In D. Donoval (ed), Proceedings of CAEE'95 : Int. Conference on Computer Assisted Engineering Education. Bratislava : Slovak Technical University, 218-225. 1995

Generation of relevant didactic explanations by the computer running a simulation for itself

Jean-louis Dessalles

E.N.S.T. - Département Informatique
46 rue Barrault - 75634 PARIS Cedex 13 - France
Tel.: (33 1) 45 81 75 29 - Fax: (33 1) 45 81 31 19
E-✆: dessalles@enst.fr

Conceptual knowledge is a fundamental part of what is taught to engineering students. However most efforts in C.A.L. research are devoted to helping students acquire new skills, not concepts. We describe here a research project that aims at providing the student with relevant conceptual explanations whenever these are needed. We try first to describe what a relevant explanation should be and how it could be generated. Then we consider the possibility of coupling the explanation module with a simulation program so that part of the knowledge used in explanations is extracted from the simulation.

keywords: ICAI, critiquing system, explanation, misconception

1. Introduction : the need for sound explanations

Learning involves many different aspects such as discovery, trials and errors, factual knowledge, extraction of recurrent patterns, understanding theoretical knowledge, and so on. During the three last decades, research in Computer Assisted Learning gave rise to several efficient devices that may assist the learner in all these tasks. Our concern here is more specifically about the possibility of providing the student with relevant explanations at the right moment. This is certainly a very important act in the teaching process. However, to our view, this fundamental aspect has not been sufficiently addressed in the recent past. Current trends in C.A.L. research are more concerned with giving access to pre-stored knowledge and multimedia data or with providing realistic simulations, rather than with the computation of tailored explanations. We consider that most C.A.L. systems could be dramatically improved by the adjunction of an efficient explanation module.

Some well-known systems such as J.R. Carbonell's SCHOLAR [[27]], the SOCRATIC SYSTEM [[2]], WHY [[24],[5]] were able to argue with the student and thus to point at misconceptions that may have occurred during the student's learning process. Unfortunately these efforts have not been systematically pursued. Among recent attempts, we find interesting studies about the development of contextual help systems [[19]] and the so-called critiquing systems [[15]]. However such studies are often technically oriented and do not rely on a systematic investigation of what a sound and efficient explanation should be. As a consequence, even if excellent realizations are produced, they appear as *ad hoc* solutions and provide us with no indication about a general method for generating efficient explanations. Moreover, such systems require a long development that is time consuming for the author. What we need is, ideally, a general-purpose explanation module that would be able to deliver relevant explanations at the right moment. Of course, this module would be given appropriate knowledge, the design of which should be quite light for the author if s/he is using powerful elicitation systems. We think that such a system is not an utopia. We even think that most of the necessary techniques are available. What we need is a better understanding of the students' conceptual needs, and a criterion to decide weather an explanation will be relevant or not.

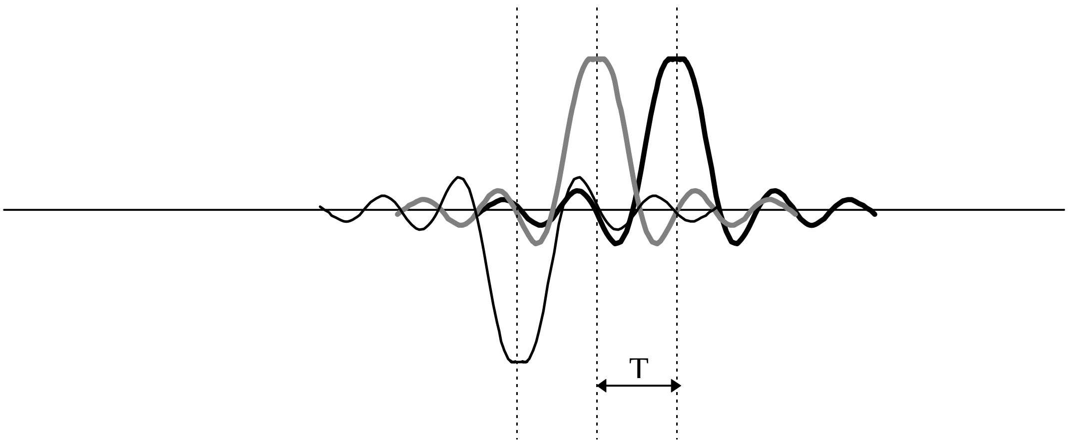
Surprisingly, most of the work on the generation of explanations has been conducted outside the C.A.L. community. The main concern was to make knowledge-based systems able to communicate not only results, but also the deductive reasoning that led to these results [[17],[20],[22]]. The trouble came from the fact that users were not at all satisfied when given the trace containing the mere sequence of deductions made by the system. Some authors attempted to perform some more elaborate treatment on the trace [[25]], while others decided to completely redesign knowledge-base systems in order to make them able to produce human-like explanations [[26],[3]]. Some of these developments were extended to C.A.L. applications [[4],[16]].

Such a transposition from the expert system technology to C.A.L. could be seen as problematic. It is understandable that the satisfaction of a knowledge-base system user requires sound explanations, but are we sure that sound explanations will systematically improve the performance of students in a learning context ? In [[23]], it was almost suggested that this was not true ! In [[11]], we gave arguments in favor of explanations. We made a qualitative distinction between two fundamentally different aspects of learning : skill learning and conceptual knowledge learning. We suggested that explanations were the very process by which concepts were acquired. In [[23]], explanations could be avoided (and were even superfluous) because it was a situation of pure skill acquisition. If we think of the importance of conceptual knowledge in itself but also as involved in complex skills [[11]] (try to imagine an ignorant physician or an ignorant engineer) we see the importance of being able to provide the learner with sound conceptual explanations.

2. Modeling relevant explanations

Our own approach contrasts with many research trends in C.A.L. in several aspects. Our first motivation was not technically oriented, but was rather to understand the structure of spontaneous human explanations [[6]]. Our starting point was the observation of natural conversations. It appeared to us that efficient explanations were quite often given or asked for during spontaneous interactions [[12]]. We also observed that this phenomenon occurred each time interlocutors considered a situation *paradoxical*. In other words, in spontaneous situations, explanations are not merely delivered. They always appear as a way to prevent or to escape from a paradoxical state of affairs. We tried to extend this principle to learning situations. Our system SAVANT3 [[7],[8]] was designed to help the student discover explanation through a dialogue. In a first phase, SAVANT3 tries to show that the situation created by the student is paradoxical, in other words, to show that the student made contradictory declarations or decisions. Then, in a second phase, the system and the student discuss together until they find a way of restoring consistency.

Besides these C.A.L. developments, we worked on a program, PARADISE, that was able to reconstruct real conversations [[10]]. The underlying principles are almost the same as with SAVANT3 : first detect problematicity, then escape from it. In the case of PARADISE, problematicity includes both paradoxical and undesirable states of affairs. We now consider PARADISE not only as computer program that is able to reproduce the interlocutors' performance, but as a plausible cognitive model of the generation of explanations. Let us consider an example.



sampling points

The above picture shows the situation encountered at the receiver in a digital communication. We see three successive pulses that overlap since *T* is smaller than the width of a single pulse. Actually, the real signal is not shown : it results from the sum of the successive pulses. The received signal is sampled at periodic moments in order to decide if the corresponding pulse is positive (meaning for instance `1`) or negative (meaning `0`). If the sampling is merely performed without specific care, then pulse overlap will corrupt the samples, and errors will ensue. However, if the signal is nicely filtered so that only the single interesting pulse is non zero at sampling time (as shown on the picture), then the message may be perfectly recovered.

What are we doing when we are teaching this aspect of digital communication ? What the student needs to "know" is certainly not only what to do in such and such circumstances. The student needs to *understand* what s/he is doing. The question, for us, is to correctly model the process by which the student understands, so that we are able to produce better explanations. According to the theory we developped, what the student has to acquire is a *set of constraints*. Here are a few examples of such constraints :

incompatibility 1 :
I.S.I. terms: p((k-n)T) are not equal to zero (for k and n different)
samples are not spoiled

incompatibility 2 :
pulses overlap at the receiver
I.S.I. terms: p((k-n)T) are equal to zero (for k and n different)
pulse shape is not modified at the receiver

incompatibility 3 :
samples are spoiled
the message is perfectly recovered

incompatibility 8 :
pulses are shorter than T
the bandwidth of the transmitted signal is not larger
the symbol rate (1/T) is not made slower

incompatibility 15 :
the bit rate is not lower
the symbol rate (1/T) is made slower
the size of the symbol set is not increased

For instance *incompatibility 2* states that it would be miraculous that I.S.I. terms (Inter-symbol Interference) are zero when nothing is done to the signal. If the student is fully aware of all the relevant constraints, then we can say he knows the subject. What can we do to help him ? The first idea, which is at the core of SAVANT3, is to consider that if the student lacks one of these constraints, then s/he will wrongly accept or create abnormal situations. Here is an example of dialogue with SAVANT3.

S- *According to what I understood: the message is perfectly recovered. But are we sure that the samples will give a reliable information ?*
L- *no*
S- *But perhaps you think that I.S.I. terms: p((k-n)T) are not equal to zero (for k and n different)*
L- *they may be equal to zero*
S- *yes... But pulses overlap at the receiver*
L- *yes*
S- *But wait... According to what you said:*
- pulses overlap at the receiver
- I.S.I. terms: p((k-n)T) are equal to zero (for k and n different)
I must say I'm confused !
L- *we can change the shape of the pulses as they are received*
S- *yes... According to what I understood: pulses overlap at the receiver. But perhaps you think that the symbol rate (1/T) is made slower*
L- *no, simply put a filter to adjust the shape*
S- *yes... Okay...*

What SAVANT3 does is to try to trap the student into a contradiction, using the 15 incompatibilities of the knowledge base. However, if SAVANT3 is able to help the student discover his misconceptions by himself through this type of Socratic dialogue, it is not, however, designed to provide complete explanations or advice. Let us now observe PARADISE's output with the same set of 15 constraints :

input

pulses overlap at the receiver
I.S.I. terms: p((k-n)T) are not equal to zero (for k and n different)

trace

constraint	decision	*comment*
12	• *pulses are not shorter than T*	understanding the situation (making the world consistent)
13	• *the symbol rate (1/T) is not made slower*	
1	• *samples are spoiled*	
3	• *the message is not perfectly recovered*	
4	• *the message is perfectly recovered*	changing the situation (making the world desirable)
3	• *samples are not spoiled*	
1	• *I.S.I. terms: p((k-n)T) are equal to zero (for k and n different)*	
2	• *pulse shape is modified at the receiver*	
10	• *pulses are filtered at the receiver*	

result

pulses overlap at the receiver
pulses are not shorter than T
the symbol rate (1/T) is not made slower
the message is perfectly recovered
samples are not spoiled
I.S.I. terms: p((k-n)T) are equal to zero (for k and n different)
pulse shape is modified at the receiver
pulses are filtered at the receiver

PARADISE is able to find out a solution. This is not a wonder, since many systems are able to satisfy a given set of constraints. What is interesting here is that PARADISE is supposed to do it "in the human way". In its modern version, PARADISE is a very simple program that makes only two things : it first makes the world consistent, then it makes the world both consistent and desirable (or at least not undesirable). Interestingly, the same technique is involved in the two phases. The trace obtained in the first phase, when PARADISE tries to make the situation consistent, looks very much as human understanding. It understands that the samples are spoiled and that consequently the message will not be perfectly recovered, using incompatibilities 1 and 3. Then, in the second phase, PARADISE tries to change this undesirable state of affairs. Doing this, it comes upon the filter solution.

What we claim here is that PARADISE is a good model of human reasoning that may have direct applications to the generation of didactic explanations. It is easy to use PARADISE's trace and output to generate sound explanations that may be illuminating for the learner. But explanations are useful when they are needed. How can we ensure that the learner will need them ?

3. Interfacing SAVANT3 and PARADISE with simulations

According to the underlying theory, explanations are needed when the student is perceiving a paradoxical situation, and advice or decisions are needed when the student is perceiving an undesirable situation. SAVANT3 tries to provoke the perception of a paradoxical situation by asking questions, and then tries to solve it with the student. PARADISE, on the other hand, has very little to say when the situation is neither paradoxical nor undesirable. It would be more natural to have both systems waiting for a suitable, *i.e.* problematic, situation. Actually such problematic situations do frequently occur when the student is using a simulation program, making choices and reacting to events. The challenge is then for us to be able to take advantage of these situations.

In a previous attempt [[13]], we tried to have SAVANT3 argue about the correctness and the efficiency of a small Prolog program written by the student. Any cause for imperfection was presented as incompatible with these two criterions. However this approach had intrinsic limitations. SAVANT3 had to be provided with a complete knowledge that it could use to prove the correctness and the efficiency of the Prolog program. Most of the time, this requirement of completeness cannot be fulfilled : the mapping *parameters* $\rightarrow$ *behavior* that is realized through a given simulation cannot be totally captured by a limited set of logical constraints. We are now trying to adopt a different approach which is reminiscent of the critiquing system approach [[15]] in this respect. First we abandon the idea of complete knowledge. Then we provide the critic, which may be SAVANT3 or PARADISE, with two kinds of interface with the simulation :

- a set of "sensors" : each sensor is attached to a logical proposition like *"pulses overlap"*. It is a way of asking the simulation for the truth value of the proposition at a given moment.
- a way of asking if a given state (partial assignation of truth values to propositions) is possible, *i.e.* can be reached by the simulation. If this is the case, the simulation may activate other sensors, *i.e.* assign a truth value to some uninstantiated propositions.

As a result, the critic is able to detect problematic situations when they occur. SAVANT3 is then able to start a fruitful dialogue with the student. Of course, due to incomplete knowledge, some problematic situations may go unnoticed, SAVANT3 being unable to detect them. In the same context, PARADISE performs almost identically. It detects a problematic situation and begins to think about it. However, it is somewhat smarter than SAVANT3, because it is able to draw some inferences. In order to make the situation consistent and desirable, PARADISE has to assign truth values and even to reverse some of them. Unfortunately, PARADISE can never be sure that what is consistent or desirable from its own point of view will remain consistent and desirable after a confrontation with the simulated reality. If pulses are filtered at the receiver, then everything seems to be okay in PARADISE's world. But filtering may be imperfectly realized, leading to non zero ISI terms, or to a worse signal-to-noise ratio. In such case, PARADISE will be able to go further by trying another way of making the world consistent and desirable, suggesting for instance to increase the set of symbols.

Such a procedure, by which PARADISE tries its solution before uttering it, seems to be very promising. We are currently working on several projects that aim at proving the feasability of the approach and its didactic value. Many technical and conceptual problems are yet to be solved, but we are very are very confident in the principles that underly the approach : the need for sound explanations that are delivered in problematic contexts, and the interface between logical contraints and simulations as a way to build up such explanations.

References

[1] Anderson John R. (1992). Interview, by J. Sandberg and Y. Barnard. *AI communications* vol. 5 n° 1, 1992, pp. 28-31

[2] Bolt Richard H. (1968). Computer-Assisted Socratic Instruction. In Orr W.D., *Conversational computer*, J. Wiley, New York 1968

[3] Clancey William J. (1981). *The Epistemology of A Rule-Based Expert System: A Framework for Explanation*. STAN-CS-81-896, Stanford University, 1981

[4] Clancey William J. (1987). Methodology for Building an Intelligent Tutoring System. In Kearsley Greg P., *Artificial Intelligence & Instruction - Applications and Methods*, Addison-Wesley Publishing Company, Menlo Park, USA 1987, pp. 193-228

[5] Collins Allan (1976). Processes in Acquiring Knowledge. In Anderson , Spiro , Montague, *Schooling and Acquisition of Knowledge*, Erlbaum Assoc., 1976

[6] Dessalles jean louis (1985). Stratégies naturelles d'acquisition des concepts et applications E.A.O.. In , *actes du colloque COGNITIVA 85*, CESTA, Paris 1985, pp. 713-719

[7] Dessalles jean louis (1990). Computer Assisted Concept Learning. In Norrie D.H., Six H.-W., *Lecture Notes in Computer Science 438 - Computer Assisted Learning*, Springer-Verlag, Berlin 1990, pp. 175-183

[8] Dessalles jean louis (1991). Conversation Assisted Learning: The SAVANT3 Dialog Module. In Forte Eddy N., *Proceedings of Calisce'91*, Presses Polytechniques et Universitaires Romandes, Lausanne 1991, pp. 159-165

[9] Dessalles jean louis (1991). SAVANT: How to Help Engineers Learn New Concepts. In Michel Jean, Pitra Zbynek, *Proceedings of CAEE'91*, Czech Technical University of Prague, 1991, pp. 202-207

[10] Dessalles jean louis (1990). The simulation of conversations. In Kohonen Teuvo, Fogelman-Soulié Françoise, *COGNITIVA 90 - Proceedings of the Third Cognitiva Symposium (Madrid)*, North Holland, Amsterdam 1991, pp. 483-492

[11] Dessalles jean louis, Rajman Martin (1992). *Concepts and Procedures in Engineering Education: Designing Specific Teaching Aids*. International Conference on Trans-European Cooperation in Engineering Education, Prague 1992, pp. 11-17

[12] Dessalles jean louis (1992). Model-Based Surprise and Explanation: a way to negotiate concepts. In Brezillon Patrick, *Proceedings of the ECAI-92 Workshop on Improving the Use of KBS with explanations,* Rapp. LAFORIA 92/21 Univ. Paris VI, Paris 1992, pp. 107-113

[13] Dessalles jean louis, Meyers pascal (1993). *Exemple d'une simulation argumentée pour l'apprentissage de Prolog*. Actes des Troisièmes Journées EIAO de Cachan, Eyrolles, Paris 1993

[14] Dessalles jean louis (1993). *From I.T.S. to I.C.S. : Learning with an Intelligent Critic, not with a Tutor*. Conference on Computer Assisted Engineering Education CAEE'93, Polytehnica University, Bucharest 1993

[15] Fischer Gerhard, Lemke Andreas C., Mastaglio T. (1991). Critics: An Emerging Approach to knowledge-based human-computer interaction. *International Journal of Man-machine Studies* vol. 35, 1991, pp. 695-721

[16] McCoy Kathleen F. (1988). Reasoning on a highlighted user model to respond to misconceptions. *Computational Linguisitics* vol. 14 n° 3, 1988, pp. 52-63

[17] Moore Johanna D. (1994). *Participating in Explanatory Dialogues*. The MIT Press, 0-262-13301-6 1994

[18] Ohlsson Stellan (1991). Interview, by J.Sandberg and Y.Barbard. *AI communications* vol. 4 n° 4, 1991, pp. 137-144

[19] Pachet François, Giroux Sylvain, Paquette G. (1994). Pluggable advisors as epiphyte systems. In Dessalles jean louis, *Proceedings of CALISCE'94 (Int. Conf. on Computer Aided Learning and Instruction in Science and Engineering),* TELECOM-Paris 94-S-002, Paris 1994, pp. 167-174

[20] Paris Cécile L. (1992). Systèmes Experts explicatifs. In , *Actes des deuxièmes journées Explication du PRC-GDR-IA du CNRS*, INRIA, Sophia-Antip. 1992, pp. 3-23

[21] Rätz Thomas, Lusti Markus (1992). Explanation Strategies: realization in a tutor for database normalization. In Brezillon Patrick, *Proceedings of the ECAI-92 Workshop on Improving the Use of KBS with explanations*, Rapp. LAFORIA 92/21 Univ. Paris VI, Paris 1992, pp. 47-56

[22] Safar Brigitte, Berthault Pascale, Sylvestre J. (1992). *Place des explications dans la conception d'une interface intelligente entre une base de données et un usager*. 12èmes Journées Internationales Avignon'92, Avignon 1992

[23] Sleeman Derek, Kelly A.E., Martinak R. (1989). Studies of Diagnosis and Remediation with High School Algebra Students. *Cognitive Science* vol. 13, 1989, pp. 551-568

[24] Stevens Albert, Collins Allan, Goldin Sarah E. (1979). Misconceptions in student's understanding. *International Journal of Man-machine Studies* vol. 11, 1979, pp. 145-156

[25] Swartout William R. (1983). XPLAIN: a System for Creating and Explaining Expert Consulting Programs. *Artificial Intelligence* vol. 21, 1983, pp. 285-325

[26] Wahlster Wolfgang (1981). *Naturlichsprachliche Argumentation in Dialogsystemen : KI-Verfahren zur Rekonstruktion und Erklarung approximativer Inferenzprozesse*. Springer Verlag, Berlin 1981

[27] Wenger Etienne (1987). *Artificial Intelligence and Tutoring Systems*. Morgan Kaufmann Publishers, INC., Los Altos, Cal.,USA 1987