What you see is what you have in mind: constructing mental models for formatted text processing

Carlo Bellettini, Violetta Lonati, Dario Malchiodi, Mattia Monga, Anna Morpurgo, and Mauro Torelli

Dipartimento di Informatica
Università degli Studi di Milano
Milan, Italy
{bellettini,lonati,malchiodi,monga,morpurgo,torelli}@di.unimi.it

Abstract. In this paper we report on our experiments in teaching computer science concepts with a mix of tangible and abstract object manipulations. The goal we set ourselves was to let pupils discover the challenges one has to meet to automatically manipulate formatted text. We worked with a group of 25 secondary-school pupils (9-10th grade), and they were actually able to "invent" the concept of mark-up language. From this experiment we distilled a set of activities which will be replicated in other classes (6th grade) under the guidance of math teachers.

1 Motivation

As repeatedly pointed out in the computer science education literature [13, 5–7], the computing discipline is often confused with literacy in basic computer applications. This is particularly evident in the context of general (non-vocational) education; for instance, in Italy, computer science teaching from 6th to 10th grade is commonly limited to the use of office automation tools (like word-processors) or communication applications (web and email).

Nevertheless, computational thinking has the potential to be very formative for pupils, and several core aspects of computing are sufficiently basic to be taught as a fundamental subject [1, 10, 11, 15]. However, two main hurdles have to be considered in introducing fundamental computing concepts in K-12 education:

1. the *abstract nature* of computing, which makes it difficult to *illustrate* the concepts by referring to physical objects;
2. the *expectations* of students, who need clear links between learning about computing and using a computer as a productivity tool, activity to which they are exposed in their daily lives.

The basic use of computers can be exploited as a chance to explore computational concepts. For instance, even the basic use of a word-processor cannot be effective unless the user has acquired some sort of interpretation model of how the application tool is structured and how it works [4]. In fact, such models are usually deduced by a *bricolage* (trial-and-error) approach which often leads to misconceptions and ambiguities that prevent users from understanding the application correctly and using it effectively [3]. Teaching such models explicitly would also provide an introduction to some fundamental computer science concepts, for instance the digital representation of information or the need of unambiguous formal description of procedures that can be executed automatically.

We tried to implement such ideas starting from one of the most common computer activities in class (the use of a word processor), and designing a sequence of learning activities with the aim of making pupils aware of the challenges posed by the automatic processing of formatted text, mainly focusing on information representation issues.

A relevant part of this didactic sequence consists of a type of kinesthetic/tactile learning activities [2] which we have called *algomotricity*. These activities are then followed by more abstract ones implying the use of symbolic meta-languages. All the activities are designed to push pupils to experiment with the problem of text representation and explore suitable techniques to represent both text and formatting. In order to close the loop from known applications to abstractions, and back to computers, the sequence ends with the practice of such techniques again on a computer.

The approach is inspired by *constructivism* [4] and the *allosteric learning model* [9], which point out that the direct transmission of knowledge is rarely the best way for learning. On the contrary:

> "Learning is a highly active process which works in a conflicting way and in an integrative mode between what the learner has in his mind and what he can find and understand through his conceptions on his environment. When a learner elaborates a new model, all his mental model[s] must be reelaborated in an interaction between the conception and the external information. [...] Most of the time they need a didactic environment to cause interferences with their sitting conceptions. And this is the main function[~~s~~] of the teacher: he has to propose or to suggest an heuristic environment that may interfere with the learner conceptions." [9]

That is exactly what we meant in designing our activities, where pupils are put in new situations with a goal to reach and few or no hints about a-priori criteria or techniques to refer to.

Notice that the general goal of our didactic proposal is not to teach how to use the application but to let the pupils discover some of the principles its designer had in mind. In fact, this can be both useful to improve one's proficiency in using the application and—above all—to understand the challenges of digital information processing.

We started by experimenting with this approach within a group of 25 pupils in 9th and 10th grades of an Italian secondary school. In Section 2 we detail the activities we proposed to pupils, in Section 3 we report some remarks and considerations on how such an experiment was designed and implemented, and in Section 4 we draw our preliminary conclusions and try to prepare the ground for future work.

2 Description of the activities

The overall activity (8 hours in 4 non-consecutive days) was organized in four phases:

1. familiarization with text formatting on a word processor;
2. a *dramatization* of the process through the use of tangible objects;
3. a game designed to force pupils to restructure their mental models and discover the power of symbolic meta-languages;
4. a final use of special software tools for formatting texts, which could also show the data structure used to record the meta-information.

We started and ended off in a computer lab, in order to partially match pupils' expectations about informatics [12]. The risk we wanted to avoid was that of being perceived as unrealistic or fruitlessly "academic". Our goal was to let pupils be confronted with the challenges of dealing with formatted digital text, and we designed the activities around two main ideas: (a) computers and software tools should be of secondary importance, but the conceptual link with them should be clear; (b) the approach should be mostly *allosteric* [9]: the direct transmission of knowledge should be kept to a minimum, and pupils should be forced to reconsider their mental models of text formatting by discovering useful techniques by themselves. Besides, the abstract nature of computing should be conveyed by concrete examples and physical activities.

Pupils were asked to work in small groups to encourage this confrontation with the challenge. Moreover, almost every task was proposed together with an accompanying meta-cognitive reflection. In particular we often asked pupils to imagine how their descriptions would be understood by someone who was unaware of the context.

2.1 Basic text formatting

We started in a computer lab and the first goal was to introduce basic formatting elements such as italics, bold, etc. The pupils already had some knowledge about "formatting", but they had rarely reflected on its semantics. Thus, we discussed with the pupils *the role of formatting as a conveyor of information.*

The pupils (working in pairs) were requested to produce a formatted text of their choice. Then they had to answer some questions about their work: Which type of formatting did you use? Why did you choose it? Did you use more than one formatting for the same text passage? They first mainly noted the aesthetic value of formatting, but by reflecting on how they had used formatting in a text, they also noted that it plays an important role in transferring information and that indeed the *meaning* of a text is communicated by the words *together with* their formatting.

2.2 Recording of meta-information

The second phase was carried out in the gymnasium of the school. Retrospectively this was a bad choice, since the acoustic of the big hall was quite disastrous; moreover, the groups were too far apart to allow a fruitful discussion.

The proposed task was the reproduction of formatting on a hand-written text through the use of objects. Each group was given a printed formatted text and assigned an area of the gym with a copy of the text written on a big poster on the floor. This copy contained no formatting, even the alignment was slightly different with respect to the printed version. In order to mimic the formatted text, formatting had to be *codified* by using objects available in the gym. The pupils were free to use their own conventions, but they were requested to make the formatting understandable to others, thus every group of pupils (6 persons) was asked to write down the rules they used in the *codification* phase. Finally, each team moved to the poster of the next team, received the rules written by that team, and had to follow them and reproduce the original formatted text.

Most rules turned out to be ambiguous, especially when more than one kind of formatting had to be applied to the text. The objects were used mainly to mimic the formatting in the prototypical text: for instance the presence of a color in the prototypical text was represented by an object of the same color. However, in some cases (see for example the last word in Fig. 1 and the previous word, covered with plastic cups, for comparison), pupils discovered a more abstract symbolic use as a means to cope with a shortage of objects: in the example, two spoons were used to mark the beginning and the end of the word respectively, since there were not enough spoons to cover the whole word (see the previous word, covered with red beakers, for comparison).

2.3 Discovering meta-language tricks

The shortage of objects had let a group discover a smarter "symbolic" use, thus we proposed a game. The game was carried out in a classroom in teams of four pupils. Teams were again requested to reproduce a formatted text (see Fig. 2) with objects and write down codification rules precisely enough to be followed by another team. However, each object had a *cost* and the goal of the game was to produce the formatting with the minimal cost. In fact, the more an object could be used to mimic the appearance of a piece of formatting, the higher was its cost: *e.g.,* since spaghetti pasta could be easily associated to the *underline* style, its cost was very high. We aimed at promoting the *symbolic* use of objects. The winner would be the team able to hand in an unambiguous codification with the lowest cost.

As Fig. 3 shows, the introduction of cost led the pupils to discover what is actually frequently done in *mark-up languages*: the use of *tags* at the beginning and at the end of (possibly overlapping) regions. Also, the pupils chose to use the cheapest kinds of pasta to represent the more frequent formatting elements. Pupils also devised some unexpected solutions reinforcing such an approach: they found a way to give more than one function to the cheapest kind of pasta,

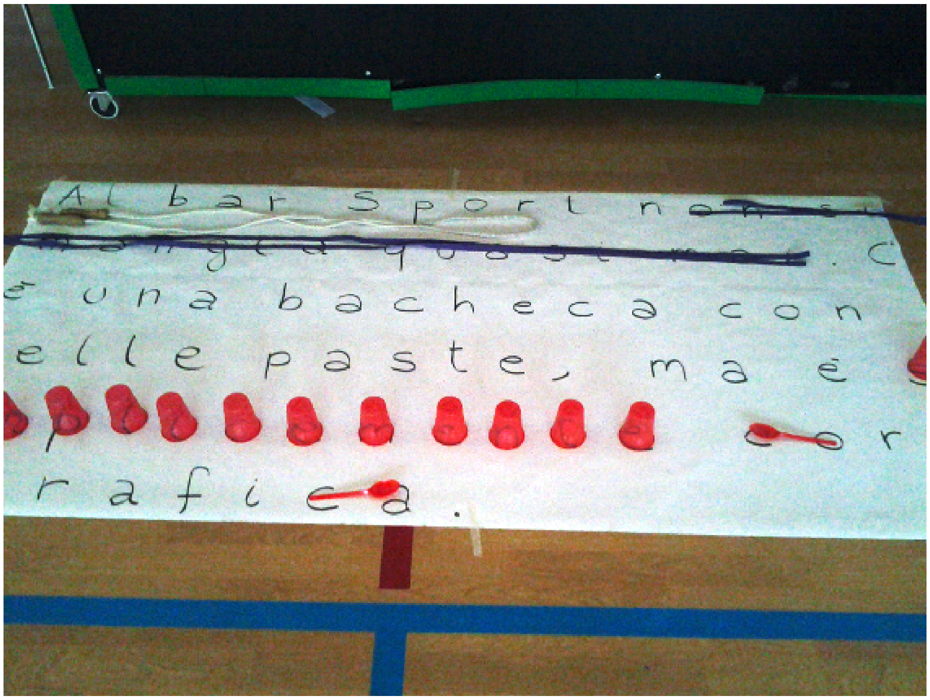

Fig. 1. Reproducing meta-information with objects

Guardando il **colore** della sua crema i vecchi riuscivano a *trarre* *le previsioni* del tempo.

Fig. 2. The text for the game

by changing its layout: in Fig. 3 the smallest piece of pasta is used both for bold (see the word "colore") and for yellow highlight (see the portion "le previsioni del tempo") and the different formatting elements are distinguished by the position of the piece, just underneath the first and last letter, or just above the first and last letter, respectively.

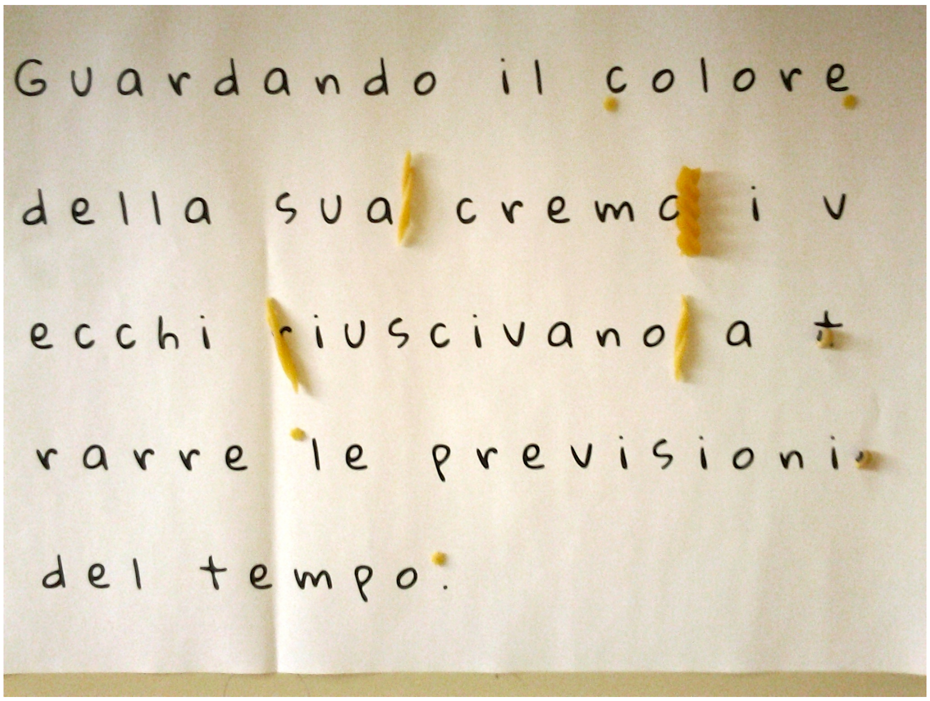

Fig. 3. Reproducing meta-information with *costly* objects

A second round of the game was proposed with cards instead of objects. The pupils only had a multi-set of alphabetic characters on small pieces of paper. Some letters and other keyboard characters (for example, the letters not included in the Italian alphabet: *j, k, w, x, y*) were not used in the text and could be easily exploited for weaving meta-information. However, this trick was not suggested by the facilitators but discovered by the pupils. Another solution proposed by the pupils was to use the characters with a meta-meaning by placing them with a different orientation: a ⌓, for example, for *bold*.

2.4 Rediscovering formatting tools

The final phase was carried out again in a computer lab. Pupils were introduced to a special software tool able to show a formatted text in three different views:

formatted, encoded using a wiki syntax, encoded with a simplified HTML-like syntax. They were again requested to reproduce formatted texts by working on the other views: the tool enabled them, however, to see immediately the effect of their (syntactical) manipulation in the formatted view.

Pupils showed no difficulties in accepting these new formal syntaxes, especially the wiki one: they immediately noticed the analogy with their own rules devised in the earlier activities; they also had already pointed out that their mates' rules, though different, all shared this same approach.

2.5 Evaluation of the experiment

Even pupils in 9th and 10th grade with a good familiarity with technological artifacts are generally unaware of two core principles in informatics: the necessity of absolutely unabiguous descriptions and the concepts of efficiency and optimization.

We think the experiment was successful. For example, all the pupils demonstrated to have grasped the idea of a meta-language expressed in the same alphabet as the language itself. This is considered quite an abstract concept, but it was accepted as rather natural (even obvious) by the pupils.

The link with word-processor and web technologies known to pupils was recognized. At the final recap we were also able to show that the same concepts are behind the scenes in several slightly different contexts, for example when editing Wikipedia entries.

The pupils' feedback was mostly positive: they did have fun and believed to have learned something. However, some of them said that the tasks were sometimes too easy, and the part in the gym (see Section 2.2) was considered boring by several participants.

3 How to design and implement an algomotorial sequence of activities

Planning a sequence of activities based on the approach we propose requires a careful preliminary design phase and the preparation of ad-hoc materials.

After setting our main didactic goals—namely to teach pupils the concept of mark-up language and that information coding and processing (as all automatic processes) need non-ambiguous rules—we had to identify a sequence of intermediate logical steps to allow pupils to gradually build up their knowledge. We identified the following logical steps:

1. given a formatted text, distinguish between text and formatting elements;
2. given a printed or hand-written plain text and a set of objects, find a way how to decorate the text in order to represent formatting elements, and formalize the rules that describe the functions of objects and the way to place them;
3. after introducing a reduction of available objects, strive to devise a more concise way to represent formatting elements, as is common in mark-up languages;

4. reproduce the whole formatted text by using the same "objects" (*i.e.* symbols, that is the keyboard characters) to represent both text and formatting, thus inventing one's personal mark-up language tag set;
5. given a fixed set of tags and a table describing their functions, apply them in an automatic wiki-like or HTML-like environment, with immediate rendering feedback.

Both steps 2 and 3 were repeated by using more and more abstract sets of objects: initially tangible objects featuring different colors, sizes, shapes (*e.g.* ropes, spoons, beakers, ...), then small objects with similar neutral features (*e.g.* kinds of pasta, see Fig. 3), and finally symbolic objects (character cards).

The careful design of the steps is critical, and we try to follow Vygotsky's concept of *zone of proximal development* [14]. In fact, if the steps are too gradual, some pupils do not find the activity interesting and get bored. On the other hand, if the steps are not gradual enough, some pupils get lost. Indeed, the same sequence can turn out to be adequate for some and not for others, as could be observed from the pupils' reports. For instance, some pupils chose step 3 as the least interesting, while others chose it as their favorite, one of them writing "It was demanding, but rewarding!"

Moreover it is very important to plan each step so that the activity makes sense *per se* in order to keep the pupils engaged [8]. With respect to this we found that the "didactic environment" [9] is critically important: we designed the dramatization phase with the goal of mark-up set-of-objects in mind, so we planned the activity in a large space and put the objects far from the working groups, expecting the distance would discourage the pupils from using too many objects. This did not happen at all. On the contrary, the setting of a game environment, where the goal was to economize on objects, naturally led them to discover the idea of marking only the beginning and the end of the text to be formatted. Similarly, we obtained poor results when we asked pupils to fill in a form writing a precise formulation of the rules according to which they placed the objects over the text: probably they did not feel interested enough in finding a good formal description, since an explicit context motivating such a request was missing.

Another critical point is to find the suitable trade-off between free exploring and external constraints. Pupils need authentic confrontations and should have the real possibility to make mistakes, but an open setting may have too many degrees of freedom and make pupils feel lost; on the other hand too many constraints may reduce their possibility to explore and investigate new solutions, hence inducing pupils to believe there is only one *right* answer. For instance, in the game activity (step 3) we asked pupils to write the rules they used to place pieces of pasta on the text, without handing out a form nor specifying the way how rules should be formalized; this freedom suggested to them not to use a verbal description, instead they opted for a more straightforward and effective description based on a tabular correspondence between objects and their function. It is worth mentioning another unexpected episode: the shortage of objects allowed a group to discover the possibility of coding formatting information by using only two objects, one at the beginning and one at the end of the text to be formatted, instead of using objects all along the text (often one for each character).

Finally, it is important that the teacher (better called the facilitator, or mediator [9]) observes what happens during the activities while keeping an open-minded prompt attitude. He/she should always clearly keep in mind the goal of the single activity and the overall sequence, helping pupils towards this goal, but should not force them along a set path. Unexpected events during the activities may be exploited to point out relevant issues which were not scheduled in the original design. For instance, the difficulty of some groups in providing a formal description of their rules was exploited to make them aware of the complexity of the task and of the amount of information we generally take for granted.

4 Conclusions

The proposed activity turned out to be a good way for conveying abstract computing concepts to pupils of secondary schools.

We are now working on refining the activity: as a mid-term goal we aim at working on producing didactic material that should be self-contained enough to be used by independent teachers in their classes. As a first step we re-proposed the activity in another school, with younger pupils (6th grade) under the guidance of a math teacher who had not participated in the conception. We are now studying reports and videos of the experiment: the first impression is that it worked also in this different context. We intend to repeat the experiment shortly with the support of a team of teachers in different schools.

We found that the *algomotricity* approach can be effective in presenting abstract symbolic manipulations in very concrete ways. By choosing activities clearly connected with the acquaintance of pupils with computers and tools we believe this approach can be very successful in fostering a thorough understanding of informatics concepts.

Acknowledgments

We would like to thank Giorgio Fattorelli and the school Marie Curie IIS for giving us the opportunity to experiment with our ideas in their school.

References

1. Barr, J., Cooper, S., Goldweber, M., Walker, H.M.: What everyone needs to know about computation. In: Lewandowski, G., Wolfman, S.A., Cortina, T.J., Walker, E.L. (eds.) Proceedings of the 41st ACM technical symposium on Computer science education, SIGCSE 2010, Milwaukee, USA. pp. 127–128. ACM (2010)
2. Begel, A., Garcia, D.D., Wolfman, S.A.: Kinesthetic learning in the classroom. In: Proc. of the 35th SIGCSE TSCSE. pp. 183–184. ACM, New York, USA (2004), `http://doi.acm.org/10.1145/971300.971367`

3. Ben-Ari, M.: Bricolage forever! In: Proceedings of the 11th Annual Workshop of the Psychology of Programming Interest Group, University of Leeds, UK (1999)
4. Ben-Ari, M.: Constructivism in computer science education. Journal of Computers in Mathematics and Science Teaching 20(1), 45–73 (2001)
5. Brinda, T., Puhlmann, H., Schulte, C.: Bridging ICT and CS: educational standards for computer science in lower secondary education. In: Proceedings of the 14th annual ACM SIGCSE conference on Innovation and technology in computer science education. pp. 288–292. ITiCSE '09, ACM, New York, NY, USA (2009), http://doi.acm.org/10.1145/1562877.1562965
6. Calzarossa, M., Ciancarini, P., Mich, L., Scarabottolo, N.: Informatics education in Italian high schools. In: Kalaš, I., Mittermeir, R. (eds.) Informatics in Schools, LNCS, vol. 7013, pp. 31–42. Springer (2011), http://dx.doi.org/10.1007/978-3-642-24722-4_4
7. Dagienė, V.: Informatics education for new millennium learners. In: Kalaš, I., Mittermeir, R. (eds.) Informatics in Schools, LNCS, vol. 7013, pp. 9–20. Springer (2011), http://dx.doi.org/10.1007/978-3-642-24722-4_2
8. De Vecchi, G., Carmona-Magnaldi, N.: Faire construire des savoirs (2ème ed.). Hachette Education (2003)
9. Giordan, A.: From constructivism to allosteric learning model. http://www.ldes.unige.ch/ang/publi/articles/unesco_AG_96/unesco96.htm (1996), UNESCO Conference on Science Education 2000+
10. Hromkovic, J.: Contributing to general education by teaching informatics. In: Informatics Education – The Bridge between Using and Understanding Computers, Proc. of ISSEP 2006, Vilnius, Lithuania. LNCS, vol. 4226, pp. 25–37. Springer (2006), http://dx.doi.org/10.1007/11915355_3
11. Hromkovic, J., Steffen, B.: Why teaching informatics in schools is as important as teaching mathematics and natural sciences. In: Kalas, I., Mittermeir, R.T. (eds.) Informatics in Schools. Contributing to 21st Century Education - 5th International Conference on Informatics in Schools: Situation, Evolution and Perspectives, ISSEP 2011, Bratislava, Slovakia, October 26-29, 2011. Proceedings. LNCS, vol. 7013, pp. 21–30. Springer (2011)
12. Taub, R., Ben-Ari, M., Armoni, M.: The effect of CS unplugged on middle-school students' views of cs. In: Brézillon, P., Russell, I., Labat, J.M. (eds.) Proc. of the 14th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education, ITiCSE 2009, Paris, France. pp. 99–103. ACM (2009), http://doi.acm.org/10.1145/1562877.1562912
13. The Royal Society: Shut down or restart? The way forward for computing in UK schools. http://royalsociety.org/education/policy/computing-in-schools/report/ (Jan 2012)
14. Vygotsky, L.: Mind in Society: Development of Higher Psychological Processes. Cambridge : Harvard University Press (1978)
15. Wing, J.M.: Computational thinking. Commun. ACM 49(3), 33–35 (Mar 2006), http://doi.acm.org/10.1145/1118178.1118215