy-brute*, these could not be directly put into Baghera because the written protocols of the students didn't respect some Baghera constraints on the proof. In order to be analysable by the system, a proof must follow some constraints on expression that we detail below in the process of transforming a copy-brute in copy-test.

The choice of using the *copy-test for* the *Baghera 1 analysis* and the *copy-brute* for the *Didactical analysis*, is due to the fact that: first, in the future we think that Baghera will be able to work directly on natural language texts written by students in the interface and therefore both analyses will eventually be carried out starting from the same *copy-brute;* second, the process of transforming a copy-brute into a copy-test, which is very interesting from the point of view of the constraints of expression at the Baghera interface, is also the object of the evaluation.

A process of segmentation, interpretation and coding had to be done in order to transform the *copy-brute* in a *copy-test*, which could then be given to Baghera. Then, the *copy-test* were analysed by Baghera which produced the diagnosis of the conception(s) attributed to a given solution. This is the part of the *Baghera 1 analysis* which is the focus of our evaluation.

6.3.1 Transformation of a copy-brute into a copy-test



Enoncé 1

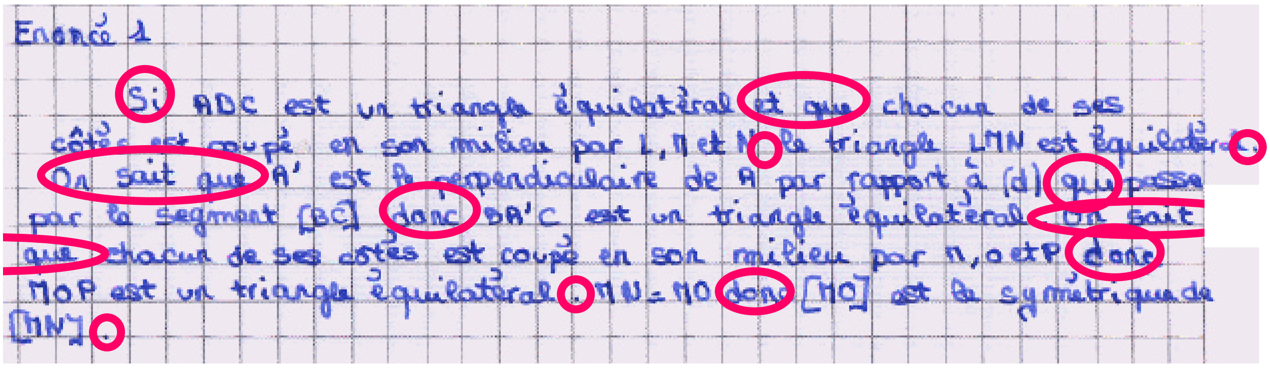
Si ADC est un triangle équilatéral et que chacun de ses côtés est coupé en son milieu par L, M et N le triangle LMN est équilatéral. On sait que A' est le perpendiculaire de A par rapport à (d) qui passe par le segment [BC] donc BA'C est un triangle équilatéral. On sait que chacun de ses côtés est coupé en son milieu par M, O et P donc MOP est un triangle équilatéral. MN = MO donc [MO] est le symétrique de [MN].

Three steps in this process: first segmenting the text to identify the statements, then cleaning the text (work on the implicit and non-standard statements) and organising the statements into steps of reasoning (identification of the status of each statement), finally coding with Baghera syntax.

Segmentation of the text

The identification of the different statements in the text of the student was made by picking out:

- the connecting words or expressions (if, then, I know that, and, by hypothesis, etc.)
- the punctuation marks (points, commas, etc.)
- the text editing (paragraphs, new lines, etc.).



Enoncé 1

Si ADC est un triangle équilatéral et que chacun de ses côtés est coupé en son milieu par L, M et N le triangle LMN est équilatéral. On sait que A' est le perpendiculaire de A par rapport à (d) qui passe par le segment [BC] donc BA'C est un triangle équilatéral. On sait que chacun de ses côtés est coupé en son milieu par M, O et P donc MOP est un triangle équilatéral. MN = MO donc [MO] est le symétrique de [MN].

Figure 28. Identification of statements in the student's proof.

This produces a list of clear-cut statements.

Interpretation of the statements

In the student's statements, some mathematical objects stay implicit. They are simply missing, it is the case of objects like the axis of reflection, or they are mentioned by a pronoun or a subordinating conjunction (cf. grammatical structure of the statement), they are present in the figure (coherence between problem and figure) or they are mentioned by a word like « figure » or « drawing » that have to be precise. In all this cases we explicit the mathematical object.

For example, the statement "Each of its sides is cut in its middle by M, O and P" becomes "M midpoint of [BC], P midpoint of [CA'], O midpoint of [BA']" (cf. Figure 29).

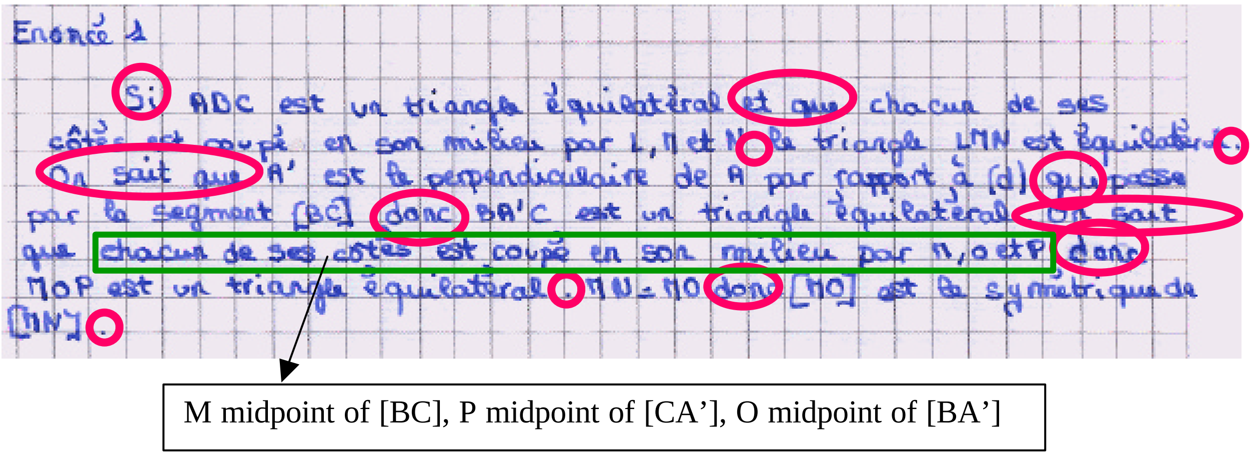


Figure 29. Interpretation of a statement.

Some statements are not mathematically standard and more problematic. We translate it when it was obvious or keep the original formulations.

For example, in the same student's protocol (Figure 29): "A' is the perpendicular of A with respect to d" can be translated into "A' is the symmetrical of A with respect to d" or not changed.

Then the list of "cleaned" statements has to be organised into a succession of proof steps. It means that we have to identify the status of each segment: hypothesis, conclusion and inference rule. Again, we used the connecting words, but in association with a logical analysis. For example, "if" can introduce an hypothesis but also the first part of an inference rule. The logical analysis allows to make the difference.

Then we rebuilt the deduction from the hypothesis to the conclusion. We organise the statements into clear-cut steps of deduction. In the student's protocol, steps can overlap and conclusion of a step is not necessarily mentioned in the hypothesis of the following step. So we move some statements and we repeat some others.

In order to be able to identify the reasoning and the steps of deduction, it is useful to add the missing inference rules. But, it is not required by Baghera.

Hypothesis	Inference rules	Conclusion
1. ABC equilateral *2.* L midpoint [AB] M midpoint [BC] N midpoint [AC]	*The triangle formed by the midpoints of the sides of an equilateral triangle is equilateral*	*3.* LMN equilateral
4. A'=sym_d(A) *5.* [BC] belong to d	*Invariance of d* *Conservation of the length*	*6.* BA'C is equilateral
6. BA'C equilateral *7.* O midpoint [BA'] M midpoint [BC] P midpoint [A'C]	*cf. step 1*	*8.* MOP equilateral
9. MN=MO		*10.* [MO]=sym_d([MN])

Table 9. List of statement after the organisation into steps of deduction. The text on a grey background are statements added by the human analysis.

Coding with the Baghera interface language

The input of the text at the Baghera's interface is the easiest part of this work. The interface of Baghera gives tools to express the sentences. This tools concern statements about geometrical objects (point, segment, line, midpoint…), geometrical properties (parallelism, perpendicular…) and a list of inference rules. The rest of the text is freely added in the editor. The inference rules not present in the written protocol are not written in the copy-test. Only the connecting words recognised by Baghera are used (according to a list of word and expressions). Finally, the conclusion of the proof is clearly marked (Figure 30).

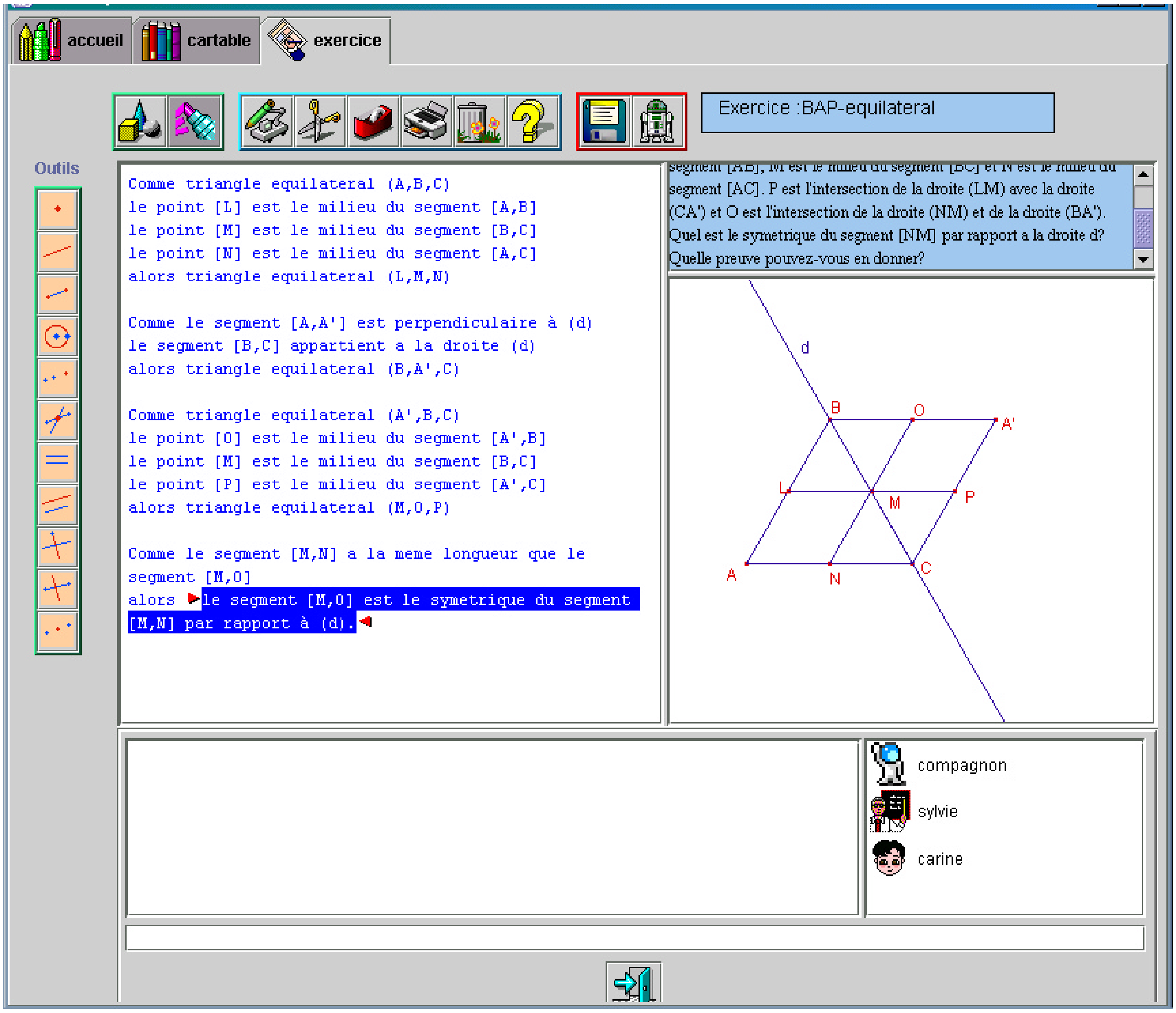


Figure 30. Copy-test in Baghera interface.

6.3.2 Baghera diagnosis of a copy-test

See section 4 where the detail of the Baghera process is presented.

6.4. Comparison between didactical diagnosis and Baghera diagnosis

In this section, we report on the comparison between the results of the *Baghera 1 analysis* and the results of the didactical analysis carried out by the three research teams on 28 protocols. This comparison will refer to the type of agreements and differences concerning the diagnosed conceptions in the human ana lysis, as classified in the previous section. We will be highlighted and commented first on the whole set of copies and then will look more carefully to two cases.

6.4.1 COMPARISON OF THE WHOLE SET OF COPIES

17 total agreements among the didactical diagnosis and the Baghera 1 diagnosis

In this 17 cases, the Baghera diagnosis agree with the didactical analyses (*total agreement*).

Copy	Pb	Student's answer	Didactical diagnosis	Baghera diagnosis
8	1	MP	Reflection	Reflection
24	1	MP	Reflection	Reflection
80	1	MO	Reflection	Reflection
82	1	MO	Central Symmetry	Central Symmetry
93	1	PC	Paralelism	Paralelism
100	1	MO	Central Symmetry	Central Symmetry
103	1	MP	Reflection	Reflection
105	1	MO	Central Symmetry	Central Symmetry
110	1	MP	Reflection	Reflection
112	1	MP	Reflection	Reflection
113	1	MO	Central Symmetry	Central Symmetry
163	1	MO	Central Symmetry	Central Symmetry
148	4	AA'B'B rectangle	Reflection	Reflection
150	4	AA'B'B rectangle	Reflection	Reflection
152	4	AA'B'B rectangle	Reflection	Reflection
124	2	RSM equilateral	Reflection	Reflection
163	2	RSM isosceles	Reflection	Reflection

Table 10. 17 cases of total agreement between the Baghera diagnosis and the didactical diagnosis

This 17 cases concern 19 of the cases where the human analysis leads to a total agreement about the diagnosis.

4 cases of partial agreement between Baghera and didactical diagnosis

In all the 4 cases of partial agreement amongst the didactical diagnosis, we can observe that the Baghera diagnosis is consistent with the didactical one. In fact, in this case, a partial agreement means that at least one conception was shared by the diagnosis of the human team and by Baghera. In the case of a second conception diagnosed by Baghera, the conception has also been proposed by two of the three human teams. This shows again a great concordance between the Baghera diagnosis and the human one.

			Conceptions			
Copy	**Pb**	**Answer**	**Grenoble**	**Pisa**	**Bristol**	**Baghera**
7	1	MO	Parallelism? Central Sym.? Oblique Sym.?	Central Sym.	Central Sym.	Central Sym.
121	1	MO and CP	Parallelism? Central Sym.? Oblique Sym.?	Central Sym.? Parallelism?	Central Sym. Parallelism	Central Sym.
160	1	MO	Central Sym.	Central Sym.? Parallelism?	Parallelism Central Sym.	Parallelism Central Sym.
149	4	AA'B'B rectangle	Parallelism	Reflection? Parallelism?	Parallelism Reflection	Parallelism Reflection

Table 11. Partial agreement. A conception is common to the didactical and the Baghera diagnosis (grey background), a second conception is common to Baghera and two of the three human diagnosis.

5 cases of difficulties in processing a comparison

In this 5 cases, a direct conclusion about the coherence of the diagnosis is difficult to state. On the human side, there is no common or partial didactical diagnosis because of difficulties coming from the specificities of the problem solved. For Baghera, it was possible to give a diagnosis but it is difficult to evaluate it against the weak results of the didactical diagnosis. However, apart from copy 6, the conception diagnosed by Baghera is evoked by a human team.

			Conceptions			
Copy	**Pb**	**Student's answer**	**Grenoble**	**Pisa**	**Bristol**	**Baghera**
6	4	AA'B'B rectangle	Non Parallelism Oblique Sym.?	Reflection? Oblique Sym.?	Reflection	Reflection Parallelism
112	4	ABB'A' parallelogram		Reflection? Oblique Sym.?	Oblique Sym.?	Reflection
113	4	AA'B'B rectangle	Non Parallelism	Reflection? Oblique Sym.?	Reflection?	Reflection
123	4	AA'B'B square			Parallelism?	Parallelism
160	2	RSM isosceles	Reflection		Reflection	Reflection

Table 12. Uncertainty amongst the diagnosis

Two cases of strong discrepancy between Baghera and didactical diagnosis

Copy	Pb	Student's answer	Didactical diagnosis	Baghera diagnosis
104	1	MO	Central Symmetry	Reflection
119	1	MO	Central Symmetry	Reflection

Table 13. Two cases of discrepancy in the diagnosis

We will comment on this two cases in section 6.4.2 below.

Conclusion : high level of consistency between Baghera and didactical diagnosis

There is a high level of consistency between the *Didactical analysis* and the *Baghera 1 analysis*.

In 17 cases there is a total agreement, i.e. total agreement among the didactical analysis and Baghera. In 4 cases, there is partial agreement leading to a common conception among the three teams and this conception is also diagnosed by Baghera. Then for 21 out of 28 cases, the behaviour of Baghera produce results coherent with the human ones.

In 5 out of the 28 cases, we have difficulties to conclude.

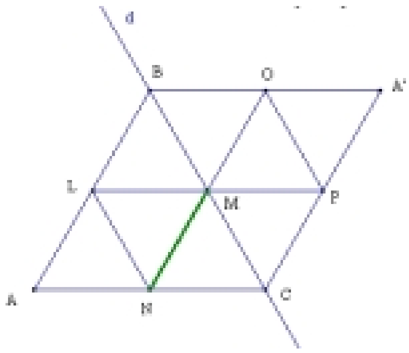
In only two cases, strong discrepancy occurred, and required detailed analysis, both in terms of evaluation and in terms of the methodology of evaluation.

6.4.1. Analysis of the two cases of strong discrepancies

In this section, we deal with the two cases of strong discrepancy, copy-brute n°104 and n°119. Let us go through the different analyses and highlight both similarities and differences in order to explain the inconsistency of the diagnoses.

Both protocols analysed in this section concern problem 1:

Let ABC be an equilateral triangle. Point A' is the symmetric of A with respect to the line d. L is the midpoint of segment [AB], M is the midpoint of segment [BC], N is the midpoint of segment [AC]. P is the intersection point of the line (LM) and the line (CA') and O is the intersection point of the line (MN) and the line (BA').
What is the symmetric of the segment [MN] with respect to the line d? How can you prove it?



Let us observe two important facts:

- the text of the problem includes a figure;
- the question contains explicit reference to symmetry: the problem asks the symmetrical of a given segment with respect to a given line.

We discuss the two cases together because of the strong similarities between them.

Didactical analysis.

In the case of the didactical analysis, a similar methodology is used for both copy 119 and copy 104.

The diagnosed conception is Central Symmetry. All the analyses report that the arguments mobilised in the student's proof do not allow to discriminate a conception. However, a diagnosis is achieved on the basis of the direct answer given by the student (that in both cases is: "the symmetrical segment of [NM] is [MO]") and the consideration of the appearance of the given figure.

Apart from the fact that, generally speaking, the Baghera diagnosis does not take into account neither the student's answer nor the appearance of the given figure, the two copies were treated in different ways.

Baghera 1 analysis of copy 104

From the Baghera analysis of copy 104 the following operators and controls emerge:

op10 = (If [A’B’] = Sym_d[AB], I midpoint of [AB] and I’ midpoint of [A’B’] then I’= Sym_dI.)
op49 = (If P is a point of the axis of symmetry d then P= Sym_d (P))
op65 = (If A’= Sym_d(A) and B’= Sym_d(B) then [A’B’]= Sym_d([AB]))
ct116 = (Symmetry preserves points alignment, with operators (op4, op10, op11))
ct120 = (Reflection preserves distances, with operators (op5, op10, op11, op46, op47))
ct124 = (The symmetrical of a point is a point, with operators (op1, op49, op64, op65))
ct125 = (The symmetrical of a segment is a segment, with operators (op64, op65))

Most operators and controls refer to many conceptions, while op10 = (If [A’B’] = Sym_d[AB], I midpoint of [AB] and I’ midpoint of [A’B’] then I’= Sym_dI.) and the related controls ct116 = (Symmetry preserves points alignment) and ct120 = (Reflection preserves distances) are ascribed only to conception Reflection.

Baghera 1 analysis of copy 119

From the Baghera analysis of copy 119 the following operators and controls emerge:

op5 = (If [A’B’] = Sym_d([AB]) then A’B’=AB)
op69 = (If A'= Sym_d(A) then [AA'] and d are perpendicular)
ct113 = (The line between a point and its symmetrical is perpendicular to the symmetrical axis, with operators (op6, op8, op46, op47, op69, op83)
ct115 = (The segment and its symmetrical segment have the same length, with operators (op5, op61, op70, op87, op97))
ct120 = (Reflection preserves distances, with operators (op5, op10, op11, op46, op47))

Op5 = (If [A’B’] = Sym_d([AB]) then A’B’=AB) and ct115 = (The segment and its symmetrical segment have the same length) are ascribed to many different conceptions whereas op69 = (If A'=Sym_d(A) then [AA'] and d are perpendicular) and the related controls (ct113 and ct120) are ascribed by the system only to the conception Reflection. This leads the system to diagnose Reflection.

Explanation

As we highlighted above, one main difference between the human analysis and Baghera one in both copies is that the former is strongly related to the answer students gave and to the appearance of the given figure. The student’s answer – “*MO* is the symmetrical of *NM*” – has been considered as a clear indication of conception Central Symmetry by the involved researchers because of the appearance of the given figure (where it is evident that MN and MO are symmetrical with respect to M). On the contrary, Baghera analyses only the student’s proof, neglecting student’s answers and possible interpretations of the given figure. However, Baghera is able to identify the conception Reflection by looking only at some key arguments in the proof (operators and controls).

What needs to be investigated and explained is whether these arguments have been observed by the researchers too and, when this is the case, whether they have been considered relevant and/or coherent with the diagnosed conception Central Symmetry.

At this point we have to distinguish between the two copies:

- **Copy 104.** The Baghera diagnosis is based on the operator op10 = (if [A’B’]=Sym_d([AB]), I is the midpoint of [AB] and I’ is the midpoint of [A’B’] then I’=Sym_d(I)) which has been ascribed only to conception Reflection. But in fact, an analysis carried out by the researchers concluded that this same argument is also consistent with the conception Central Symmetry. The operator 10 could also be ascribe to conception Central Symmetry.

- **Copy 119.** The Baghera diagnosis is based on the operator op69 = (If A'=Sym_d(A) then [AA'] is perpendicular to d). This argument does not appear in the *Didactical analysis*. Re-analysing *a-posteriori* the student's written protocol, we can find the following statement: "one knows that A' is the perpendicular of A with respect to d".

  This statement has been interpreted in two different ways by humans:

  - Those, who prepared the *copy-test* from the *copy-brute* for the *Baghera 1 analysis*, interpreted this statement stressing the presence of the key word "perpendicular" and obtained the following translation: "[AA'] is perpendicular to d"; this new statement led to obtain op69.
  - Those, who carried out the *didactical analysis* of the *copy-brute*, interpreted the same statement suggesting a linguistic confusion between "perpendicular" and "symmetrical". So the statement has been interpreted as a simple recall of a hypothesis: "one knows that A' is the symmetrical of A w
    not related to an operator.

Conclusion

To summarise, some elements responsible for the differences between the *didactical analysis* and the *Baghera 1 analysis* of copy 104 and copy 119 are highlighted below.

According to our analysis, the strong discrepancies are mainly due to the following factors:

- The attribution of some operators and controls to conceptions made by Baghera system is different from that of the involved researchers (when it happens, usually Baghera's ascriptions are more restrictive). It questions some didactical choices, related to modelling conceptions in terms of operators, controls and problems in the specific case of symmetry but does not directly concern the implementation of emergence. It should lead to an improvement of the table relating operators to conceptions implemented in Baghera.

- The process of translation of the *copy-brute* into the *copy-test* may cause problems because of linguistic ambiguities, possible different interpretations of the content of the copies and the existence of specific expression-related constraints of Baghera. The possible ambiguities in the interpretations of a written statements revealed in this example are likely to occur with a text produced within the Baghera environment. It shows some a priori limits on the Baghera diagnosis. Moreover, even if the process of translation concerns particularly this experimentation which dealt with written protocols, some of its aspects reveal a more general phenomena about the constraints coming from the use of the interface of a computer environment. When a student solves the problem directly within the Baghera environment, he has to find the way to express himself and he has to formulate his arguments within the possibility of the Baghera interface. This is an unavoidable phenomena which must be taken into account.

- Differently from the *Baghera 1 analysis*, the *didactical analysis* were not limited to students' written proofs. They took into account more elements, e.g. the students' answers, the appearance of the given figure and the relations between the figure and the problem. This aspect will have strong impact in the analysis of particular types of problems (e.g. problem 1). Baghera should in the future overcome this by integrating other registers into its diagnosis process. But the didactical analysis were also able to take into account some pragmatic and contractual aspects of the student's work. In the didactical and theoretical analysis of the problem, the possible answers are identified taking into account some constraints of the didactical situation (cf. "didactical contract" (Brousseau 1997)). For example, the fact that the student produces his proof for a teacher and not to convince a peer has effect on the proof he proposes. The human analysis is aware of these aspects of the situation and not the machine.

6.5. Conclusions about evaluation

The key-issue at stake in BAP are:

- modelling conceptions;
- implementing emergence as the key process to get relevant didactical behaviours of the system.

The evaluation carried out and described in the previous section tried to give answers to these issues.

According to the methodology adopted, we focused on the diagnosis phase, comparing the functioning of the system with human expertise, in particular didacticians' diagnosis expertise. The obtained results seem to confirm the Baghera effectiveness in diagnosing conceptions. In fact a large agreement was found between the didactical analysis and the Baghera analysis while the few cases of discrepancy may be referred to contingent factors. In particular:

- the "translation" from the *copy-brute* to the *copy-test* to be inserted as input in the machine.

The process of translation is related to the fact that we had to deal with written protocols, not generated within the Baghera environment. In the future, the comparison between human diagnosis and Baghera one will be carried out on the same object, a *"Baghera-copy"*. As a consequence some interference of the translation process should disappear. However, this process reveals some important aspects of the system interface and its constraints which have to be studied.

- the rigidity in the attribution of operators and controls to conceptions.

This latter aspect is related to the complexity and the difficulty of the implementation process, as witnessed by the limits that we found in the modelling of the conceptions about symmetry. In principle, such limits can be overcome but this aspect must be taken into account in the general evaluation of the effectiveness of the system. In any case it must be stressed that the functioning of emergence is not affected by this aspect.

An interesting aspect emerging from the evaluation process and which needs further investigations concerns the cases of uncertainty, i.e. when the three teams were unable to give a definite answer in the diagnosis. On the contrary, in these cases Baghera provided a definite answer ascribing one or two conceptions related to the operators and controls identified. This apparent more restricted diagnosis of the system should be carefully evaluated when the diagnosis of conceptions will be carried out in relation to the same student solving a sequence of problems. In this case emergence should guarantee a flexibility comparable with that of humans. An other way to enlarge Baghera diagnosis could be to take into account all the coalitions (groups of agents which vote for the same conception) and not only the strongest one as done in the current version of the mock-up.

However the problem of uncertainty of diagnosis should be considered both in the implementation phase and successively in the evaluation phase. At this moment the constraints of the mock-up did not allow such an analysis.

Other interesting aspects to be considered in the future come from the analysis of the potentialities of the logical diagnosis from HOARD-ATINF (see section 4.3). The evaluation carried out could not take into account the integration of HOARD-ATINF analysis into the existent Baghera platform. Thinking about such integration, it is possible to highlight the following aspects:

- the possibility that HOARD-ATINF identify implicit statements in the student's proof with respect to the formal system implemented;
- the possibility to identify new conceptions (see section 4.3.2.4).

According to the previously described methodology, it will be necessary to compare the didactical analysis and that carried out by Baghera integrating the contributions of HOARD-ATINF. We expect that the greater possibilities of the integrated system will allow a better consistency between the human diagnosis and that produced by the machine. However the main contribution of HOARD-ATINF is expected in relation to didactical interventions. In this sense, the evaluation of the system should be carefully designed. In particular among the potentialities of HOARD-ATINF the possibility of generating counterexamples seems very interesting for the contribution that it can give in the construction of didactical situations.

References

Brousseau, G. (1997) *Theory of didactical situations in mathematics,* Dordrecht : Kluwer Academic Publishers.

Grenier, D. (1988) *Construction et étude du fonctionnement d'un processus d'enseignement sur la symétrie* Thèse. IMAG, Université Joseph Fourier.

Tahri, S. (1993) Modélisation de l'interaction didactique : un tuteur hybride sur Cabri-géomètre pour l'analyse des décisions didactiques. Thèse. IMAG, Université Joseph Fourier.

7 Open BAP workshop for external validation

Leader of the task: Ricardo Caferra

As stated in task 3 of the project proposal, a workshop for external validation of our work by computer science and education specialists from the academic world and from industry was organised in Grenoble on November 2002 the 28th and 29th.

The list of the invited specialists is the following:

Peter Brusilovsky	School of Information Sciences and Intelligent Systems Program, University of Pittsburgh, USA.
Antonio Rizzo	University of Sienna, Italy.
Giorgio Olimpo	Institute for Educational Technology, Italian National Research Council, Italy.
Cyrille Desmoulins	Institute of Computer Science and Applied Mathematics of Grenoble, University Joseph Fourier, France.
Vanda Luengo	Institute of Computer Science and Applied Mathematics of Grenoble, University Joseph Fourier, France.
Lucien Lumbroso	General manager of AffixCCE Society, France.
François Bonnin	Engineer, AffixCCE Society, France.

In fact, Peter Brusilovsky was not able to attend the November meeting and came to Grenoble latter in February 2003.

All the participants of the project were present and most of them gave presentations. The workshop was structured on the following detailed presentations and demos (available at: http://www-baghera.imag.fr/seminBAP.htm):

- Presentation of the project by Nicolas Balacheff
- Key-ideas and presentation of Baghera platform by Sylvie Pesty and Nicolas Balacheff
- Demonstration of the mock-up by Sylvie Pesty, Carine Webber and Sophie Soury-Lavergne
- Theoretical framework: the cK¢ theory of didactical situations by Nicolas Balacheff and Hamid Chaachoua
- Implementation of emergence in Baghera by Carine Webber and Sylvie Pesty
- Implementation of ATINF in Baghera, by Nicolas Peltier and Jana Trgalova
- Emergence, theoretical framework by Jean-Pierre Muller
- Evaluation methodology by Maria Alessandra Mariotti
- Evaluation process and analyse by Sophie Soury-Lavergne and Mirko Marracci
- Questions raised by the evaluation process by Rosamund Sutherland
- The future of BAP with algebra by Jean-François Nicaud
- Synthesis and feed-back from the external evaluators

After his visit, Peter Brusilovsky sent us a report containing his evaluation (see below). He included valuable recommendations concerning the future of the project.

Baghera Assessment Project report

by
Peter Brusilovsky
School of Information Sciences and Intelligent Systems Program
University of Pittsburgh

I have read the draft report of BAP project and visited the Leibniz Lab in Grenoble. During this visit I have multiple talks with several members of the BAP team and observed several components of the target system. This is my brief report.

Since my major expertise areas do not completely cover the whole set of ideas that were to be explored by BAP, I want to focus on four aspects of the project that I consider most essential.

1. The use of agent-based technology as a basis for developing Web-based educational systems.

I consider the overall agent-based approach (top layer of agent-based architecture) that Baghera explores as an important direction to follow that needs additional research. Historically, modern Web-based courseware management systems such as Blackboard of WebCT are monolithic Web application. However, the monolithic nature of these systems brought these systems to a number of problems. One problems is re-usability – the monolithic systems are not able to use more successful components (such as better discussion forums or quiz systems). It drives away users who want to use their favorite components instead of the inferior components provided as a part of the system. Another problem is an inability to run external content components (known as "lessons" in AICC terminology). There are a number of attempts to solve this problem. Several university consortia are currently developing more open, component-based systems. Research groups explore various component-based architectures. Blackboard itself attempts to develop an approach to communicate with external components such as Question Mark's perception system. In this context, the work on high-level agent-based architecture of Baghera is important. It is commonly accepted that agents provide a most natural programming platform for developing distributed application (especially, distributed intelligent educational application). French researchers were first to start exploring seriously the agent-based architectures for learning environments. Baghera architecture is one of the two major projects that explore agent-based architectures in Web context, I also dare say, the more technically solid of the two. Continuing and enhancing this stream of work is important for the world WBE research community in general and also important to retain European and French leadership in this field.

2. The idea of conceptions and its role for developing more intelligent learning environments

While I am not an expert in the educational psychology and instructional design, I can clearly see the value of the cK¢ theoretical framework for driving the design of more intelligent learning environments, As pointed out by the proposal authors, modern approach to developing intelligent tutoring systems (ITS) is based on the idea of an overlay student modelling. An overlay model describes student knowledge as a subset of expert knowledge in the same field. This model has been originally developed for the case of diagnosing student knowledge. Unfortunately, it has since become dominating in developing pedagogical approaches behind ITS. More exactly, the dominating approach to knowledge expansion and problem sequencing in ITS assumes that students have to add components of teacher's expertise one by one to their knowledge. As pointed out by the authors, a real education is not that simple. The student knowledge, while being incomplete are rarely just a part of a whole – they are often self-contained compendium that allows a student to work on a subset of known problem. This is very close to the history of establishing of humankind knowledge. Understanding of this aspect is important to provide a proper instructional sequencing in ITS. Early "titans" of ITS field has got to understand this problem. It was on was discussed in an important and immensely cited work of Goldstein on *genetic user models*. Unfortunately, at that time there was a clear lack of computational support for this paradigm outside of primitive cases. While often cited, the spirit of this work has been forgotten and the field moved to simpler computationally supported overlay models.

I was very happy to see that this stream of work has been revived now on a more advanced level of understanding, backed by authors solid expertise in educational psychology and instructional design. I think that principles formulated by Balacheff and Gaudin in section 2 of the BAP report are very important for the development of future high-quality ITS and other advanced instructional systems independently of the success of their implementation side that I will address in the next section. Having a better understanding of the knowledge establishing process will help us to develop better sequencing techniques and avoid primitivistic non-scaleable approaches. Most important these ideas are for such fields as physics, math, and other sciences subjects – the subjects that are most critical for a scientific well-being of any leading nation, but are also the hardiest to teach. Teaching of these subjects is currently undergoing a serious crisis (at least at the United States). National Science Foundation allocates a solid portion of its budged to develop better educational systems in these fields and to explore the process of learning. In this context the stream of work on conceptions and the cK¢ model is really important. The only thing that I wish this team to achieve is to publicize their work more broadly and in a form that is more open to a larger community of researchers (i.e, with more supported examples) so that it can influence larger number of researchers. At the moment, the presentation, as it is in Section 2 is very dense with ideas, formulas, and assumes some reasonable understanding of educational psychology.

3. Agent-based approach to diagnosing concepts

As I have mentioned above, the pioneer ideas of Goldstein have not been explored further due to the lack of proper computational platform. The Baghera group, however, was able to continue practically the work in the same direction that has originally motivated genetic student models. They suggested an interesting and creative way of diagnosing multiple conceptions using a specific agent-based approach that they were exploring in the context of this project. I have found their approach not only very original, but also very meaningful. It is cognitively grounded in the cK¢ theory, but it is also computationally possible and scaleable since the number of required agents grows linearly with the growth of the field complexity. I think, an idea of having one agent for every element of the cK¢ framework is a really clever one.

At this moment, there is no guarantee that this approach will be a complete success – as it is with any really serious advance piece of research. However, from my point of view, the work that has been done over the duration of the BAP project provides very good evidence that the project will succeed. In this sense, this one year of work was a year quite well spent. I had a chance to see the work of the diagnostic component of the project and have asked the team run a number of examples of diagnosis for me. Altogether, the Grenoble team and the teams in the UK and Italy have done a marvelous work to analyze a solid number of student problem-solving examples both by humans and by the multi-agent technology. It has provided me with a great material to evaluate the performance of the multi-agent framework. I should tell clearly that I was quite impressed with the performance of the framework – even more than after reading the project draft. From the project draft itself, it is clear that the agents were able to replicate human diagnosis in a large number of cases, though not in all. What I have seen during my work with the framework is more exciting. While I have not run and analyzed all cases, I still have seen a lot of them spending at least 2 hours on this part of my visit. After all that, I have got a clear impression that the performance of the team of agents is very close to the performance of the team of human evaluators. I.e., if the case is quite clear and several human teams expose a very high level of confidence and a coherent decision between the teams, the agents are also able to get to a definite diagnosis. If the case is incomplete or unclear and the human team has low confidence and disagreement, the agent framework is able to replicate this pattern by ending up with more then one strong teams of agents. Myself, I consider it as a strong evidence in favor of the framework. Any good diagnostic systems should be able to replicate not only a clear agreement in clear cases, but also the lack of agreement in unclear cases. I feel that this positive aspect of the system has not been well presented in the draft of the report. Overall, I think that the results of the agent-based diagnosis are impressive and provide a very good support in favor of the agent-based diagnosis in general and the groups original approach to developing agents following the cK¢ framework. I hope that the team will get support to continue work in this direction.

4. Machine learning for building the expertise in ITS

My last comment is about the work on the use of machine learning in the context of analysis that the group has started relatively late in the course of the project. While the results of this work themselves are too premature to comment on, I feel important to comment that the authors are doing all the right steps in the direction that I consider important and timely. One of the problems of modern ITS is the acquisition of expertise. Moder ITS can provide an impressive level of learning support, but this support is based on manually collected and extracted knowledge. While this is sufficient for small domains and lab systems, a future of ITS depends on their ability to accumulate and build expertise automatically or, at lest, semi-automatically. This need has caused a few leading ITS research groups that are already working on large practical ITS (such as Bev Woolf group in U. of Massachusetts and Jack Mostow group in Carnegie Mellon) to start work on using machine learning for developing ITS expertise. This direction of work is doing its first steps in the research world, but I found it very important. I was pleased to know that the BAP team has embarked to this direction of work and does all the correct first steps. Their machine learning team is solid and professional and, given the overall attitude to research in Leibniz lab, I may expect some major research results coming from this project. I may only wish the group started the work on Machine learning earlier – but this group is hard to blame. I think, at the moment, less than a handful of ITS teams has realized the importance of machine learning and started exploring this research direction.

Overall, I was quite impressed with the work of the BAP project. I think, that the team made a solid step towards their ultimate goal and that their original idea that may look as an advanced dream just a year ago is now getting its practical shape. I think that the one year work has provided some good results and made the groups chances to achieve their ultimate goal much stronger. Moreover, I was pleased to see that on the way to their ultimate goal the group has produced some interesting results that were not originally planned and has started to investigate some other critical research direction. For me these are the features and the signs of a successful solid research project. I wish the authors success and hope that the team will get a chance to continue their pioneer work

Peter Brusilovsky
School of Information Sciences and Intelligent Systems Program
University of Pittsburgh

8 Conclusion and perspectives

Author: Nicolas Balacheff

The initial challenge of BAP was scientific and technological; in the end it has also been a human challenge in an epistemological sense. We have been, all along the project, driven by the need to build a common culture. This process has focused on the role of models, their nature, their constraints and benefit. Especially, the exigencies of modelling for computational purposes have been confronted to those for educational purposes of which sense making is an essential preoccupation. At all the stages we have succeeded in integrating the different views, even if the time needed for that imposed finally to reduce the initial scope of the assessment of Baghera to the issue of student modelling. Baghera, can now count on a multidisciplinary and coherent European team, fully integrating education and computer sciences.

As a result of the integration process, and also because the issue was raised during the external evaluation workshop, it is worth clarifying that:

- BAP does not ignore the social and cultural dimensions of teaching and learning. These will be taken into account in the implementation phase of Baghera in actual schools, especially under the competencies of the research groups from Pisa and Bristol.
- The potential interaction between the tools and the teachers is not an issue addressed at this stage of the project, but it will be considered as soon as the project has available a prototype robust enough to go to schools.

Concerning the results of this assessment project, each section of the report presents the outcomes of the work carried out, and especially the evaluation section suggest that going ahead is worth considering. In the near future our plan is to achieve the cK¢ model by developing fully its didactical dimension, to specify and implement the multi-agents model of the didactical functionalities of Baghera, to strengthen the mock-up so as to have a robust prototype to use with real students.

One severe limitation of the current work comes from the fact that we have not been able to get a full version of a micro-world, which would have allowed *to trace in details the students problem solving activities* in geometry modelling. This possibility is indeed a crucial point, because the diagnosis process is much more relevant and efficient if it does not rely only on a production but on the whole process which led to it. This issue was pointed by the participants in the external evaluation workshop; we agreed on this remark. Turning to algebra will be a solution, since it is a domain for which we have a full access to an environment, Aplusix, which is under development at the Leibniz laboratory (see Annex I).

To ensure the relevance and effectiveness of this evolution, which we understood rather early in the project life, in parallel to BAP, a team of the Leibniz laboratory together with other partners has undertaken a project exploring emergence (in the machine learning sense) and student modelling in the case of algebra (project funded by the French Ministry of Research, leader J.F. Nicaud). The very promising results of this project speak for the success of its merge with BAP. More, it appears that algebra is clearly in the domain of competencies of the Pisa Group (which is also developing an algebraic environment) and the Bristol Group (which widely known for its work in this field.) Algebra will open an easier way to mathematical modelling in various domains from scientific problems to everyday life, hence broadening the present approach.

A good indication of the potential of the project may be interesting to get at the end of these lines: the cK¢ model, which is the core of the Baghera approach, is currently used to support the design of the learning platform of the Flexible University project VOEU (IST FP5, project officer Wim Jansen) in the domain of orthopaedic surgery. This contribution to VOEU has been positively evaluated by the referees and let us expect an efficient dissemination of the Baghera outcomes once we have achieved the whole project as intended.

Annex I
Looking forward to Algebra

Author: Jean-François Nicaud

Two computer programs for helping students to learn algebra are developed in the teams that participate to the BAP project. These programs are **Aplusix** (did@TIC team, Leibniz laboratory) and **L'Algebrista** (Department of Mathematics, University of Pisa). They are currently experimented with students and researches are made to investigate how they help students to learn algebra. In this annex, we describe one of these two systems, the Aplusix system, and the associated researches. Then we envisage to plunge this system in the BAP project.

A.1. Presentation of Aplusix

Aplusix is both a microworld and a training system for algebra. In the microworld mode, the student may input an exercise that concerns an algebraic expression, i.e., an equation, and solve it. In the training mode, the student has to solve exercises chosen by the teacher and stored in a file.

A.1.1. Main characteristics

The main characteristics of Aplusix are:

- Visual fidelity (same language)

  The representation of algebraic expressions is identical to the usual paper representation, see Figure 31.

- Reification of the reasoning process

  The reasoning process of the student is represented with boxes containing expressions that are the steps of the student. A link indicates form what step a given step is derived, see Figure 31. This representation allows backtrack, see Figure 32.

- Smart edition

  Aplusix includes an advanced editor in two dimensions of algebraic expressions. This editor is very intuitive including the usual feature of editors (insert, delete, select, copy, cut, paste, drag & drop). It takes into account the structure of algebraic expressions, for example, only well-formed sub-expression can be selected. See an example in Figure 33.

- Strong and non intrusive feedback

  A feedback concerns the equivalence of the expressions. This may be interpreted as correctness indicator. When the expressions are not equivalent, the link between the two steps has the form of an equivalence crossed in red, see Figure 32.

The other indicators concern properties of the expression of the current step and indicates whether the expression is reduced or not, is expanded or not, is factored or not, etc. At the bottom of Figure 31, the Reduced indicator is filled, which means that the current expression (where the insertion point is) is reduced, and the Equation indicator is half filled, which means that there is a progression in the resolution of the system of equation but that it is not yet solved.

- Commands

Aplusix includes commands in a menu to make calculations. When the student applies a command to a selected expression, the computer makes the calculations of this command. This feature of Aplusix is close to CASs (Computer Algebra Systems).
See an example in Table 14.

- Saving the interactions

Aplusix record in a file the interaction: each keystroke, movement of the insertion point, selection, command. The file may be analysed off line.

- Depending of parameters that are set by the teacher

Aplusix allows the teacher to choose the values of many parameters in order to customise the system to his wishes. For example, the teacher may set Aplusix with/without the indication of the equivalence, with/without the indicators, with/without the commands for the calculations.

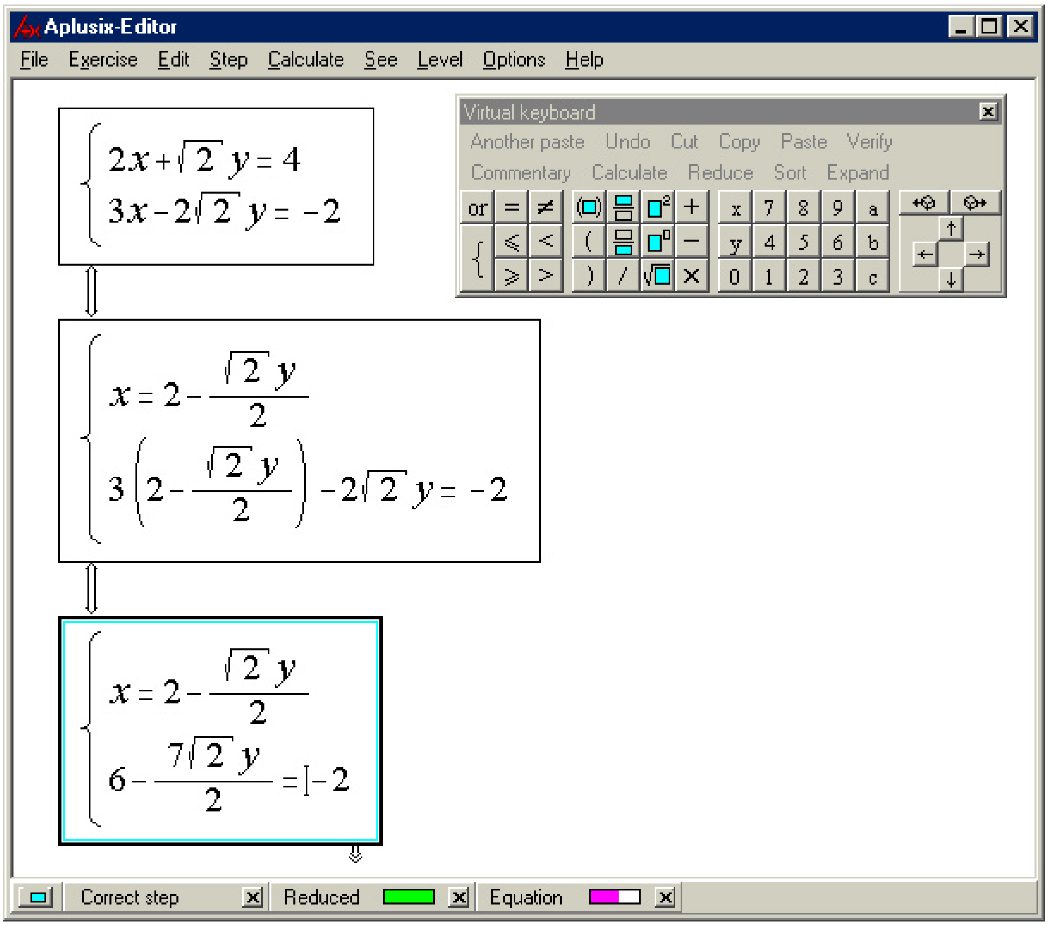


Figure 31. Aplusix Editor interface, a problem resolution

Figure 31: Mary has loaded an exercise, which is displayed in step one (the first box). She duplicated step one in step two and modified step two getting what appears on the screen (in the first equation, she selected $\sqrt{2}\,y$, dragged and dropped it at the right of 4; she made a fraction with it, typed 2 has denominator and typed – before the fraction; she deleted the 4 and typed 2 at this place; she deleted the 2 before x). Then, she selected the right hand side of the first equation, made a copy, selected x in the second equation and pasted. Such paste in the APLUSIX-EDITOR is a substitution, that is why parentheses appear. Last, Mary duplicated step two in step three, selected the left hand side of the second equation and clicked on the "expand" menu, then on the "reduce" menu. The double arrow between the steps indicates that the calculations are correct. The status bar provides information about the current step. Here, the status bar tells that the step is correct, that the expressions are reduced and that a significant step has been executed toward a solution.

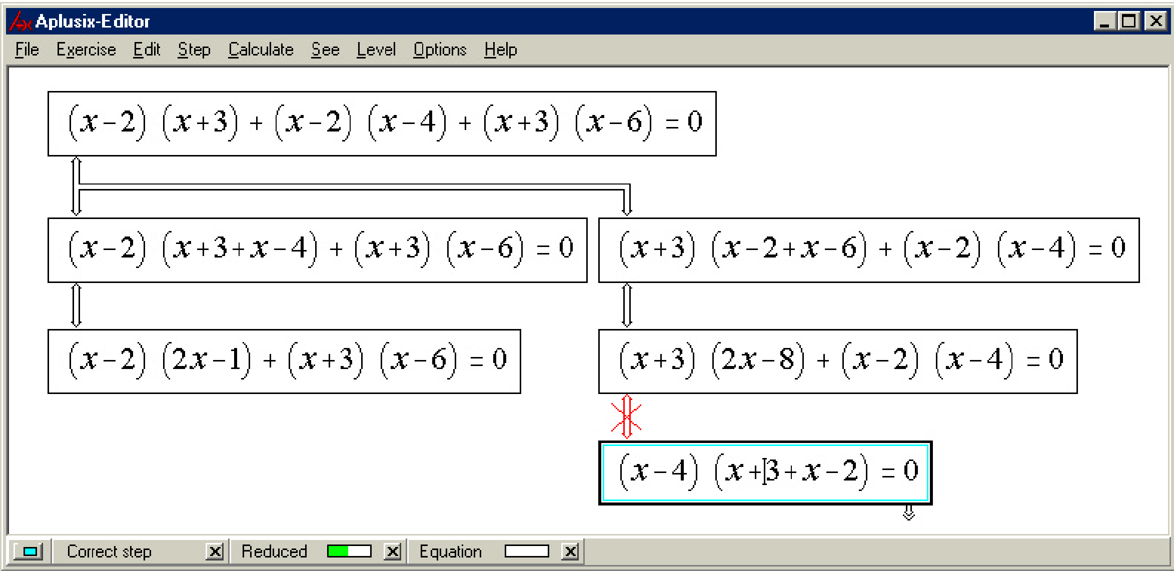


Figure 32. Peter's resolution in Aplusix

Figure 32: Peter is at a level where efficient methods for factoring ax2+bx+c forms are not known. At this stage, a good strategy for factoring polynomial expressions consists of applying factorisation rules. So Peter used first the partial common factor x-2 expecting a common factor after. As his expectation failed, he went back to use the other partial common factor x+3. His calculations are correct, except that last one. The error is indicated by the red cross over the equivalence sign.

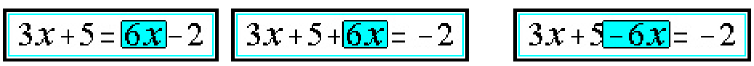


Figure 33. Moving additive sub-expressions

Figure 33: Here is the technique for moving additive sub-expressions in an equation. First, select the sub-expression, second drag & drop it on the other side, third hit the minus key to change the sign.

$4(x-2)^2-(x-2)(x+1)=0$	The student has selected a sub-expression.
$4x^2-16x+16-x^2+2x-x+2=0$	The student has clicked on the *expand* command. The result has replaced the selection and remained selected.
$3x^2-15x+18=0$	The student has clicked on the *reduce* command.

Table 14. Applying commands to an expression.

AVAILABILITY

Aplusix is currently available as free system limited in time. It will become a commercial product in 2003. It runs in French and English and will be translated in other languages.

FEEDBACK FROM USERS

The enquiries we made show that teachers and students like well the Aplusix system.

A.2. The cognitive-algebra project

GOALS AND FRAMEWORK

The goals of cognitive-algebra project are:

- To model students in algebra,
- To find prototypical conception,
- To elaborate teaching strategies.

The project is funded by the French Ministry of Research. The researchers involved in the project belong to several disciplines and several teams:

- Computer science
  - Knowledge base systems: did@TIC team, Leibniz, Grenoble
  - Machine learning: *Appentissage* team, Leibniz, Grenoble
- Didactique
  - Did@TIC team, Leibniz, Grenoble
  - LIRDEF, IUFM de Montpellier
- Psychology
  - *Cognition finalisée* team, University of Paris 8

THE FIRST EXPERIMENT AND THE FIRST ANALYSES

The project began in October 2002. We made an experiment with:

- 91 students 15 years old
- 108 students 16 years old
- 47 students 17 years old

Each student had 2 sessions of 1 hour, the first one without verification (indication of the equivalence) and the on second with verification on request (when the student clicks on the Verification menu). Some students had a third session with a permanent verification.

The interaction files are currently analysed to model students. We use several ways for that.

1) Replay. We have realised a computer program that allows to replay the student behaviour. This is a fine tool to study in detail the behaviour of a few students.

2) Making specific table to study by hand. We have realised a computer program that selects particular steps in the student's behaviour and presents them in a tableau with complementary information.

3) Applying hierarchical classification algorithm to the students' beha viour. We have realised a computer program that determines the rules of a reference knowledge base that are applicable to an algebraic expression. We have enriched the students' file with this information and applied a hierarchical classification algorithm. This provides clusters (see Figure 34) that groups students in a way we are currently analysing.

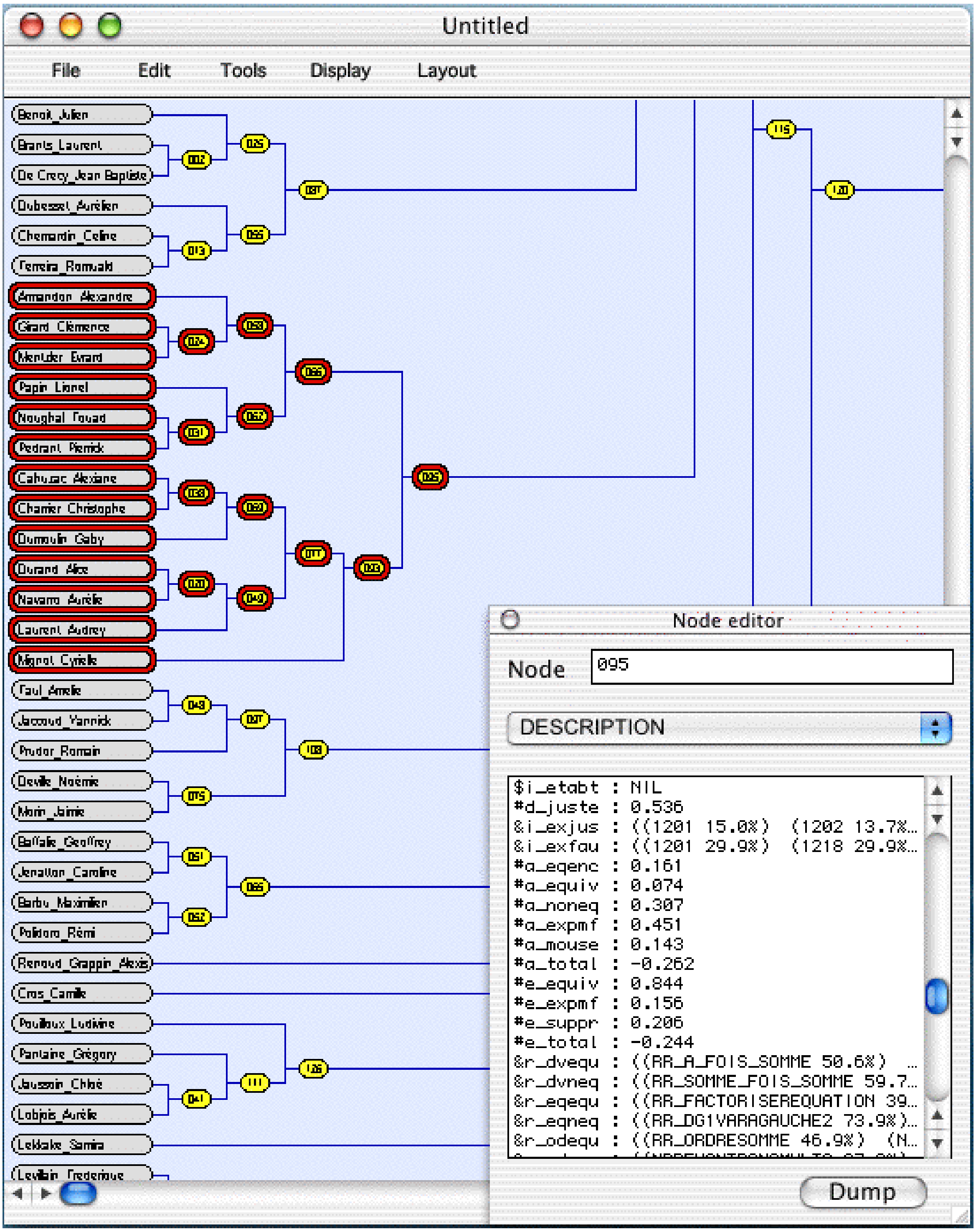
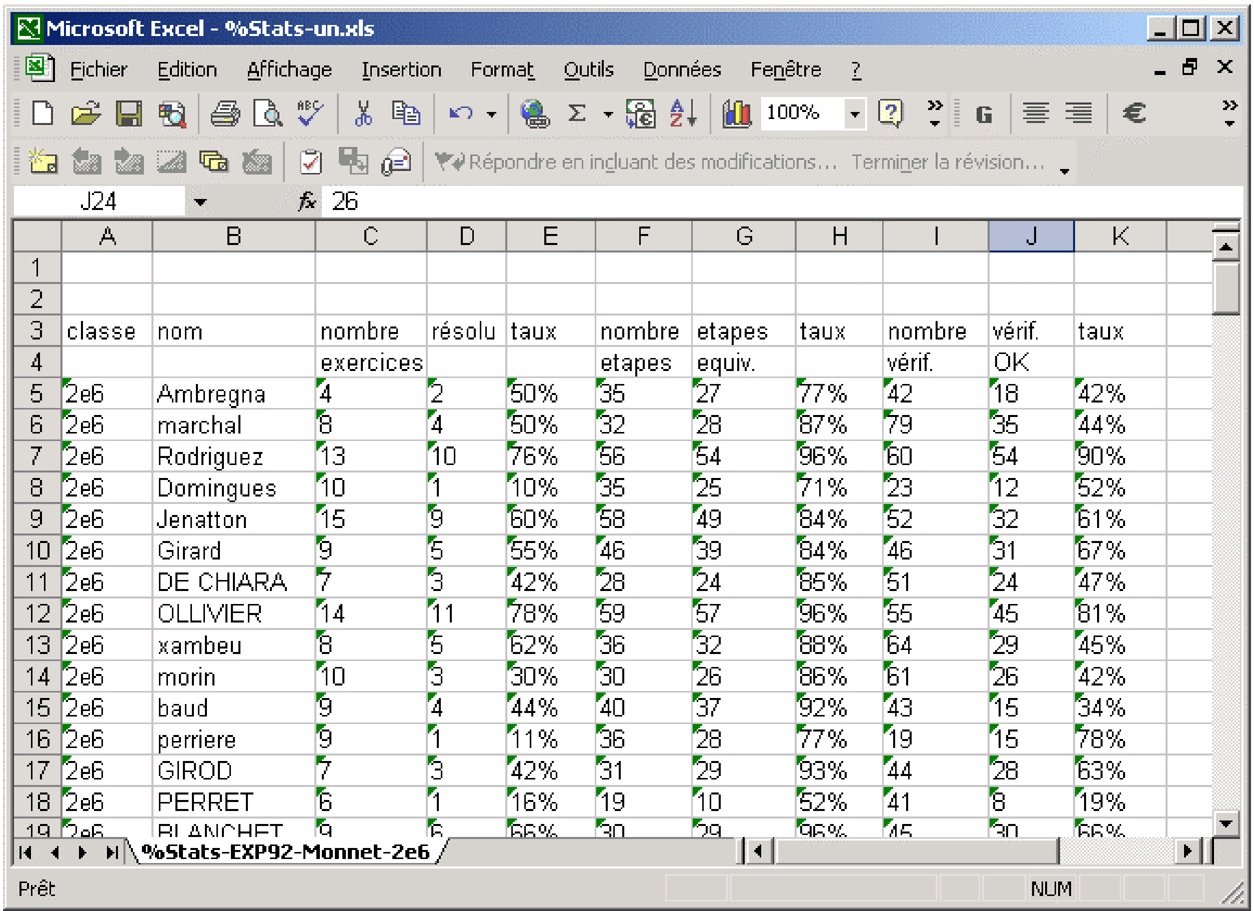


Figure 34. Making clusters with the HAC algorithm

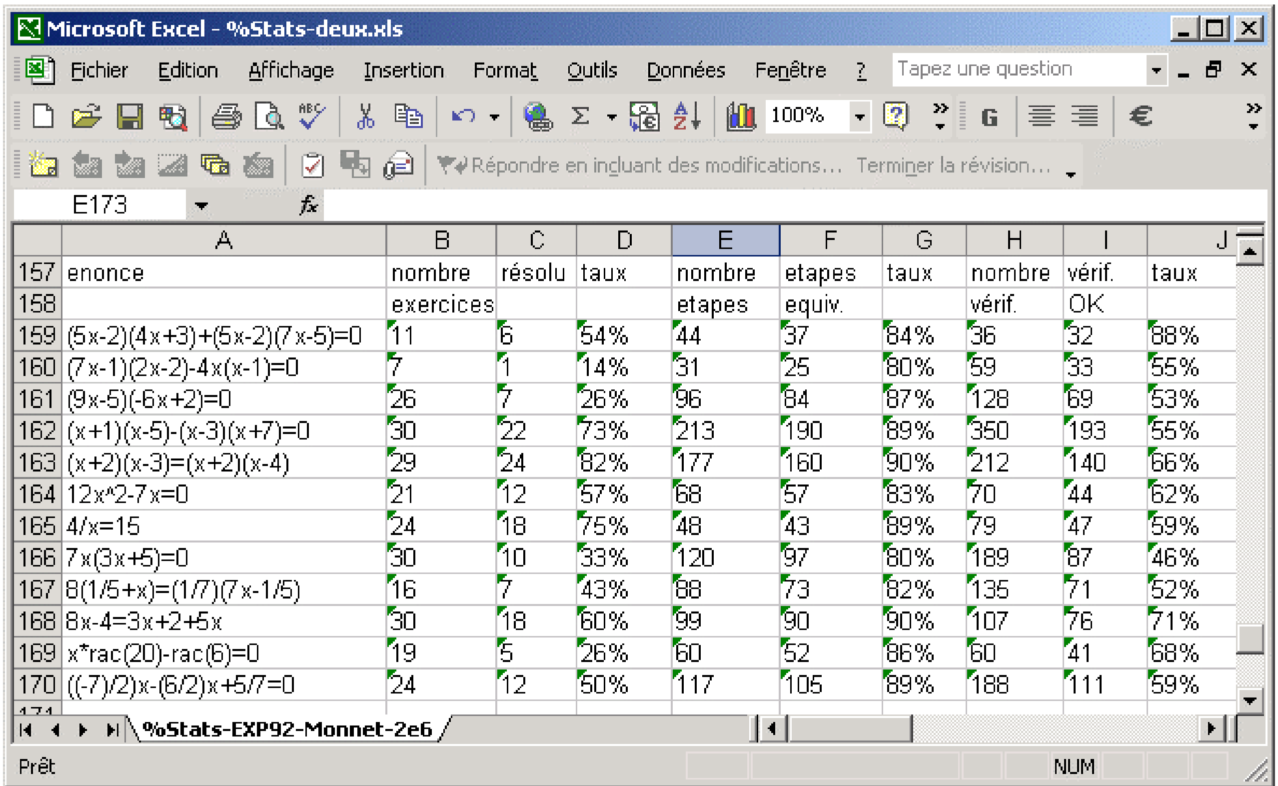
6) Making simple statistics by program. We have realised a computer program that calculates some statistics related to the students or the exercises. See
Figure **35** and Figure 36.



Microsoft Excel - %Stats-un.xls

classe	nom	nombre exercices	résolu	taux	nombre etapes	etapes equiv.	taux	nombre vérif.	vérif. OK	taux
2e6	Ambregna	4	2	50%	35	27	77%	42	18	42%
2e6	marchal	8	4	50%	32	28	87%	79	35	44%
2e6	Rodriguez	13	10	76%	56	54	96%	60	54	90%
2e6	Domingues	10	1	10%	35	25	71%	23	12	52%
2e6	Jenatton	15	9	60%	58	49	84%	52	32	61%
2e6	Girard	9	5	55%	46	39	84%	46	31	67%
2e6	DE CHIARA	7	3	42%	28	24	85%	51	24	47%
2e6	OLLIVIER	14	11	78%	59	57	96%	55	45	81%
2e6	xambeu	8	5	62%	36	32	88%	64	29	45%
2e6	morin	10	3	30%	30	26	86%	61	26	42%
2e6	baud	9	4	44%	40	37	92%	43	15	34%
2e6	perriere	9	1	11%	36	28	77%	19	15	78%
2e6	GIROD	7	3	42%	31	29	93%	44	28	63%
2e6	PERRET	6	1	16%	19	10	52%	41	8	19%

Figure 35. Simple statistics on students made by program.



Microsoft Excel - %Stats-deux.xls

enonce	nombre exercices	résolu	taux	nombre etapes	etapes equiv.	taux	nombre vérif.	vérif. OK	taux
(5x-2)(4x+3)+(5x-2)(7x-5)=0	11	6	54%	44	37	84%	36	32	88%
(7x-1)(2x-2)-4x(x-1)=0	7	1	14%	31	25	80%	59	33	55%
(9x-5)(-6x+2)=0	26	7	26%	96	84	87%	128	69	53%
(x+1)(x-5)-(x-3)(x+7)=0	30	22	73%	213	190	89%	350	193	55%
(x+2)(x-3)=(x+2)(x-4)	29	24	82%	177	160	90%	212	140	66%
12x^2-7x=0	21	12	57%	68	57	83%	70	44	62%
4/x=15	24	18	75%	48	43	89%	79	47	59%
7x(3x+5)=0	30	10	33%	120	97	80%	189	87	46%
8(1/5+x)=(1/7)(7x-1/5)	16	7	43%	88	73	82%	135	71	52%
8x-4=3x+2+5x	30	18	60%	99	90	90%	107	76	71%
x*rac(20)-rac(6)=0	19	5	26%	60	52	86%	60	41	68%
((-7)/2)x-(6/2)x+5/7=0	24	12	50%	117	105	89%	188	111	59%

Figure 36. Simple statistics on exercises made by program.

THE CURRENT WORK AND FUTURE WORKS ON STUDENT MODELLING

We are now searching the rule or succession of rules a student applies to go from step n to step n+1. These rules may be correct rules or mal-rules. A library of mal-rules is being built according to three ways: (1) modification of correct rules by researchers, (2) hand identification of mal-rules in some students' behaviour, (3) getting mal-rules published in articles. We have realised a computer program that search with a classical AI technique what succession of rules allow to go from step n to step n+1. When our library of mal-rules will be filled, we will be able to apply it to the entire behaviour of each student and to calculate the occurrence of each mal-rule. Of course, we do not expect a 100% success and we will probably have to refine the search algorithm to have a good result.

After these works, we will search prototypical conceptions. At the rule level, we call prototypical conceptions a set of rules (correct or not) that is consistent in some way and that we can attribute to some students. Automatic classification algorithms will be used for that. If we succeed to identify a few prototypical conceptions that cover a large set of students, we will have a strong result at a cognitive and didactical level because we will be able to describe these prototypical conceptions in a way teachers can understand which will help them in their work.

When prototypical conceptions will be determined, we will be able to compare groups of students by observing the frequency of the prototypical conceptions, in particular students of different countries.

When prototypical conceptions will be determined, we will build teaching strategy to help students to correct incorrect conception or to progress.

A.3. Plunging the Aplusix system and researches in the BAP project

The work concerning prototypical conceptions and teaching strategies of the above cognitive-algebra project is totally compatible with the BAP project. Plunging the Aplusix system and researches in the BAP project has two main interests. First it groups researchers working on the same theme making a stronger research force. Second, it goes to a domain that concerns a very important part of the learning of many students and a domain that may be connected to many others, in particular:

- word problems that are solved with algebra techniques,
- electricity, mechanics that provides equations to solve,
- calculus.

Annex II
The notion of emergence

Author: Jean-Pierre Muller

After introducing the motivation, we shall present the history and some most interesting definitions about emergence without being exhaustive. After discussing these definitions we shall propose a new one opening the possibility to derive a methodology for designing multi-agent systems (MAS), in particular for problem solving (part 2).

1. The notion of emergence

A traditional approach in Artificial Intelligence is to propose problem solving methods inspired by the strategies developed by the natural systems (in particular the humans conceived as thinkers). These methods rely mostly on an a priori formalization of the problem domain. In dynamic and uncertain domains, this a priori formalization becomes difficult and requires an increased adaptive capability. Nature seems to have solved this problem by emergence of collective behaviours selforganizing from the dynamics of entities in interaction between themselves and with an environment. Unfortunately, the notion of emergence itself is problematic for the modeller as well as the designer. For the time being, there is no constructive definition of emergence. This notion is more frequently defined by an absence of "something" as composability, predictability, and so on when the notion itself is not criticized.

Most often, the notion of emergence is associated to a misunderstanding or intuition facing some phenomena observed in nature; it is therefore rejected for a reductionist interpretation. For example, Daniel Memmi [Memmi 96] limits emergence to a problem of description or explanation and assumes that emergent phenomena are examples among others of the variety of scientific explanations. In this sense, emergence is not in nature but in the change of observation focus on the phenomena. However, even if emergence is used to explain some phenomena, these phenomena have to pre-exist to our observation. In other terms, the emergence of an entity, a structure, a function or a process is in the system independently of our observation even if it is a change of point of view of the observer which reveals the emergence. Ants colony with their bridges and the termites with their nests produce emergent phenomena for million years without any observers.

To better understand and exploit these emergent phenomena, one should provide not only an alternative approach to computer modelling, but also for problem solving. This approach, whose principle is to build a collective of agents, immersed in an environment, which by their interactions will evolve towards a stable state representing a solution, has been initiated by R. Brooks in robotics [Brooks 89a, Brooks 89b, Brooks 91]. It is opposed to the classical problem solving approach where the global resolution task is decomposed into subtasks. The program then codes the resolution steps; while executing, the process follows the predefined path until the solution is reached. In the "emergentist" approach, the program codes the agents, the environment and the interactions; while executing, the process self-organizes and builds a solution.

Another trait of the emergentist approach is the adaptative capability of emergent phenomena or structures to changes of the environment. These phenomena are in dynamic interaction with the environment but are not totally dependent on it. Generic regularities and properties are abstracted away through self-organization and are applicable in other environments. In reality, the environment instantiates behavioural and structural rules raising the emergence of a global phenomena. For example, a bacteria following a sugar gradient can go towards a sugar source but as well follow such a source without modifying the "program" controlling the behaviour of the bacteria. This adaptability to external changes is immediate because it does not rely to internal representations nor internalisation of paths, and then, does not necessitate updating or modification of the representation. On the basis of the notion of emergence inspired by natural emergent systems as social insects, we can hope to build collective of miniature robots (with little intelligence and limited sensori-motor apparatus) which could be introduced in hostile and inaccessible environment and realize complex tasks of exploration or maintenance. Programmed with some simple behavioural rules, they could organize themselves in the space and adapt their behaviours depending on the context and therefore palliate robustness and flexibility problems usually encountered with traditional robots.

In the French multi-agent community, the concept emergence has gained considerable interest. We can cite J. Ferber [Ferber 95], A. Drogoul [Drogoul et al 92a, Drogoul et al 92b, Drogoul et Ferber 91 et 94], Y. Demazeau [Demazeau 93], N. Ferrand [Ferrand et al 97], L. Magnin [Magnin 96], P. Marcenac [Marcenac 96], J.P Müller [Müller et Pecchiari 96a et 96b] or R. Marcelpoil et al [Marcelpoil et al 94]. In the artificial life domain, we can cite the work of Deneubourg [Deneubourg et al 91a, 91b et 92], E. Bonabeau [Bonabeau et al 95, Bonabeau et Theraulaz 94], J.L. Dessalles, L. Steels [Steels 90 et 91]. The two papers of [Bonabeau et al 95]: "Characterizing emergent phenomena", give a number of examples of phenomena considered as emergent and propose un frame of study for a better understanding of these phenomena. Finally, we can make reference to the work of S. Forrest [Forrest et Miller 90, Forrest 90], in the domain of emergent computation we shall come to later on.

1.1 EMERGENCE IN PHILOSOPHY

The related concepts

As we shall see, the concept of emergence in philosophy is essentially opposed to the one of reduction, which consists in considering a phenomenon as nothing more than the underlying processes. The temperature is a classical example. The term "temperature" is used to denote a property of matter, which is nothing more than the average cinematic energy of constituting molecules. We shall see that the concepts of survenance, realization and epiphenomenona are also linked to the one of emergence. Before exploring the various concepts and in order to relate the domain of discourse, we begin with several notations and definitions:

Given two domain of discourse (or theories because we use this term in philosophy) D and D' each composed of a set of entities E and E' respectively, of a set of properties P and P' and of laws L and L'. Among these laws, we have to distinguish causal laws C and C' allowing to describe the dynamics and the general laws G and G' allowing to describe the composition or the conservation (invariants) in the considered domains, such that L=C U G respectively L'=C' U G'. Note that D and D' are not necessarily distinct by all the elements above. One can consider the same entities with different properties (P/=P'). Several relations are possible between these two domains of discourse, that we will explore by mentioning whether they are opposed to or simply different from the concept of emergence. It is possible that there is no relationship at all between two domains D and D' as, for example, between the ondulatory or corpusculary theories of light.

a) The concept of reduction

One distinguish several types of reduction [Scaglione 96]:

• E' is *ontologically reducible* to E when all the entities of E' are an entity or a composition of entity from E (for example, a gene is a set of segments of DNA);

• P' is *reducible by property* to P when all the properties of P' are properties or composition of properties of P (for example, the temperature is the average kinetic energy of the molecules);

• C' is *causally reducible* to C when all the causal laws of C' is a causal law or a composition of causal laws of C (for example, the laws of gas are reducible to the statistical laws of thermodynamics).

If one has simultaneously the ontological, causal and by property reducibility, then D' is reducible to D, or D' is just another way to talk about D [Nagel 61].

A number of remarks are necessary. In the preceding definitions, we do not make any difference between ontological distinctions (gene per se versus a set of segments of DNA per se) and epistemological distinction (what we call genes versus what we call segments of DNA independently of their actual existence). Effectively, to assume the existence per se of entities, properties and so on does not seem to add anything to the discussion. The notion of reducibility is asymmetric (D' reducible to D does not imply that D is reducible to D'). The asymmetry seems to go in the direction of a simplification by replacing composed structures by single entities, for example the molecular structure by a single object. This replacement is not arbitrary because it relies on the composition laws (included in G).

To situate emergence in this framework, the notion of emergence is coming from the irreducibility of some domains to others (for example, psychology to physics or life to biochemistry).

b) The concept of survenance

The concept of survenance only applies to the properties and more precisely when E=E'. It says that a property p2 survenes from a property p1 if when two entities have the property p1 then they also have the property p2. this notion is more than a simple covariance because it assumes a dependency of p2 on p1, therefore the survenance relation is asymmetric. The given definition is potentially ambiguous. The resolution of this ambiguity gives the following versions:

- *weak survenance*: if two entities have the property p1 then they have the property p2, everything else remaining the same.
- *global survenance*: if two entities in identical worlds from the point of view of p1 have the property p1, then they have the property p2.
- *strong survenance*: if two entities in two different worlds have the property p1 then they also have the property p2 (context independence).

Emergence could be a survenance but in the literature, emergence is defined as being different.

c) The concept of realisation

The concept of realisation applies both on properties and on laws. For the properties, the property (or properties) p of an entity e realises the property q if and only if e has both the properties p and q and it exists between p and q a necessarily explicative connection. We have to precise what constitutes an explanation.

First of all, an explanation is related to a theory if it explains changes and not only facts. This theory possesses both causal laws (C) and relationships between properties (G). Some causal laws cannot be explicative because it is not always necessary to understand why e acquired the property q and similarly for non-causal laws (for example, the coupling between electric conductivity and thermal conductivity does not explain anything). For example, an explanation linking the property p: "to have a given atomic structure" and q: "to be transparent" is imaginable. This last example suggests that p must be associated to the microstructure of e and q at the macrostructure.

The realisation relationships for laws are more complicated. If p realises q and p' realises q', which relationship must be established between the laws linking p and p' and the laws linking q to q' such that a law on p and p' realises the law between q and q'. Note that again the realisation on laws assumes in a way that D and D’ describe successive levels. We will not explore this path further here but the fact that the brain “realise” our mental states can be expressed by the realisation of the causal law linking a belief to an intention by the causal laws changing the brain state.

Is realisation always the converse of reduction ? It seems that it is not the case because of the example of transparency. In effect, the atomic structure can realise the transparency property but the phenomenon of transparency cannot be observed at the atomic level and is therefore emergent from the atomic microstructure. This defines realisation as a very general concept which relates two domains of discourse D and D’, D’ (or given aspects of D’) being either reducible to or emergent from D. Moreover, realisation explicitates the condition of survenance between two properties.

d) The concept of epiphenomenon

Let’s take the definition of the “ Petit Robert: ”

Epiphenomenon: 2. (end XIXème, from english.) *Philo*. Phenomenon which accompany an essential phenomenon without being part of its production or development.

An example is the shadow of the hand which is an epiphenomenon because its presence does not change the possible movements of the hand.

The British emergentism

We have said that the concept of emergence was essentially defined by opposition to the concept of reduction. To make it operational, the problem is to find a definition at least positive, at most constructive. One has to keep in mind that the concept can be defined differently depending on the emergence of an entity (a structure), a property or a law. We shall first describe part of the history of this notion before reviewing some definitions.

The first mention of the term “emergence” comes from the so-called British emergentism which takes place in the debate between four accounts of life:

- the *substantial vitalism* which states that it exists a substance linked to the living called entelechy;
- the *component theory* which pretends that it exists a component linked to the living as a variant of the vitalism;
- the *mechanistic theory* for which we are only machines (life is reducible to biochemical processes *)*;
- and finally the theory of *emergentism* (Lewes, S. Alexander, Broad, Stuart Mill: book of Broad "the mind and its place in nature", 1923).

This last theory is based on the work by Stuart Mill which distinguishes between two types of laws organising nature :

- the *homopathic* or *resultant* modality we can explain by causal laws or composition of them ;
- the *heteropathic* or *chemical* or *emergent* modality we cannot explain by causal laws as the acquisition of the properties of water out of the properties of oxygen and hydrogen respectively.

The debate is on the existence of these heteropathic laws which raises the problem of relating specific sciences (chemistry, biology, etc.) to physics.

The argumentation of the emergentists uses a vision of nature in levels articulated on top of one another, the physical level being the most fundamental, and on the ontological existence of the entities at a given level, emerging from the level immediately under it.

For the British emergentist, the entities at a certain level n+1 emerging from a level n exist if and only if they have a causal power and therefore allows the explanation of the phenomena at their level. If we can find a cause between n and n+1, we do not have emergence but reducibility. For having emergence, one must find a causality between n+1 and n or n+1 and n+1 (which is equivalent to finding a cause between n+1 and n). If there is only a cause between n and n+1, it is an epiphenomenona (e.g. the shadow of the hand).

In this argumentation, the nature and the existence of a descending causality is fundamental. One can distinguish two types of descending causalities (effect of the level n+1 on n):

- An effective macro-micro causality: one possibility is the teleological argument: something x realised y therefore y caused x, but this argument mixed up efficient and final causes producing a circular causality.
- The macro-structure has a causality on the micro-structure when the macro-structure allows the microstructure to exist over time: for example, a social organisation (macro level) ensures the existence of individuals which are part of it (micro level).

Note that in both case we have a circularity which is logical for the first case (and accordingly must be ruled out) and material in the second case.

Before going on, let's introduce the distinction between epistemological emergence which relates two discourses about nature and the ontological emergence which states the actual existence of the emerging entities. Given our formalisation, we can add two remarks on british emergentism:

- The domains D, D', etc. first constitute a strict hierarchy and second relate complete scientific fields (chemistry, biology, etc.)
- The concept of emergence applies on entities (E and E'), the problem being to validate the existence of the entities in E' by the existence of causal laws at their level (C').

More modern definitions shall extend this vision.

Recent definitions of emergence

Most of the modern tentative to define emergence rely on the critics of the British emergentism and more specifically to the not explicit :

- Nature of the relationships between parts and wholes;
- Nature of the properties we take into account;
- State of knowledge or more precisely the theoretical framework we are situated in.

But the concept of emergence seems to be deeply related to these aspects. Some definitions will take these critics into account.

1/Hempel & Oppenheim [Hempel & Oppenheim 48]

The occurrence of a property W of an object w is emergent relatively to a theory T, a relation to parts PT, a class of attributes G of the parts, if this occurrence cannot be deduced from the parts of W given by PT, respectively from all the attributes of G.

This definition has the advantage to show the characterisation of emergence of a property is linked to an object and more precisely to its composition. Then we find w in E', PT gives us the components Ci of E and, finally, we have G in P and W in P'. The theory T is not explicitly situated but is clearly in D (therefore identical to L). The fact that T does not allow to induce W on w justifies the construction of another domain of discourse D' in which W takes its meaning. The next definition makes more precise the relation PT and the nature of the theory T we want to consider.

2/ Cummins [Cummins 83]

Given the micro-structure of an object x : x = [C1, C2,....Cn; R] such that Ci are its components and R their organization. We have the following definition of what it means for micro-structure to possess a property :

X necessarily has the property F iff [if x = [C1, C2,....Cn; R] then F(x)] is a law.

To explain that [C1, C2,....Cn; R] produces F (and not just to notice it), we need a theory of the components Tc which, given F allows to deduce the [C1, C2,....Cn; R] having F. We define the micro-reduction as follows:

F is micro-reducible to [C1, C2,....Cn; R] iff it exists Tc (the component theory) allowing for each law relative to F to build its image in [C1, C2,....Cn; R].

We finally come to the definition of emergence:

F is an emergent property of the macrostructure x iff x = [C1, C2,....Cn; R] follows F:

1. When all the x= [C1, C2,....Cn; R] have the property F
2. When F is not micro-reducible

3/ Teller [Teller 92]

A property is emergent if it is not reducible to the non relational properties of its parts.

For example, the property to be the longest pen in a pen box is non-emergent in this definition but would be emergent if non-relational was not mentioned. We obtain a more precise characterisation of the properties to consider in D. Regarding D', Teller precises that the functional properties are never emergent. A part acquiring a function only in a whole (to discuss later), we get the case of the emergence of a property of an entity of D given a structure in which it is contained.

4/ Mario Bunge [Bunge 77]

Given P a property of a complex thing x distinct from the property « being a component of », then:

1. P is a *resultant* or *hereditary* property if P is a property of one of the components of the thing.
2. Said otherwise, if no component of x has the property P, then P is *emergent*, *collective* or *gestaltist*.

Bunge adds as a postulate that all the emergent properties can be explained from the properties of its components and the couplings between them.

5/ John R. Searle [Searle 95]

Let us cite what J. Searle writes: "Given a system S, composed of the elements a, b, c... For example S could be a stone, and its elements the molecules. In general there are some characteristics of S which are not, or not necessarily, characteristics of a, b, c… For example, it could be that S weights five kilos, but not the molecules taken individually. Let us call these characteristics "the characteristics of the system". The form and the weight are characteristics of the system. Some characteristics of the system can be deduced or conceived, or computed from the characteristics of a,b,c on the simple basis of their arrangement or of their composition (and sometimes from the relations they have with their environment) – for example the form, the weight, the speed. But other characteristics cannot be conceived only from the composition of its elements or their environmental relations; they must be explained in terms of causal interactions produced between the elements. Let us call them the "causally emergent characteristics of the system". Solidity, liquidity and transparency are some examples. From these definitions, consciousness is an emergent property of the system…This conception of causal emergence, we will call "emergence 1", must be distinguished from a more adventurous conception, we will call "emergence2". A characteristics F is emergent 2 if F is emergent 1 and F has causal power which cannot be explained by the causal interactions a, b, c… If consciousness is emergent 2, consciousness could cause things, which would not be explained by the causal behaviour of the neurons. The naïve idea which is at stake here would suggest that consciousness comes from the effect of the behaviour of the neurons of the brain, but just after, would live of its own life."

From Searle, consciousness is emergent 1, but not emergent 2. He precises that: "… I cannot thing of anything which would be emergent 2, and it is little probable that we could found characteristics which would be emergent 2 because such characteristics would be in violation of the weakest principle of the transitivity of causality."

6/ Miriam Scaglione [Scaglione 96]

From M. Scaglione [Scaglione 96], emergence is the inverse (in the mathematical sense) of the realisation. The distinction comes from the descending characteristics of realisation while emergence is ascending. Emergence becomes compatible with survenance which is one of its properties. One of the important consequences is that emergence is explainable without loosing its emergent nature. An important part of the above discussion comes from [Scaglione 96].

Conclusion

From this philosophical overview on emergence, it comes out that one has often searched, in a first time, to characterise an emergent phenomenon (or a property) (epistemological emergence), to finally raise the question of its existential truth (ontological emergence).

Let us notice the various given definitions to characterise an emerging "Whole" are often negative and relative to a set of parts constituting the "Whole" and/or to a theory T. The main criteria on which these definitions are grounded are the non-causality parts-whole, the irreducibility of the whole to the parts and the unpredictability of the emergent phenomenon, very often disputable.

The general tendency is the critics made to the definition of emergence as irreducibility, this definition being qualified as irrational. A rational definition would be the Bunge's one because it does not make reference to a theory but it greatly limits the definition. Another tendency is to make the notion of emergence relative to the considered theory. It does not necessarily means that emergence is a waiting room for better times (i.e. a better theory), but simply that the notion is relative.

1.2 OTHER RECENT DEFINITIONS OF EMERGENCE

One can notice that the notion of time, which is necessary for emergence of a phenomenon, is often lacking or implicit in the definitions in philosophy. Conversely, in definitions closer to our MAS domain, the dynamical aspect is essential: the definitions are on the process of emergence and not on the definitions of emergent properties. Let us look at some of these new definitions.

Emergent computation

A definition of Stephanie Forrest [Forrest 90] of the notion of emergent computation is particularly interesting because it is conceptually closer from what we want to formalize in MAS. Emergent computation is defined as:

- A set of entities in interaction: the process;

- An epiphenomenon produced by this process: a stable state, an invariant or an execution trace;
- The interpretation of this epiphenomenon as a computation or the result of a computation.

A first remark concerns the distinction between an epiphenomenon and an emergence; what is emergent is not the stable state, the invariant or the trace but its interpretation in a given vocabulary distinct from the vocabulary in which the process is programmed. For example with the ants, the emergent phenomenon is not the pheromone trace but its identification by the observer as a path between the nest and the food source.

As a second remark, an epiphenomenon is not an emergence because it as been created by the process but does not interact with the process itself. Automatically, we can create an emergence precisely when this feedback takes place. In order to do that, it is enough that the trace (it is the only physical reality, the stable state or the invariant can only be potentially perceivable when they let a trace) interacts with the process. At this moment something new is produced which nor in the process (because we need the trace) nor in the trace (because we need the process) and, if it stabilizes, becomes emergent simultaneously as a structure (the trace) and as a dynamics (the process). We have an illustration of the duality structure/dynamics. This kind of emergence is different from the first because it does not depend on the observer.

Emergence in cognitive science

Yves-Marie Visetti [Visetti 96] in the workshop « Emergence et explication » of the Cognitive Science Association defines three types of emergence:

- Emergence by the transition from a micro point of view to a macro point of view;
- Emergence by restructuring of an explanation of a system following an event;
- Emergence by the inscription of a system A into a system B in interaction with A. In the case where A=B, we have an autopoesis otherwise we can talk of an heteropoieses.

The first case is the emergence of Stephany Forrest in a definition which would be static and dependent on the observer. The last case is the extreme case, which would be dynamical and independent of the observer. Between the two cases, an intermediate case would be dynamical and dependent on the observer (because he must restructure his understanding of the system). A priori, we can use any definition and even all of them in the same system.

However, it remains to precise, in the case of the observer, what it means that two points of view are emergent from one another. At one end, to build the macro point of view as a sum (or any composition as complex as it could be) of some values of the micro entities is not emergent. At the other end, the average kinetic energy of a set of molecules is not emergent but calling it “temperature” and associate to this notion a phenomenology from the molecules would make it emergent.

For Minsky, the careful analysis of emergent phenomena “makes generally apparent that these phenomena can completely be explained when one take into account the interactions between the parts – as well as the particularities and limitations of the perception and expectations of the observer” [Minsky 88]. It is the precise position of Bertalanffy when he is talking about emergent properties: “the knowledge of the set of parts contained in the system and of the relations linking them allows to deduce from the behaviour of the parts, the behaviour of the whole" [von Bertalanffy 93].

Emergence in Economy/Management

“The economic order is an emergence, it is the non intentional and not desired consequence of a sufficient number of people driven by their own interests [Friedman 92]”. By defining this way the economic order, M. Friedman clarifies the notion of emergence in which we would be tempted to qualify the actions as local. Moreover, a number of other notions are associated: stability, lack of intentionality, novelty, unpredictability, continuity, irreversibility.

Let us first consider the idea of a relative stability. F. von Hayek considers that social emergent orders exist and for H. Mintzeberg [Mintzber 81] an emergent strategy is a non intentional order. In fact, to be identified, a phenomenon, a behaviour, a process, etc. which emerges must be sufficiently stable, hence this idea of order, evoking stability.

We also find the idea that there is non central direction associated to the idea of emergence: it is not possible to pilot nor to control it. At the most could we attempt to predict that there will be emergence (without being able to predict what will emergent nor when it will emerge); it is the notions of novelty and unpredictability. A phenomenon, a behaviour, a process, etc which emerges at the global level, results from the interactions of the behaviours, phenomena, local processes, both egoistic and myopic (for example, each behaviour in its own context and in reference to its own finalities).

But, still maintaining the idea that emergence cannot be piloted, one propose to go further than spontaneous (emergent) social order (Hayek): "Human action but not human design" to "human action for human design". Said otherwise, one try to restore the idea of intentionality (teleology). By exposing myself to "noise", I create the conditions for the emergence of non-programmed occurrences (occurrences of not yet known possibilities). For example the nomad (unlike the sedentary), by its behaviour, creates the conditions of the emergence of not yet known possibilities. Still maintaining the idea that emergence cannot be piloted, one add that we can, intentionally, attempt to create favourable conditions.

Finally, the notion of emergence is also characterised by the ideas of continuity and irreversibility. Effectively, this notion has to be understood in its temporality, resulting from processes, phenomena, behaviours taking place in time, and consequently as being irreversible.

1.3 TOWARDS A DEFINITION OF EMERGENCE IN MAS

At the light of the various definitions, in philosophy, in computer science, in economy/management or in cognitive science, it seems essential to come up with a positive, temporal (where time appears explicitly) and constructive definition of emergence.

A definition initiated by Ch. Lenay [Lenay 96] proposes this positive and dynamical approach of emergence. To begin, it is important to distinguish some characteristics of emergent systems implying the subject and its environment and where emergence results from the interaction between both.

The first essential feature of a multi-agent system is that no agent controls entirely the dynamics of the population. The agents are limited and there are differences of a global system they are unaware of. Therefore, there is an exteriority relative to each agent: an environment.

The second feature is that, by definition, the agents act and therefore modify this environment. But, the agents can only perceive and act locally in this environment. Said otherwise, each agent interprets the environment given his limited means (using the distinctions it is able to make).

The third feature is that the exterior of each agent contains other agents. There are several agents in a common environment (they are exterior to each other). The interpretation of the environment by the various agents can possibly be different. In the case of reactive agents, the environment contains the objects and the other agents. In the case of cognitive agents, the environment can also contain messages.

Therefore, the dynamics proceeds by iteration of interpretation of the local environment by the agents, action of the agents on this environment, new interpretation of the modified environment, new actions, etc.

When such a dynamics (or some of its components) stabilizes, we can talk of emergence of a structure or of a global functionality. Notice that at any moment, it is the environment possibly modified by all the other agents (and itself) that each agent submits to its interpretation. It is the condition for the global dynamics to by more than a simple sum of independent dynamics. If it was the case, we could only speak of emergence in a very weak sense and only for an external observer. But whenever, through the environment, the whole feedbacks on the parts, there is emergence in a strong sense, emergence for the agents if it is the global and stable emergent state which selects the individuals behaviours of each agent.

In this definition, the dynamics of interaction is postulated as a basic condition or emergence of phenomena, structures, etc.. Also notice the importance of the link whole-parts which characterises the various kinds of emergent phenomena. In the following, we will derive a more operational definition by characterising the whole and the parts and most importantly the feedback whole-parts. This definition is inspired by the preceding definitions and moreover the ones by S. Forrest and M. Bunge and postulates:

- A system of entities in interaction whose expression of the states and dynamics is made in a vocabulary or theory D;
- The production of a phenomenon which could be a process, a stable state, or an invariant which is necessarily global regarding the system of entities;
- The observation of this global phenomenon either by an observer or by the entities themselves.

This observation can only be done through an inscription of the phenomenon on one hand and on the other hand the interpretation of this inscription by the observer or by the entities themselves in a vocabulary or a theory D' distinct from D. A theory of emergence would be a theory D0 of inscription by a system of entities in interaction and its interpretation.

The preceding definitions of emergence in philosophy allow to clarify which relation must exist between D and D' to be able to really talk of emergence, namely that D realise D'. In all cases, this definition allows to take into account the definitions 1 and 3 of Visetti [Visetti 96]. Finally, notice that this definition distinguishes precisely two levels: a micro level (the one of the interacting entities) and a macro level (the set of entities). From this set of entities, one considers a global production (stable state) from this global set.

Another issue raised by this definition is the problem of the interpretation of the inscriptions which provides two possible meaning to emergent phenomena. A weak sense where the inscriptions refer to a same reality understood at two levels that it is useful, or even necessary to distinguish. The problem is summarised as "another way to talk about things" which would be simpler as the temperature regarding the kinetic energy of the molecules. It is the emergence in the definition 1 of Visetti. Finally notice that this emergence can be linked to the ignorance of the observer or its inability to formalize an underlying compositionality. Without assuming a two radical dualism, we can consider emergence in the strong sense if the inscription which is globally produced does not have the same order of reality than the individual productions.

How to use these definitions ? Or how to apprehend emergence in MAS? The natural approach would be to situate it with respect to emergence observed elsewhere. In effect, the classical approach of modelling and explaining emergence goes from the observation of natural phenomena to the reproduction with artificial systems. In certain cases, modelling is restrained to an interpretative attitude of the observed phenomena without being able to validate the proposed hypotheses. It is the case for some collective behaviours observed in ants societies (altruism,…) or in human societies. The models in Artificial Life attempt to reproduce these hypotheses and their conditions of emergence of some phenomena in simulation, that allows to get a better understanding of these phenomena (but uses strong presuppositions on the sense of the term emergence). The extension of the obtained results allows then to adopt a predictive attitude by creating the conditions of emergence of new artificial phenomena by simulation or by experimenting (where there are other technological constraints). It is the case of collective robotics where the specification and the combination (sometimes random) of basic behaviours of a set of robots, can produce emergent global behaviours. It is also the case in programming when it is the only way to solve the problem of trivial compositionality: all what is made in computer science is necessarily the combination of the execution of elementary instructions. But talking of computer science, is to go further than physical processes. The emergentist approach provides a way to go further. The relationship between emergence elsewhere and emergence in computer science goes through inscriptions.

Finally, it is important to notice that the justification of a global production (i.e. the behaviour of the system) as emergent is not only its adequacy to the definition, but also its subordination to the existence of an emergence in the strong sense.

1.4 Conclusion et perspectives

Emergence is invoked each time we have to deal with complex systems and when we have to realize systems able to adapt to dynamic and uncertain environments. In both cases, the very concept of emergence raises problems because of the lack of constructive definitions. In this part, we made an attempt of reviewing most of what has been said on emergence. This allowed to propose a positive definition (and not by absence of something) from which we hope to derive a concrete methodology of description and design of complex systems. At least it allows us to argue that it could exist an emergentist alternative to the classical method for problem solving.

The emergentist alternative relies on a set of interacting entities between themselves and with an environment and on an interpretation mechanism mediated by an inscription substrate. By its main focus, i.e. sets of interacting entities, MAS constitutes with neuro-mimetic networks and connexionism in general, a privileged mean to study and to deploy an emergentist alternative. In other terms, MAS does monopolize the concept of emergence which is also invoked, as we have seen, in Artificial Life and the natural sciences. This last point suggests that we can envision emergence as a new paradigm, transversal to several fields of computer science, whenever the concept of emergence stops to be the waiting room of human understanding to become a constructive concept.

Regarding observation, we have seen that an observer of the global phenomena is necessary for having emergence. This point raises an ambiguity because the observer can stay external to the system in which case he is requested an interpretation effort in a direction which is not intrinsically contained into the system dynamics, otherwise the observer can be himself part of the dynamics of the whole through its capabilities to interact with the system, in which case he is himself producer of the emergent phenomena. In this second case, the interactions of the user/observer can produce by emergence more than the user or the system could perform independently.

Computer science already requests from the user to look at the computer in terms of the tasks he is performing, independently of the detailed manipulation of the machine. In the same way, emergence in computer science requests the observer/user to look, independently from the simple (= local/micro) interaction phenomena, to the global phenomenon these interactions realize. This makes explicit the interpretation phenomena in computer science (and their acceptability). That opens doors towards a better understanding of human-machine cooperation. Nevertheless, this opens a new perspective in addition to those consisting in developing tools to validate and exploit our proposition of concept of emergence in concrete realizations.

2. Designing emergentist multi-agent systems

Designing multi-agent systems is distinct from designing other types of classical computer systems, in particular distributed systems, by at least two dimensions:

- The importance of interactions between the agents which should exceed the importance of the agent architecture itself if we want the whole to be more than the sum of its constituents;

- The role of the environment simultaneously as a place of inscription and a set of constraints on the multi-agent system dynamics. In distributed systems, the environment is often reduced to the communication channels between the processes. However, if these channels can be dynamically reconfigured, it is essential to introduce spatiality and not only connectivity in order for the connectivity to be deduced from the spatiality and not the converse.

The design of multi-agent systems raises new challenges and becomes a major issue after a somewhat exploratory phase necessary to any new and expanding domain. We shall distinguish three types of design approaches:

- The agent-oriented approaches which focus on the individual agents and propose specification formalisms of their behaviours with various tools (agent-oriented programming by [Sho93] or Rao [Rao96], temporal logics approach by Wooldridge [Wool96]. These approaches are distinct from pure mono-agent approaches by the introduction of communication up to complex negotiation protocols.
- The organisational approaches which deal with the specification of interactions through the notions of roles, relations between roles and groups either to statically specify interaction networks as in [Dur96] and more recently in [Gut98], either in a more dynamical perspective as the Cassiopee method [Col96] or the Aalaadin framework.
- The emergentist approaches distinguishing a micro-level of interacting agents from a macro-level where the desired global phenomenon is produced, either an organisational structure, the realisation of a task or the building of a solution to a problem. These approaches must therefore articulate these two levels thanks to a positive definition of emergence as proposed earlier and, for example, used in [Lab98a].

These various approaches are in fact complementary in the sense that the agent-oriented approaches specify the entities in interaction, the organisational approaches are an important tool to specify what we want to obtain at a macro-level, et finally, the emergentist approaches insist on the interactions and on the micro/macro articulation, making the link with organisational structures (at the global level) using the interactions (at the local level). Separately, they suffer limitations because the agent-oriented approaches can hardly take into account group dynamics, the direct implementation of organisational structures hardly grasps reorganisation dynamics and the emergent approaches still lack sufficiently general methodologies to articulate the micro-level and the macro-level.

To go towards a design methodology of multi-agent systems for problem solving by emergence (excluding multi-agent systems for simulation or cooperative work as with software agents and CSCW), one can start either from the proposed definitions of emergence, or from an analysis of existing systems, natural or artificial, exhibiting emergent properties in order to determine their common features and to deduce a heuristic methodology if not a systematic one. This last approach is proposed in this chapter taking the definition of emergence developed earlier as a reading grid.

After having fixed the vocabulary on problem solving and having reminded and commented the definition of emergence exposed earlier, we are going to present a number of multi-agent systems obeying the following criteria:

- Their aim is to perform problem solving, and even optimisation;
- The structure of the state space is not explicitly manipulated but emerges from the MAS dynamics;
- The solution that results can adapt dynamically to changes of the problem data.

We will present a synoptic table of their common structure and derive a systematic methodology. Finally we will conclude before applying it to the problem of designing a teaching system.

2.1 PROBLEM SOLVING

In order to fix the vocabulary, we shall define formally what is problem solving (see, for example, [Chr82]). This clarification will allow us to compare more easily the classical algorithmic approaches from the multi-agent approaches in general and emergentist ones in particular. A *problem* is specified by:

- A *search space E* constituted by an finite or infinite, discrete (combinatorial) or continuous set of states *{ei}* ;
- A subspace *S* in *E* called the *space of solutions* (or admissible states).

The problem is dynamic if the state space and/or the solution space evolve over time.

The search space must be described by a set of components and of parametrisation and composition operators allowing to generate the search space (i.e. the set of states represented as structures made of these components). This structure is often given as a set of variables *vi* and their domain of definition *Di*. In our case, it will be easier to describe the set of components *C* and their parametrisation and composition operators we will call the *structure S of the search space*.

The set of solutions is expressed, in general in intention by a set of constraints. When the definition does not allow to directly build one or several solution states, we need a *search method*. One distinguish to classes of search methods:

- *Restriction search* consists in reducing the search space by incrementally fixing the parameters, the components and their composition until one obtains a state or a sub-space of the solution set. When one obtain a subspace in which no state can be a solution, a backtrack is performed (see, for example, CHIP or PrologIII);
- *Repair search* that consists in building a random state and to modify whenever it does not satisfy the solution criteria (i.e. whenever it is not in the solution space) (see [Ghe93]). Notice that the process goes along a trajectory in the search space and therefore can be reformulated as a *control problem* [Dea91], *optimal* if we want to optimise the search time and *stochastic* if the search is stochastic.

Any combination of these two methods is possible.

One can define on the search space an *objective function F* and search a state that is the best solution state in the sense that it optimises this function *F* (minimum or maximum). In such a case, we have an *optimisation problem*.

2.2 DEFINITION OF EMERGENCE

The concept of emergence is important to describe how the macro level relies on the interactions of the micro-level. The advantage is the possibility to design a micro level which auto-organizes depending on the circumstances while ensuring an invariant at the upper level, hence providing flexibility and robustness. Here, we are just reminding the previously proposed definition by saying that an phenomenon is emergent if:

- There is a set of interacting agents whose dynamics is not expressed in terms of the emergent phenomenon to produce;
- The dynamics of the interacting agents produces a global phenomenon which can be a stable structure, an execution trace or any invariant either static or dynamic, synchronic or diachronic;

- The global phenomenon is observed either by an external observer or by the agent themselves in terms that are distinct from the underlying dynamics.

For the global phenomenon to be observed, it must be inscribed on a support. In natural systems (and more and more in artificial ones), the environment plays this role of inscription medium.

As a first example of natural multi-agent system producing an emergent phenomenon, we can take the ants that produce a path between a food source and their nest. An ant transporting food deposits pheromones. These pheromones diffuse, producing a gradient attracting other ants. As a result, a path is produced between the source and the nest as long as there is food available. The individual behaviour of the ant entirely relies on local perception of the pheromonal gradient. The interactions take place between the ants and the environment or more precisely between the ants but mediated through the environment. The global phenomenon, i.e. the path, is identified as such by an external observer (most probably, the ants are not aware of the path as such). We have the three conditions satisfied:

- A set of agents interacting in terms of pheromones (and not in terms of paths);
- The production of a stable global phenomenon (as long as there is remaining food): i.e. the ants going back and forth between the nest and the food source;
- The observation of this global phenomenon in terms of a path through an inscription which is the pheromone trace, or, more easily observable, the ants themselves along the trace.

Notice that the system is flexible because its behaviour is related to the pheromone gradient and not the produced path. The path being a side product can dynamically change depending on new obstacles or change of food source position; If the dynamics would depend on the path and then of a representation in the ants'heads, the decision process to change both the path and its representations would be much heavier. It is this advantage of emergence we would like to exploit in emergentist approaches of problem solving and not only the intrinsic efficacy of parallelism.

These thoughts allows us to specify what we will be interested in the examples we want to analyse in the next section, namely:

- The agents and the nature of their interactions, i.e. in which terms are their behaviours expressed;
- The global phenomenon, namely the search space E in its components C;
- The mode of composition of the global phenomenon by the elementary interactions.

2.3 SOME EMERGENTIST MULTI-AGENT SYSTEMS

Among the first multi-agent systems to make explicitly reference to the emergence of a solution to a problem by side-effect of the interaction dynamics, we must cite the eco-resolution of Ferber [Fer90] from ethological inspiration. It has been applied to a number of problems among which:

- The blocks world whose problem consists in finding a sequence of executable actions to go from an initial configuration to a final configuration. In the proposed solution, the agents are the blocks themselves that interact on the basis of their relationships. These interactions produce movements whose succession will generate a plan and even, in this case, an executable and acceptable plan. In this case, the search space is the space of possible configurations of the blocks and the emergent solution is the trajectory in the search space [Fer90].
- The magic square consists in putting the tiles in a final configuration by moving one tile at a time. In this case, the agents are the tiles that will push one another in order to go to their final place and, hence, generate the necessary movement sequence [Dro92].

This approach is also applied to chess to show the possibility to have the emergence of a global strategy from the local interactions of the chess pieces [Dro93].

Another kind of emergentist multi-agent systems relies on dynamical systems like PACO [Dem93] and SMARPS [Ferr97], which have been applied to:

- Contrast line detection in computer vision. The agents are located on the pixels of the image and will climb the local intensity gradient trying to keep a given distance from a given number of agents (in general, two) and being influenced by their movements. Lines of agents will eventually stabilize on the contrast lines and even follow them if the image changes sufficiently slowly with respect to the agent own dynamics. It is one of the examples where we have an explicit environment, i.e. the image [Dem93].
- Cartographic generalization in which the agents are geographic entities (houses, trees, road segments, rivers, etc.), which will enter a competition to appear on a map of a given resolution. The place occupied by the graphics (lines, icons, etc.) will constrain the other agents to move slightly and/or to change their graphic up to invisibility if there is not enough place. The search space is constituted of all the possible positions with all the possible graphics. The actual position and graphical representation will emerge from the interactions under the resolution constraint [Bae95].
- Production of a set of possible positions of linear structures (roads, power lines, etc.) under multi-criteria constraints in land planning. Here, we have multiple environments describing the spatial constraints from various points of view [Ferr97].

We can also cite AMROSE [Ove94], a multi-agent system to control an articulated robot to sold steel “plaques”. In this case, the rigid parts of the robot are the agents trying to go to a certain position without touching the obstacles or the other parts of the robot. The result is a trajectory of the end tool of the robot resulting from a sequence of commands to the joints.

We can also mention more applied multi-agent systems using optimisation processes as:

- MARSA which is dynamic scheduling system for flow-shop workplant [Dao95]. The system is coupled with a workplant simulator providing various events as: command arrivals, the beginning and end of lot treatments, set-up and breaks. The scheduler provides back the next lots to produce. By providing only a few lots in advance, the scheduler is able to dynamically revise its schedule on incoming events. The agents are the commands trying to by produced in accordance with their associated deadlines and the machines trying to minimise their set-up time. The interactions are formulated in terms of allocation (and not of temporal order) with an implicit gradient produced by the optimisations to realize. The result is a schedule of the commands as produced for a production campaign. A similar system has been deployed by Daewoo in Korea using 30 machine agents and about 700 task agents [Chu97].

- AMACOIA is a design system of assembly lines. Given the description of a product to assemble and contract cycle time (time between the finishing of two parts), the system computes a functional description of an assembly line with minimal cost. Two multi-agent systems are coupled: one to explore the space of assembly sequences taking into account the product constraints, the other one to explore the space of possible assembly lines taking into account the production constraints. In the first system, the agents are the links between the sub-assemblies (and not the sub-assemblies themselves) trying to place themselves into the assembly sequence. In the second system, the agents are the operation realizing the links competing to be placed on assembly posts, themselves into cells, and the latter into the assembly line. In each case, the dynamics is in term of assignment, but the result is respectively a temporal and a topological structure.

These cases easily allows to see how emergence can provide a better adaptivity of the multi-agent system as a whole. Now we shall make the synthesis of these examples on the basis of the definition of emergence and problem solving as defined earlier.

2.4 TOWARDS A METHODOLOGY

The definition of emergence suggests a number of steps to defined a multi-agent system:

- The formal description of the global phenomenon the multi-agent system must realize;
- The projection of this global phenomenon on the interaction structure at the micro-level to determine the identities of the agents and the interaction dynamics;
- The specification of individual behaviours of the agents to produce the interactions generating the global phenomenon we want to observe.

The lack of direct connection between the macro and the micro levels calls for validation tools to guarantee effectively the emergence of the desired global behaviour. This methodology has been entirely developed in [Lab98a] and [Lab96,Fer97,Lab97].

The preceding examples allows to detail further this methodology. In effect, all the case have been described as search processes of a solution in search space. The structure *S* of the search space can be:

- Spatial: as a path, a map, contrast lines;
- Relational: as a schedule or any relational configuration as a logical proof considered as a deduction relation on formulas;
- Spatio-temporal: as the tasks to be performed by a collective of robots [Lab98a].

This structure is a composition of elementary entities *C*:

- Spatial: pheromones deposits, graphics, proximity links;
- Relational: relations, assignments;
- Spatio-temporal: movements, force applications, etc..

In classical programming, we would represent a state of the search space as a data structure to be built and modified by a given algorithm. This algorithm would depend implicitly or explicitly on all the constraints, initial hypotheses and state to elaborate its solution. The problem has to be closed. Any modification would require to stop computing or even to change the algorithm itself. If the initial data change over time and that new constraints are added or removed dynamically, the approach becomes extremely difficult to problem because the problem becomes open.

In the multi-agent systems we just described, the state is not explicitly manipulated by the agents. The agents interact with one another and with the environment in a way which is indirectly related to the state we want to manipulate. For example, the agents "task" or "command" are seeking to be placed with given criterion as the availability of the resource and their deadline and not to placed before or after another one. Therefore, the schedule is only an indirect outcome of these interactions. The ants follow the pheromone gradients and bring food. The resulting pheromone deposits which constitute the path is only an indirect effect which feedbacks locally on the ant behaviours. In AMACOIA, the agents are the links between the parts when the result is the order on the operation. In a similar way, the assembly line is a structure imposed on the tools, stations, cells and not the tools, stations and cells themselves (which are the agents). The same logics can be seen in the case of AMROSE.

We have systematically reported these observations on the described multi-agent systems in the next table. The first column is the search space structure S, the second are the components C, the third column exhibits the chosen agents and the last one on what they interact. On easily distinguish the multi-agent systems having an explicit environment from the others. In this last case the agents interact depending on their own internal state.

MAS	Structure	Components	Agents	Interactions
Ants	Path	Pheromone deposits	Ants	Pheromones
Termites	Nest	Ground deposits	Termites	Pheromones
Allocation	Global assignment	Individual assignment	Tasks and resources	Availability
MARSA	Total order	Temporal position	Commands and machines	Deadline and set-up time
AMACOIA	Assembly sequence	Relative position	Links	Product constraints
	Assembly line	Posts	Operations	Production constraints
Blocks world	Configuration	Block relations	Blocks	Freedom
Magic square	“	Tile positions	Tiles	Freedom
PACO	Contrast lines	Points positions	Points	Intensity gradient
	Map	Positions and graphics	Geographic entities	Freedom
AMROSE	Trajectory	Parts positions	Parts	Gravity, obstacles

Table 15. Observations on multi-agent systems

We would like to comment further on the role of the environment and the distinction to make between agents and processes:

1. We have two distinct roles of the environment in the described systems: one is to contain the state of the search during the solving process such that it can indirectly feedback on the multi-agent system dynamics, the other is to provide exogeneous constraints on this dynamics as the time line in the scheduling systems, the obstacles in the ants paths and, more generally, restrictions on the search space. Notice that the agents are themselves situated in the environment and this “physical” presence is also source of interaction and constraints, for example, in the blocks world or the magic square where interactions take place because the agents are in the way of other agents.
2. We have to know whether we have to really specify an agent or just a dynamical process taking place in the environment (because the environment can have his own dynamics). It is enough to refer to the definitions of agents (as, for example in [Fer95]), which insist on the autonomy of the agents in the form or more or less explicit representation of their goals. In the case of problem solving, it is enough to identify the choice points, i.e. the components of the search space structure, which allows to potentially explore the whole search space (it does not mean we will have to do it exhaustively). These choices must result from the interactions between the agents. The rest of the search space structure can be processes propagating the consequences of theses choices if necessary. For example, in dynamic scheduling the choice is a date of production, recomputing the dates of the other jobs placed before or after is just the computation of the consequences of this choice and should not be taken in charge by an agent. In the case of simulation we are not talking about in this paper, what is a process and what is an agent is less clear and is more related to the interpretation of the observer of the system [Bat96].

We are now at the position to propose a methodology derived from this analysis:

1. To specify the search space and the structure of its possible states S;

2. To determine the elementary components C from which the states of the search space are made and among them the choice points determining the change from a state to another and potentially guaranteeing the exhaustive search through the search space (ergodicity condition);
3. To determine the entities whose interaction will produce these components. We obtain the agents of the system as a kind of negative image of the structure to produce.
4. To determine the objectives and the dynamics of these entities allowing to go through the search space. We obtain the production mechanisms of the interactions which will by side-effect go through the search space and possibly converge towards a solution;
5. To determine the exogeneous constraints guiding the trajectory and potentially forbid some parts of the search space. This allows with the inscription of the search space to completely define the environment.
6. To determine the processes propagating the consequences of the actions of the agents, defining the dynamics of the environment itself. It is also possible that the exogeneous constraints are directly linked with the world external to the multi-agent system (for example in the workplant scheduling, to the workplant itself), in which case the proper dynamics becomes more than a simple reaction to the agents interactions.
7. To validate the design either experimentally if the multi-agent system is too complex or theoretically by using either non-linear dynamics, Markov chains [Lab98b] or statistical dynamics.

This methodology demonstrates clearly why a multi-agent system is more adaptable and flexible than a classical algorithm. In effect, the solution is not computed explicitly by the multi-agent system but emerges from the interactions between the agents, which are dynamical relationship with the problem data and the constraints on the possible solutions. This dynamic formulation of accounting for the data and the constraints allows the multi-agent system to react spontaneously to the modifications either while trying to find a solution, or after a solution has already been found. In reference to the section on problem solving, we are in a logic of search by repair and therefore in the paradigm of control. In effect, a solution appears as an invariant or a stable state of the dynamics of the multi-agent system. However, this formulation raises the problem of the observability of the stable state, which represents the solution we are looking for. It is the reason why the notion of emergence puts forward the notion of observer. We will detail two reasons to this:

- The multi-agent system may not know that he found the solution (as it is the case for contrast lines) in the sense that no single agent can locally decide it but only a global observer of the system can know it.
- The medium and the inscription process takes all its importance because it is this way the observer will be able to observe the solution (as the drawing of the links between the agents on the image to visualize the contrast lines). The inscription process can also constitute a discretisation process allowing to observe a stable state where the multi-agent with a continuous dynamics is in a chaotic attractor.

This last remark justifies the conclusion of the part on the notion of emergence [Jea97] that suggests that a theory of emergence has to use a theory of inscription and interpretation and, therefore, calls for semiotic thinking.

References

[AMY 92] B. AMY, “Emergence et Immergence : les adrêts et les ubacs de l’organisation”, Actes des journées : Emergence dans les modèles de la cognition, ENST Paris, 15-16 avril 1992.

[AVENIER 92] M-J. AVENIER, “Le management stratégique revisité : La problématique de -management” N.R. GRASCE n°92-16.

[AVENIER 94] M-J. AVENIER, “L’ingénierie d’actions stratégiques en milieu complexe : Repères” N.R. GRASCE n°94-11.

[AVENIER 96] M-J. AVENIER, "La stratégie tâtonnante" : des interactions récursives entre vision et action stratégiques ", N.R. 96-01, GRASCE.

[BIRNER 96] J. BIRNER, "Mind, Market and Society, Network structures in the work of F. A. Hayek", working paper, 1996.

[BONABEAU et al 95] E. BONABEAU, J.L. DESSALLES, A. GRUMBACH, "Characterizing emergent phenomena (1) and (2),: A critical review", in Revue internationale de systémique, vol. 9, N°3, pp 327-346, et pp 347-371.

[BONABEAU et THERAULAZ 94] E. BONABEAU, G. THERAULAZ (Eds.), "Intelligence collective", Ed. Hermes, 1994.

[BOURGINE et VARELA 92] P.BOURGINE, F.VARELA, "Toward a practice of autonomous systems". Ed. F.varela, P.Bourgine, MIT Press, 1992.

[BROOKS 89a] R.A. BROOKS, "A Robot that Walks: Emergent Behaviors From a Carefully Evolved Network", Massachusetts Institute of Technology, Cambridge, MA, A.I Memo. 1091, 1989.

[BROOKS 89b] R.A.BROOKS, A.M.FLYNN, "Fast,Cheap and Out of Control" ref ? 1989.

[BROOKS 91] R.A.BROOKS, "Intelligence Without Reason" IJCAI'91, 1991.

[BUNGE 77] M. BUNGE, "Emergence and the mind", in Neuroscience, 2:501-509, 1977.

[CAMPS et GLEIZES 96] V. CAMPS, M.P. GLEIZES, "Attitudes coopératives individuelles pour une capacité d'adaptation collective", journées françaises sur les systèmes multi-agents, JFIADSMA, Avril 1996, Sète, France, Pages 153-166.

[CARIANI 90] P. CARIANI, "Emergence and Artificial Life". Artificial Life II; pp. 775-797, Ed. C. Langton, C.Taylor, J. Doyne Farmer, S. Rasmussen, Addison-Wesley Publishing Company, 1990.

[CARIANI 91] P. CARIANI, "Adaptivity and emergence in organisms and devices", in World Futures, vol. 31, pp. 49-70, Gordon & Breach science Publishers, 1991.

[CUMMINS 83] R. CUMMINS, "The nature of Psychological Explanation", MIT Press, 1983.

[DEMAZEAU 93] Y. DEMAZEAU, "La Plate-forme PACO et ses Applications", 2èmes Journée Nationale du PRC-IA sur les Systèmes Multi-Agents, PRC-IA, Montpellier, France, décembre 1993.

[DENEUBOURG et al 91a] J.L. DENEUBOURG, S. GOSS, N. FRANKS, A. SENDOVA-FRANKS, C. DETRAIN, L. CHRETIEN, "The Dynamics of Collective Sorting : Robot-Like Ants and Ant-Like Robots". From Animals to Animats (SAB'91); pp. 356-363, Ed. J.E. Meyer, S.W. Wilson, MIT Press, 1991.

[DENEUBOURG et al 91b] J.L. DENEUBOURG, G. THERAULAZ, R. BECKERS, "Task Differentiation in Polistes wasp colonies: a model for self-organizing groups of robots". From Animals to Animats (SAB'91); pp. 346-355, Ed. J.E. Meyer, S.W. Wilson, MIT Press, 1991.

[DENEUBOURG et al 92] J.L. DENEUBOURG, G. THERAULAZ, R. BECKERS, "Swarm-Made Architectures". Toward a practice of autonomous systems; pp. 123-133, Ed. F.varela, P.Bourgine, MIT Press, 1992.

[DROGOUL et al 92a] A. DROGOUL, B. CORBARA & D. FRESNEAU, "Applying EthoModelling to social organization in ants". Biology and Evolution of Social Insects, pp. 375-383, Billen Ed., Leuven University Press, Begium,1992.

[DROGOUL et al 92b] A. DROGOUL, J. FERBER, B. CORBARA, D. FRESNEAU, "A Behavioral Simulation Model for the Study of Emergent Social Structures". Toward a practice of autonomous systems; pp. 161-170, Ed. F.varela, P.Bourgine, MIT Press, 1992.

[DROGOUL et FERBER 91] A. DROGOUL, J. FERBER, "A Behavioral Simulation Model for the Study of Emergent Social Structures", European Conference on Artificial Life, Paris, 1991.

[DROGOUL et FERBER 94] A. DROGOUL, J. FERBER, "Multi-agent simulation as a tool for studying emergent processes in societies", in Simulating Societies: the computer simulation of social phenomena, N. Gilbert and J. Doran eds, North-Holland, 1994.

[DROGOUL 93a] A. DROGOUL, "When Ants Play Chess (Or Can Strategies Emerge from Tactical Behaviours?)", in "from Reaction to Cognition", : 5th European Workshop on Modelling an Agent in a Multi-Agent World - MAAMAW'93, C. Castelfranchi and J-P. Müller eds, Springer Verlag, 1996.

[DROGOUL 93b] A. DROGOUL, "De la simulation Multi-Agents à la résolution collective de problèmes. Une étude de l'émergence de structures d'organisation dans les systèmes Multi-Agents", Thèse de l'Université de Paris VI, Paris, Novembre 1993.

[DUPUY 92] J-P DUPUY, "Introduction aux sciences sociales; Logique des phénomènes collectifs", Edition Ellipse, 1992.

[FERBER 95] J. FERBER, "Les systèmes multi-agents", InterEditions, 1995.

[FERRAND et al 97] N. FERRAND, Y. DEMAZEAU, C. BAEIJS, "Systèmes multi-agents réactifs et résolution de problèmes spatialisés", Revue d'Intelligence Artificielle (à paraître 1997).

[FLYNN et BROOKS 89] A.M. FLYNN, R.A. BROOKS, "Battling Reality". 1989.

[FOISEL et al 96] R. FOISEL, V. CHEVRIER, J.P. HATON, “De l’organisation d’une société à sa réorganisation”, Journée thématique du PRC-IA sur les systèmes multi-agents, Toulouse, Février 1996, pp 121-128.

[FORREST et MILLER 90] S. FORREST, J.H. MILLER, “Emergent behavior in classifier systems”, in “emergent computation”, MIT Press, Cambridge, MA, 1990

[FORREST 90] S. FORREST, “emergent computation: Self-organizing, Collective, and Cooperative Phenomena in Natural and Artificial Computing Networks”, introduction to the proceedings of the ninth annual CNLS Conference, in “emergent computation”, MIT Press, Cambridge, MA, 1990, pp 1-11.

[FRIEDMAN 92] M. FRIEDMAN, “Free to choose”, 1981, cité par J-P Dupuy, in “Introduction aux sciences sociales; Logique des phénomènes collectifs”, Edition Ellipse, 1992,p199.

[GRUMBACH 96] A. GRUMBACH, Emergence et niveaux d'organisation : un problème d'observateur, in Journées thématiques de l'ARC (émergence et explication), 1996.

[HEMPEL & OPPENHEIM 48] C.G. HEMPEL & P. OPPENHEIM, “Logic of explanation”, Philosophy of science, 15:135-175 - 1948

[HOFSTADTER 79] D. HOFSTADTER, “Gödel, Escher, Bach: An eternal golden braid”, basic books, 1979

[IKEGAMI et KANEKO 90] T. IKEGAMI, K. KANEKO, “Computer symbiosis - emergence of symbiotic behavior through evolution”, in “emergent computation”, MIT Press, Cambridge, MA, 1990, pp 235-243.

[KANADA et HIROKAWA 94] Y. KANADA, M. HIROKAWA, “Stochastic Problem Solving by Local Computation based on Self-organization Paradigm”, 27th Hawaii Int. Conf. on System Sciences, 1994, pp. 82-91.

[KAMPIS 91a] G. KAMPIS, “self-modifying systems in Biology and Cognitive Science”, Oxford, Pergamon Press, 1991.

[KAMPIS 91b] G. KAMPIS, “Emergent computations, life and cognition”, in World Futures, Vol.31, Gordon and Breach Science Publishers S.A. 1991, pp. 33-48.

[LACROUX 96] F. LACROUX, “Contribution à une théorie de la planification adaptative : La stratégie procédurale”, thèse de doctorat, Université d’Aix-Marseille III, 1996.

[LABBANI et al 96] O. LABBANI, J.P. MULLER, A. BOURJAULT, “Describing collective behaviours” in Workshop ICMAS, 09-13 Dec., Kyoto, Japan, 1996.

[LE STRUGEON et al 96] E. LE STRUGEON, G. AGIMONT, R. MANDIAU, P. MILLOT, “Un exemple d’organisation évolutive parmi les robots fourrageurs”, journées françaises sur les systèmes multi-agents, JFIADSMA, Avril 1996, Sète, France, pages 167-176.

[LANGTON 90a] C.G. LANGTON, “Computation at the edge of chaos: phase transitions and emergent computation”, in “emergent computation”, MIT Press, Cambridge, MA, 1990, pp 12-37.

[LANGTON 90b] C.G. LANGTON, “Introduction to Artificial Life II”. Artificial Life II, Proceedings of the Workshop on Artificial Life Held Feb., 1990 in Santa Fe, New Mexico, Ed. C. Langton, C.Taylor, J. Doyne Farmer, S. Rasmussen, Addison-Wesley Publishing Company, 1990.

[LENAY 96] C. LENAY, “Coopération et intentionalité”, journées françaises sur les systèmes multi-agents, JFIADSMA, Avril 1996, Sète, France, Pages 265-272.

[LIU et SYCARA 93] J. LIU, K. SYCARA, “Emergent Constraint Satisfaction Through Multi-Agent Coordinated Interaction”, in “from Reaction to Cognition”, : 5th European Workshop on Modelling an Agent in a Multi-Agent World - MAAMAW'93, C. Castelfranchi and J-P. Müller eds, Springer Verlag, 1996.

[MAGNIN 96] L. MAGNIN, “Modélisation et simulation de l’environnement dans les systèmes multi-agents. Application aux robots footballeurs”, Thèse de l'université Paris VI, 1996.

[MARCELPOIL et al 94] R. MARCELPOIL, E. BEAUREPAIRE, S. PESTY, “La sociologie cellulaire : modéliser et simuler une “société cellulaire” pour étudier le vivant”. Intellectica, pp. 53-72, n°19, 1994/2.

[MARCENAC 96] P. MARCENAC, “ Emergence of Behaviors in Natural Phenomena Agent-Simulation”, Complex Systems 96, From local interactions to global phenomena, IOS-Press, R. Stocker, H. Jelinek, B. Durnota & T. Bossomaier eds, Albury, Australie, 1996, pp 284-289.

[MARCIA 96] GROUPE MARCIA, “Auto-organisation:=évolution de structure(s)”, Journée thématique du PRC-IA sur les systèmes multi-agents, Toulouse, Février 1996, Pages 139-152.

[MARTINET 90] A-C MARTINET (Ed) , “Epistémologies et Sciences de Gestion”, Economica, 1990.

[MARTINET 93] A-C MARTINET, “Les paradigmes tratégiques : l’éternel retour ?”, papier de Recherche n°28, 1993.

[MATARIC 92a] M. MATARIC, “Minimizing complexity in controlling a mobile robot population”, in Proc. of the 1992 IEEE Int. Conf. on Robotics and Automation, Nice, France, May 1992.

[MATARIC 92b] M.J. MATARIC, “Designing Emergent Behaviors: From Local Interactions to Collective Intelligence”. From Animals to Animats 2 (SAB'92), MIT Press, 1992.

[MEMMI 96] D. MEMMI, “Emergence et niveaux d'explication”, in Journées thématiques de l'ARC (émergence et explication), 1996.

[MEYER et GUILLOT 91] J.A. MEYER, A. GUILLOT, Simulation of Adaptive Behavior in Animats: Review and Prospect. From Animals to Animats (SAB'91); pp.2-14, Ed. J.E. Meyer, S.W. Wilson, MIT Press, 1991.

[MITCHELL et HOFSTADTER 90] M. MITCHELL, D.R. HOFSTADTER, "The emergence of understanding in a computer model of concepts and analogy-making", in "emergent computation", MIT Press, Cambridge, MA, 1990, pp 322-334.

[MÜLLER et PECCHIARI 96a] J.P. MÜLLER, P. PECCHIARI, "Un modèle de systèmes d'agents autonomes situés : application à la déduction automatique", dans AI distribuée et systèmes multi-agents, JFIADSMA'96, coordinated by J.P. Müller and J. Quinqueton, Ed. Hermes, 1996.

[MÜLLER et PECCHIARI 96b] J.P. MÜLLER, P. PECCHIARI, "Un modèle de systèmes d'agents autonomes situés : application to automated deduction", dans Proceedings ICMAS'96, Kyoto, décembre 1996.

[MINTZBERG 81] H. MINTZBERG, " Structure et dynamique des organisations ", Les Editions d'Organisation, 1981.

[MINSKY 88] MARVIN MINSKY - La société de l'esprit - Ed. Intereditions - 1988

[NAGAO et al 93] K. NAGAO, K. HASIDA, T. MIYATA, "Emergent Planning: A Computational Architecture for Situated Behaviour", in "from Reaction to Cognition", : 5th European Workshop on Modelling an Agent in a Multi-Agent World - MAAMAW'93, C. Castelfranchi and J-P. Müller eds, Springer Verlag, 1996.

[NAGUEL 61] E. NAGEL, "The structure of Sciences", Ed. Harcourt Brace World, New York, 1961.

[PIQUEMAL et GLIZE 96] C. PIQUEMAL-BALUARD, P. GLIZE, "Des aptitudes non cognitivistes d'agents pour l'auto-organisation", Journée thématique du PRC-IA sur les systèmes multi-agents, Toulouse, Février 1996, pp 129-138.

[SCAGLIONE 96] M. SCAGLIONE, "L'intentionalité dans les systèmes artificiels. Emergence et réalisation: deux côtés de la même pièce", Thèse de l'Université de Neuchâtel, Suisse, 1996

[SEARLE 95] J.R. SEARLE, La redécouverte de l'esprit, Gallimard (éd.), 1995.

[STEELS 90] L. STEELS, Cooperation between distributed agents through self-organisation, in Decentralized A.I., Ed. by Y. Demazeau and J.P. Muller, Elsevier Science Publishers B.V. (North-Holland), 1990.

[STEELS 91] L.STEELS, Towards a Theory of Emergent Functionality. From Animals to Animats (SAB'91), Ed. J.E. Meyer, S.W. Wilson, MIT Press, 1991, pp. 451-461.

[TAYLOR 90] C.E. TAYLOR, "Fleshing Out" Artificial Life II. Artificial Life II; pp. 25-38, Ed. C. Langton, C.Taylor, J. Doyne Farmer, S. Rasmussen, Addison-Wesley Publishing Company, 1990.

[TAYLOR 91] C. TAYLOR, "Fleshing out Artificial Life II", in Artificial Life II, C. Langton, C. taylor, D. Farmer, & S. Rasmussen eds, Addison Wesley, 1991.

[TELLER 92] P. TELLER, "A contemporary look at Emergence", dans A. Beckermann, H. Flohr et J. Kim, "Emergence or reduction?" Walter de Gruyter 1992, pp. 139-153.

[VISETTI 96] Y.M. VISETTI, "Emergence et hétérogéniété : le problème des passages", in Journées thématiques de l'ARC (émergence et explication), 1996.

[VON BERTALANFFY 93] LUDWIG VON BERTALANFFY - Théorie générale des systèmes - Editions Dunod - 1993

[WILSON 90] S.W. WILSON, "Perceptron redux: emergence of structure", in "emergent computation", MIT Press, Cambridge, MA, 1990, pp 249-256.

Annex III
Publications 2002 related to BAP

Balacheff N., Gaudin N. (2002) Students conceptions: an introduction to a formal characterization, *Les Cahiers du Laboratoire Leibniz* n°65. **http://www-leibniz.imag.fr/LesCahiers/Cahiers2002.html**.

Gaudin N. (2002), Conceptions de Fonction et Registre de Représentation: Etude de Cas au Lycée, *For the learning of Mathematics* 22, 2, pp. 35-47.

Pesty S., Webber C., Balacheff N. (2002) Baghera : une architecture multi-agents pour l'apprentissage humain. (à paraître) *Revue Cognitique.*

Webber C., Pesty S. (2002) Emergent diagnosis via coalition formation. In: *IBERAMIA 2002 - Proceedings of the 8th Iberoamerican Conference on Artificial Intelligence*. Garijo, F. (ed.), LNAI 2527, Springer Verlag, Spain, November 2002. pp.755-764.

Webber C., Pesty S. (2002) Emergence de diagnostic par formation de coalitions - Application au diagnostic des conceptions d'un apprenant. In: *Journées Francophones pour l'Intelligence Artificielle Distribuée et les Systèmes Multi-Agents 2002.* J.P. Muller (ed), Hermes, Lille, 2002. pp.45-57.

Webber C., Pesty S. (2002) A two-level multi-agent architecture for a distance learning environment. In: *ITS 2002/Workshop on Architectures and Methodologies for Building Agent-based Learning Environments*, E. de Barros Costa(Ed.), Spain/France, 2002, pp.26-38.

Webber C., Pesty S., Balacheff N. (2002) A multi-agent and emergent approach to learner modelling. In: *ECAI 2002 - Proceedings of the 15th European Conference on Artificial Intelligence.* F. van Harmelen (ed.), IOS Press, Amsterdam.

Le laboratoire Leibniz est fortement pluridisciplinaire. Son activité scientifique couvre un large domaine qui comprend aussi bien des thèmes fondamentaux que des thèmes très liés aux applications, aussi bien en mathématiques qu'en informatique.

Les recherches sur les Environnements Informatiques d'Apprentissage Humain et la didactique des mathématiques ouvrent cette pluridisciplinarité sur les sciences humaines, elles jouent un rôle particulier en favorisant les coopérations entre différentes composantes du laboratoire.

* mathématiques discrètes et recherche opérationnelle
* logique et mathématique pour l'informatique
* informatique de la connaissance
* EIAH et didactique des mathématiques

Les cahiers du laboratoire Leibniz ont pour vocation la diffusion des rapports de recherche, des séminaires ou des projets de publication réalisés par des membres du laboratoire. Au-delà, Les cahiers peuvent accueillir des textes de chercheurs qui ne sont pas membres du laboratoire Leibniz mais qui travaillent sur des thèmes proches et ne disposent pas de tels supports de publication. Dans ce dernier cas, les textes proposés sont l'objet d'une évaluation par deux membres du Comité de Rédaction.

Contact Gestion & Réalisation : Jacky Coutin
Directeur de la publication : Nicolas Balacheff
ISSN : 1298-020X - © laboratoire Leibniz