A case study of collaborative modelling: building qualitative models in ecology

Paulo Salles

Universidade de Brasilia, Instituto de Ciências Biológicas

Campus Darcy Ribeiro, Brasilia - DF, 70.910-900, Brasil

E-mail: paulo.bretas@uol.com.br

Bert Bredeweg

University of Amsterdam, Department of Social Science Informatics

Roetersstraat 15, 1018 WB Amsterdam, The Netherlands

E-mail: bert@swi.psy.uva.nl

Keywords

Collaborative Model Building, Qualitative Modelling/Reasoning, Distance Learning, Ecology

Publication details: (in book) Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies. U. Hoppe, F. Verdejo and J. Kay (eds), pages 245-252, 2003, IOS-Press/Ohmsha, Japan, Osaka.

A Case Study in Collaborative Modelling: Building Qualitative Models in Ecology

Paulo SALLES[1] and Bert BREDEWEG[2]
[1]*Universidade de Brasilia, Instituto de Ciências Biológicas, Campus Darcy Ribeiro, Brasilia - DF, 70.910-900, Brazil*
[2]*University of Amsterdam, Department of Social Science Informatics, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands*

Abstract. To investigate the difficulties students have during the process of building qualitative models we describe a study in which artificial intelligence undergraduate students from the University of Amsterdam and graduate ecology students from the University of Brasilia were engaged in a collaborative model building activity. Their objective was to build models about the carbon cycle and the greenhouse effect using a process-based approach and possibly to implement the models in the qualitative simulator GARP. Almost all the students (94%) reported an increase in their understanding of the ecological problems after the modelling activity. Two types of difficulties were identified: some problems were related to the complexity of the ecological system being modelled. Other problems were related to the modelling language, particularly with the design of model fragments. Collaboration was considered an added value for the modelling success. Identifying causal relations, processes and describing the system's behaviour were easier for ecologists. The results stimulate us to carry on with research and practice on collaborative modelling.

Introduction

Model building is becoming an important educational activity. According to Forbus *et al.* [6] it is important that students become modellers because during the modelling process they have to articulate relationships between entities and dependencies between their beliefs. This is important for both understanding the phenomenon being modelled and for developing a broader understanding of complex, interrelated systems. This way, models provide means to externalise thoughts and to support questioning, discussion and justification of decisions. Finally, modelling provides students with practice in using formal representations; a skill needed for mastering mathematics and programming.

Qualitative Reasoning (QR) has a role in introducing modelling into the classroom and a new generation of QR related tools are being developed for that purpose: *Mobum* [1], *VisiGarp* [2], *Vmodel* [6], and *Betty's Brain* [8]. As QR-related learning environments become available and more people start building qualitative models, it becomes important to actively develop tools to support such model building activities. Our research questions are therefore: what are the most difficult concepts for the students to include in a qualitative model? Is there any added value in collaboration of students with different background? To carry out an exploratory study on these topics, a collaborative modelling effort was conducted with students from the University of Amsterdam (UvA) and the University of Brasilia (UnB). They worked on a pencil and paper basis and at the end of the course some of the students actually implemented their models in GARP [3]. In order to evaluate the modelling activity the students answered a questionnaire consisting of 41 questions, including personal characterisation, course evaluation, the modelling effort, and an evaluation of the collaboration.

1. Educational Context and the Domain Knowledge

The exploratory study described here involved 10 undergraduate and MSc students enrolled in the discipline 'Model Based Reasoning' (MBR) at the UvA and six MSc and PhD students enrolled in the discipline 'Models in Ecology' at the UnB in 2001. Due to their different backgrounds, the approach each group took to qualitative modelling was somewhat different. UvA students took it in terms of an artificial intelligence curriculum and UnB students took the modelling effort in the context of an ecology curriculum. UvA students had a good introduction to the QR literature in their MBR course, whereas UnB students had only basic knowledge on that topic. The lecturers also prepared a tutorial on GARP, particularly focussing on the notion of model fragments and qualitative behaviour graphs. Five groups were formed, each consisting of two Dutch and one Brazilian student. The overall modelling problem was divided into sub-problems and each group had to tackle a specific sub-problem. In order to facilitate the interaction and communication the e-group facility offered by Yahoo! was used. Thus, participants could communicate with each other using regular email as well as other Computer Supported Communication (CSC) tools provided by the Yahoo! e-group facility.

Nutrient cycling is an important area in ecology. The carbon cycle is particularly relevant because it includes a very broad set of phenomena and involves the interaction of biological, physical and chemical processes related to the production and use of organic matter. One of the most interesting aspects of this cycle and a big issue nowadays is the fact that compounds of carbon, specially the carbon dioxide (CO_2), retains heat and therefore affects the climate – the greenhouse effect. One of the tasks for the students was to acquire the knowledge relevant to this domain (see Figure 1).

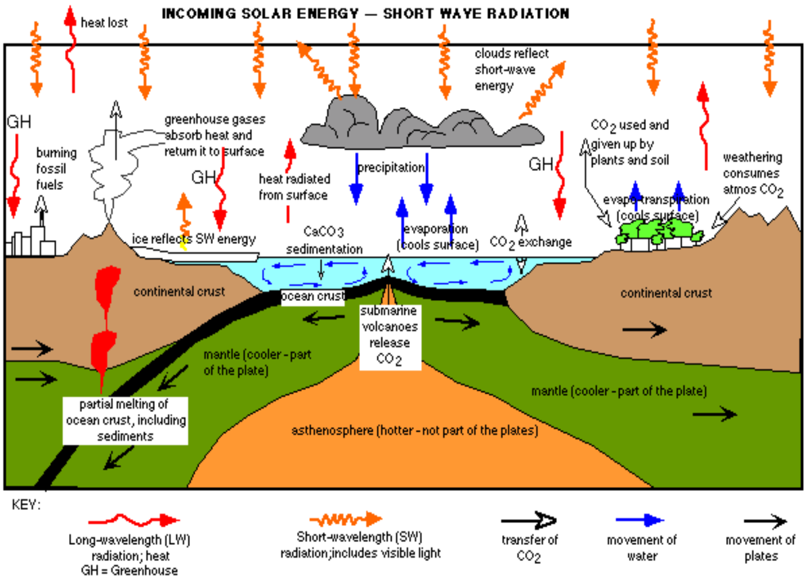

Figure 1: One of the many pictures found by the students on the WWW illustrating the problem situation.

2. Building Models with Qualitative Reasoning

We adopted the Qualitative Process Theory (QPT) [5] as the basic ontology for building the models and the domain independent qualitative simulation engine GARP [3] to implement the models. According to QPT the world consists of objects that interact among themselves and changes can only occur when a process is active. Characteristics of objects are represented by *quantities*, and the qualitative state of a quantity is a pair <*amount,derivative*>. Changes in their values mean changes on qualitative states, and therefore change in behaviour. Each quantity is associated with a set including the most relevant qualitative values, its *quantity space*. Following the compositional modelling approach [4] qualitative models are built from a set of *model fragments*, stand-alone parts that combine to form simulation models. There are two kinds of model fragments, *views* and *processes*. The former represents characteristic aspects of the system and typical situations. The latter describes mechanisms of change. Both types of model fragments include descriptions of objects, their quantities, quantity spaces and statements about inequalities involving quantities, which can be used to determine whether or not a situation holds or a process is active. Objects and model fragments can be organised hierarchically so that properties can be inherited (*subtype-hierarchies*). Some modelling primitives describe the relations between quantities. The most important are *influences* and *qualitative proportionalities*. Both express causality and have mathematical meaning. The former indicate what happens when a process is active: if, for example, *I+(A,B)* then *B* is the process rate and should be added to the derivative of *A*, and if *I–(A,B)* then the rate *B* is to be subtracted from the derivative of *A*. The effects of processes are propagated to other quantities through indirect influences modelled as proportionalities: if *P+(A,B)*, when *B* changes (e.g. increases), then it causes *A* to change in the same direction (that is, increase), and if *P–(A,B)*, when *B* changes (e.g. increases), then it causes *A* to change in the opposite direction (that is, decrease).

Simulation models in GARP consist of: (a) an initial scenario with a description of which objects are present, how they are structurally related, and initial values of relevant quantities (the system structure); (b) a library of model fragments containing knowledge about the system (a domain theory); and (c) a set of domain-independent rules governing state transitions. Simulations produce a *behaviour graph* with all the possible sequences of qualitatively different states the system can assume, given the conditions set in the initial scenario. This way, it is possible to infer the behaviour from a description of the structure of the system. GARP has been used in a number of educational studies (*e.g.*, [9] and [10]).

3. The Model Building Methodology

Starting the collaboration. Initially the students should get to know each other, particularly to 'meet' the colleagues from the other university. Each student therefore had to perform the following tasks: register at Yahoo! and join the e-group (mbr-ecology), send an introductory message to all the e-group members, presenting themselves and submit at least two bookmarks to the (mbr-ecology) e-group concerning the carbon cycle and/or the greenhouse effect. These bookmarks would give the group an initial overall understanding of the problem. In addition, and mainly to provide focus, the teachers gave the students a technical paper discussing some of the most important aspects of the 'carbon cycle' [7].

Forming teams and assigning subsystems. Based on information provided by the students, teams were formed to tackle the sub-problems that constitute the overall problem. Notice that this step already enforced students to decompose the main problem into a set of sub-problems, even though their knowledge on the domain was limited at this point. In order to prevent a potential deadlock the lecturers deliberately intervened both concerning the problem decomposition and the forming of teams. After all, this step was an important one and needed to be solved in order to progress with the main model building activities. Five themes were identified and each group started working on one of them: (a) a global model about the carbon cycle; (b) forests; (c) water and oceans; (d) human activities including industries, transport, agriculture; and (e) the greenhouse effect.

Domain related publications. In order to learn about their specific part of the overall problem the teams had to study the domain related material (WWW pages, including some online articles) and produce a written document (4 pages) discussing their part. As all documents, this document had to be submitted to the (mbr-ecology) e-group at Yahoo! so that all teams could read about the knowledge acquired by the group as a whole. Notice that each team consisted of students from both universities. During the seminars (locally at each University) group members had to present and discuss their ideas with the members from the other teams. This activity had two goals: first, to share insights among teams and second, to avoid too much diversity between the groups (part of the discussion focussed on the relations between the sub-problems). Students were also instructed to download the qualitative simulation software[1], inspect GARP's syntax for building model fragments and run simulations with the models already implemented.

Structural model and global behaviours. To arrive at a qualitative model of the system under study, the model building activity continued with a relatively strong focus on the knowledge representation underlying the qualitative simulator that we intended to use. The idea was to divide that goal into four steps (see also below). The first step consisted of three sub-activities: (a) structural model of the system. It is basically a concept map, including all the objects relevant to the problem organised in a subtype hierarchy. In addition definitions of structural relations, such as is-a, part-of, contains, etc, between those entities as far as needed; (b) global description of behaviour (processes). Textual oriented descriptions of typical behaviours, in fact processes (e.g. respiration) or agent models producing certain behaviour (e.g. farming); (c) scenarios and related behaviour graphs. Students had to develop two scenarios using the previously defined objects and behaviour descriptions and show how those lead to a particular behaviour graph relevant for understanding the issues concerning the problem of global heating.

Detailed behaviour model. The second step is to further detail the behavioural aspects. This consisted of three sub-activities: (a) define quantities; (b) define quantity spaces for each quantity and point out important landmarks; (c) construct an influence diagram (a 'causal model'). Using the previously defined quantities specify how they are causally related, mainly using the notions of influences and proportionalities (see Figure 2). Students also had to further detail the 'global behaviours' defined during the previous step. The idea being that this should lead to a first global description (using text) of the model fragments that will be part of the final model.

Detailed specification of model fragments. The third modelling step consisted of constructing detailed descriptions of the required model fragments. Although this step was performed using paper and pencil the students had to follow a specific syntax (provided by the teachers). Students who were more experienced with PROLOG were encouraged to formulate their model fragments directly in GARP, however, without running the simulator. The focus of this step was to conceptually clarify the set of model fragments (and not to focus on the overall effect of those model fragments on the behaviour graph potentially generated by the simulator).

Towards detailed implementation and running a simulation. During the fourth step the idea was to actually run models using the simulator. This mainly included the following sub-activities: (a) defining and implementing the possible scenarios (input systems); (b) analysing behaviour graphs generated by the simulator; (c) debugging and finalising the library of model fragments.

Writing documentation. Finally students had to write a report discussing their model. The goal was not to simple copy and paste the documents produced before, but to examine

[1] For software details see: http://web.swi.psy.uva.nl/projects/GARP/

them and reformulate these documents following all the discoveries and modifications made during the model building activities. In other words, a description of the latest results had to be produced by each team.

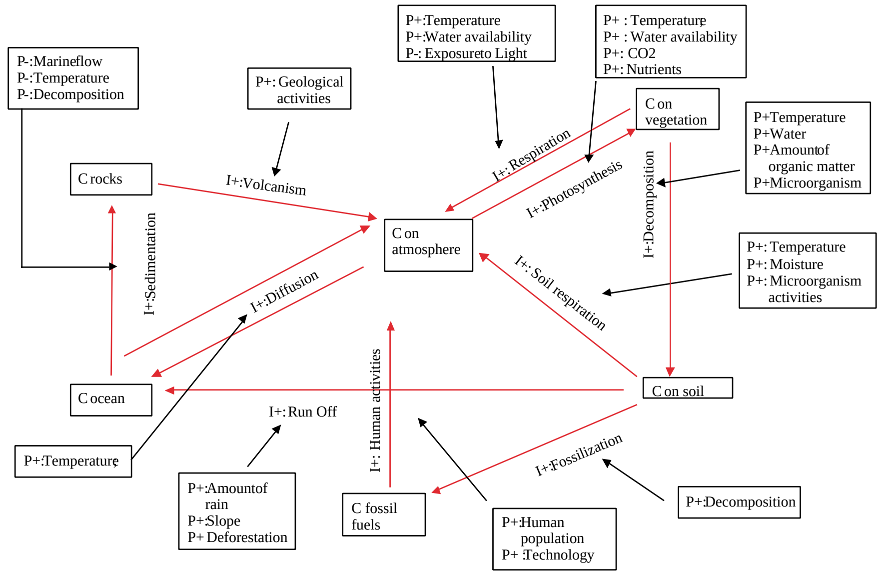

Figure 2: An example of an influence diagram constructed by students (stressing the fluxes).

4. Evaluating the Modelling Effort

The modelling effort was surveyed by means of a questionnaire with 41 topics presented to the students at the end of the activities. We used a five-point scale for the answers corresponding to {easy; ± easy; medium; ± difficult; difficult}. The answers given by the students are represented as follows: UvA: [0:2:5:1:2] = zero students from UvA answered ‘easy’, two students answered ‘± easy’; five answered ‘medium’, one answered ‘± difficult’ and two answered ‘difficult’ (see also Table 1).

4.1 Building Qualitative Models

‘Building model fragments’ was the most difficult task both for 60% of UvA and 50% of UnB students. ‘Understanding the overall problem’ was also mentioned as difficult. Identifying ‘physical objects’, ‘typical situations and scenarios’, and ‘building causal models’ were not mentioned by any student as the most difficult part. The easiest parts for UvA students were ‘constructing subtype-hierarchies’ and ‘identifying physical objects involved in the system’, and for UnB students it was ‘identifying the processes’.

4.2 The Effects of the Model Building Activity

In order to evaluate the effects of the modelling effort both on the understanding of ecological problems and of qualitative models, we asked the students specific questions to be answered with the options {beginner; ± beginner; medium; ± expert; expert}. The results are the following: (a) knowledge of these ecological systems BEFORE model building: UvA:[4:5:1:0:0] / UnB:[0:3:1:2:0]; (b) knowledge of these ecological systems AFTER model building: UvA:[0:3:3:4:0] / UnB:[0:0:2:3:1]; (c) knowledge of qualitative modelling BEFORE the course: UvA:[4:5:1:0:0] / UnB:[6:0:0:0:0]; (d) knowledge of qualitative modelling AFTER the course: UvA:[0:0:6:4:0] / UnB:[1:0:5:0:0]. The results show that the modelling effort increased their understanding of both the ecological problem and the qualitative models. The 'hands on' approach was apparently a good choice: *"building models is a good way to learn"*, said one of the students. Anyway, we insisted and presented them two more statements. Most of the students agree with the statement: 'building a qualitative model made me understand better the problem'. A UvA student said *"it made me able to abstract the problem"*. Another one said *"I fully agree, it gave me insight in the causality and the different perspectives from which we can look at the problem."* The students also agreed with the idea that 'it would be easier for me to build qualitative models about a similar problem (e.g. nitrogen cycle)'. A UnB student explained: *"I agree because of a better understanding of the processes."*

Table 1: Overview of the answers given by students.

Questions	Answers	Comments
In the beginning understanding the overall problem	UvA: [0:2:5:1:2] UnB: [0:0:1:4:1]	In the beginning, nobody found the problem to be modelled easy to understand, and the majority of answers were between 'medium' and 'difficult'. One of the students said: "*the start was difficult. After that when you finally figured out what is needed, it was easy*".
Identifying the most relevant aspects of the problem	UvA: [0:2:4:3:1] UnB: [0:0:1:4:1]	Also, identifying relevant aspects was not easy for anybody and the answers were also between 'medium' and 'difficult'. UvA students referred to consulting their Brazilian partners via email "*to have a confirmation that we had chosen the right aspects*".
Identifying typical situations of the system	UvA: [0:0:7:1:2] UnB: [0:0:5:1:0]	Recognising typical scenarios was a bit more difficult for UvA students than for UnB students. They said it was *"difficult to identify at an early stage"* but it became "*easier as a result of identifying the most relevant aspects of the problem*".
Drawing state-graphs of the system's behaviour	UvA: [0:0:4:4:2] UnB: [0:2:3:1:0]	Representing the system's behaviour was not easy. However, it was more difficult for UvA students "*mostly because of lack of ecological / chemical fore knowledge*". A student mentioned the fact that the whole problem was "*fragmented over different groups*".
Identify the objects involved in the system	UvA: [0:4:4:2:0] UnB: [0:3:2:1:0]	Answers in the two groups had similar distribution around 'medium'. One of the AI students of UvA said "*the difference between objects in GARP / PROLOG and a OOP language as Java make some conceptions hard.*"
Identifying quantities and define quantity spaces	UvA: [0:2:3:2:3] UnB: [0:3:1:2:0]	None of the students found it easy to identify quantities and to define quantity spaces, and it was a bit more difficult for UvA students. One of them noticed that *"more perception and interpretation are needed"*. Some tried to explain: "*there are too many quantities, it is difficult to choose the relevant ones*". Another pointed out that it is "*difficult to identify quantity spaces, it is hard to imagine. Would there be a maximum or not? Is zero an option?*"
Building is-a hierarchies	UvA: [0:5:3:1:1] UnB: [1:2:2:0:1]	Similar distribution of answers was observed in the two groups. Some UvA students mentioned their previous experience in doing this type of knowledge representation, while these were new concepts for the UnB students.
Identifying processes	UvA: [1:1:5:2:1] UnB: [4:1:1:0:0]	Identifying processes was more difficult for UvA students than for UnB students. One of the former group members said it was "*difficult, because many processes also occur in other subsystems*".
Building causal models, including influences and proportionalities	UvA: [0:1:4:4:1] UnB: [2:3:1:0:0]	*"This is the point were you need the domain knowledge and a lot of modelling knowledge"*, said one of the UvA students. Another student commented that *"especially problematic was the distinction between indirect / direct influences"*. A third one said it was difficult to say if the influences "*are going on all the time."*
Defining model fragments	UvA: [1:0:1:4:4] UnB: [0:1:3:2:0]	Creating model fragments is one of the most difficult parts of the modelling effort, as shown by the answers. *"There you need to have a great understanding of GARP"*, admitted one of UvA students. It is "*difficult to translate the causal relations into a model fragment, because this has certain limitations on the representation you can use*" said another student. UnB students found it easier because they defined the model fragments in natural language while their colleagues from UvA used PROLOG.

4.3 About the Collaboration

The students found it important to build models collaboratively: *"it is relevant, you can learn from each other the skills that you lack"*. However they said it is not necessary to have *international* collaboration. Still, 75% of the total of the students think that the lectures should continue with the international collaboration and 25% had no opinion. As one UvA student said, "*it was nice and it has never been done before*".

We asked the students to evaluate the international collaboration using the scale {not effective at all, not effective, medium, effective, very effective}. The answers were distributed as follows: UvA students, [3:2:2:2:1]; UnB students, [0:0:2:3:1]. Why did ecologists find the interaction 'effective' while their partners did not? The students had an uneven collaborative modelling experience. Out of the five groups, the students of one group had a strong interaction, and eventually implemented part of their models in GARP; students of three groups had some interaction, but this was not regular; and the students of one group had no interaction at all. From the questionnaires, explanations can be drawn on different aspects: (a) UvA students started the modelling effort earlier than their colleagues in Brazil. Therefore, they were not at the same stage of the modelling process and for some groups it was difficult to catch up; (b) there were problems with the language. The Brazilian students had more difficulties with written English, and messages did not flow smoothly and quickly; (c) moreover, even students who did not interact, could finish their work separately in spite of their different background. Given that e-mails exchanged in the e-group went to all the members, there was some sort of feedback for everybody; (d) the teachers could not anticipate all the problems, and there was some delay in delivering didactic material.

The students evaluated the performance of the Yahoo! e-group using the scale {very well, well, medium, bad, very bad}. The distribution of the answers were UvA:[0:4:2:4:0] and UnB:[5:1:0:0:0]. The students complained about the interface and about problems they had downloading files.

5. Discussion

This paper describes an exploratory study aimed at identifying the difficulties students have with qualitative modelling and to investigate the added value of collaborative model building activities, involving students with different background from different universities via the internet.

During the modelling process, the students had two kinds of problems: domain-specific and modelling language related problems. Domain specific problems were related to the complexity of the carbon cycle and the greenhouse effects. Understanding the overall problem and selecting the most relevant aspects to include in the model in the beginning were mentioned both by UvA and UnB students. If identifying objects was not pointed out among the most difficult tasks, assigning them quantities and selecting values with some special meaning to include in the quantity spaces was considered one of the most problematic aspects of model construction. In fact, ecological systems are complex and often there is a large set of variables to be considered. Selecting some to include in the model requires experience and clearly defined objectives, things that were missing to our students. Also, compared to physics, few ecological quantities have values with a physical representation, such as for example the freezing and boiling temperatures of water. That's why the students complained that *"it is hard to say when something is normal, high or maximum"*; "*Would there be a maximum or not? Is zero an option?*" Other ecology related problems include time scale and spatial reasoning. The students commented that in the carbon cycle there were some processes that happen all the time and others only in particular moments. Also, some processes occur in all the parts of the system and others occur only in some parts. Among the modelling language related problems the students mentioned difficulties in expressing causality. It was sometimes hard to recognise direct influences caused by processes and the propagation of their effects, via qualitative proportionalities. Above all, students elected model fragment creation as the most difficult task. This is understandable because these partial models encode the core of the knowledge being modelled: objects, conditions for things to happen, typical situations and processes.

As the students have different backgrounds, collaboration could be profitable for all. For example, it was easier for ecology students to identify typical situations, processes and to build causal models, while for AI students it was easier identifying objects and organising them in subtype-hierarchies. Collaborative model building can be seen as a kind of 'learning by teaching' activity. As a UvA student said, "*maybe it is a good idea if the ecologists put more emphasis on explaining the modellers (=AI students) their domain knowledge. Then the modellers could spend more time on making the models and hierarchies, and maybe explain things about that to the ecologists*". Collaboration was considered to be important by the group. We feel encouraged to carry on with this practice and with investigations on how to do that.

Acknowledgements

We would like to thank the UvA and the UnB students for their collaboration and valuable suggestions and observations.

References

[1] Bessa Machado, V & Bredeweg, B. (2001) Towards Interactive Tools for Constructing Articulate Simulations. *Proceedings of the International Workshop on Qualitative Reasoning*, QR'01, pages 98-104, San Antonio, Texas, USA, May 17-19.

[2] Bouwer, A. & Bredeweg, B. (2001) VisiGarp: Graphical Representation of Qualitative Simulation Models. In J.D. Moore, G. Luckhardt Redfield, and J.L. Johnson (eds.), *Artificial Intelligence in Education: AI-ED in the Wired and Wireless Future*, pp. 294-305, IOS-Press/Ohmsha, Osaka, Japan.

[3] Bredeweg, B. (1992) *Expertise in Qualitative Prediction of Behaviour*. PhD thesis. University of Amsterdam, Amsterdam, The Netherlands.

[4] Falkenhainer, B. & Forbus, K. (1991) Compositional Modeling: Finding the Right Model for the Job. *Artificial Intelligence*, 51(1-3): 95-143.

[5] Forbus, K.D. (1984) Qualitative process theory. *Artificial Intelligence*, volume 24, number 1-3, pages 85–168.

[6] Forbus, K.; Carney, K.; Harris, R. & Sherin, B.L. (2001) A qualitative modeling environment for middle-school students: A progress report. In Biswas, G. (ed.) *Proceedings of the Fifteenth International Workshop on Qualitative Reasoning*, St. Mary's University, San Antonio, TX.

[7] Grace, J. (2001) The Carbon Cycle. http://www.ierm.ed.ac.uk/ierm/teaching/ccycle/ University of Edinburgh, Scotland.

[8] Leelawong, K.; Wang, Y.; Biswas, G.; Vye, N. & Bransford, J. (2001) Qualitative Reasoning Techniques to support learning by teaching: The Teachable Agents Project. In Biswas, G. (ed.) *Proceedings of the Fifteenth International Workshop on Qualitative Reasoning*, St. Mary's University, San Antonio, TX.

[9] Salles, P. & Bredeweg, B. (1997) Building Qualitative Models in Ecology. In Ironi, L. (ed.) *Proceedings of the 11th. International Workshop on Qualitative Reasoning (QR'97)*. Instituto di Analisi Numerica C.N.R., Pubblicazioni no. 1036 , Pavia, Italy.

[10] Salles, P. & Bredeweg, B. (2001) Constructing Progressive Learning Routes through Qualitative Simulation Models in Ecology. In Biswas, G. (ed.) *Proceedings of the 15th. International Workshop on Qualitative Reasoning (QR'01)*. Saint Mary's University, San Antonio, TX, USA.