César Córcoles; Antonia Huertas

Mathematical Communication and Representation in a Virtual Learning Environment

In: Petr Sojka (ed.): Towards a Digital Mathematics Library. Paris, France, July 7-8th, 2010.
Masaryk University Press, Brno, Czech Republic, 2010. pp. 95--104.

Persistent URL: http://dml.cz/dmlcz/702578

Mathematical Communication and Representation in a Virtual Learning Environment
A Case Study

César Córcoles and Antonia Huertas

Computer Science Department, Open University of Catalonia (UOC)
Rambla Poblenou, 156, 08018 Barcelona, Spain
http://www.uoc.edu/

Abstract. Abstract. At an exclusively online university such as the UOC the necessity for communicating mathematics in the web is pressing. In an environment that does not allow for face to face communication, things implicitly communicated when using a blackboard, such as the canonical verbalization or handwriting of formulae, are lost and become a big obstacle. Also, the editorial process for the creation of learning/teaching resources is suited for a generalist approach and, consequently, needs such as those presented by formula typesetting, especially for web-based materials, are not deemed a priority. In the last two years a series of innovation projects and initiatives have been set off in the UOC in order to improve the situation: the use of LaTeX and MathML standards in writing web resources, the use of LaTeX, MathML and other technologies in verbalization and locution of formulae, and the study of current possibilities for mathematical handwriting recognition.

Key words: mathematical communication and representation, eLearning, mathematical markup standards, web publishing, automatic verbalization

1 Introduction

At a virtual learning institution such as the UOC (http://www.uoc.edu/), an exclusively online university, the necessity for communicating mathematics online is pressing: students enrolled in mathematics-dense courses, mainly in computer and telecommunication engineering degrees, have a need to learn mathematics, and in order to do so they have to overcome the barrier posed by mathematical formulae [1]. In an environment that does not allow for face to face communication, things implicitly communicated when using a blackboard (such as the canonical verbalization of formulae) and commonly considered a "non problem" for teaching suddenly become a big obstacle. The detection of these difficulties, plus the perceived advantages digital resources could bring to the teaching/learning process led us to start exploring the available means to improve digital communication of mathematics in our learning environment. Different aspects make that a hard to overcome hurdle.

– The virtual learning environment has been in place since the birth of the institution fifteen years ago, and only recently an initiative to adhere to the OKI project has reached usable state, allowing for the use of tools widely available for other learning management systems.

Petr Sojka (editor): DML 2010, Towards a Digital Mathematics Library, pp. 95–104.
 ISBN 978-80-210-5242-0

– The institution has a strong, homogeneous pedagogical model, which has a list of strong advantages, but also means that the needs of subjects with special needs, both pedagogically and technologically, such as mathematics, will not automatically gain full support. Students, even those enrolled in technological degrees, do not necessarily have an inclination towards technology, so any "different" solutions they do not usually see in other classrooms is met with distrust or even rejection by a number of students. Additionally, a significant amount took their last mathematics course in secondary a long time ago and, as a result, common mathematical language has been forgotten. This situation is especially problematic when students are adults with professional experience, with not much time and with insufficient mathematical background who can only study at a distance [1], the typical student profile in an online distance university and more and more in the life long learning paradigm.
– The editorial process for the creation of learning/teaching resources is, again, suited for a generalist approach and, consequently, needs such as those presented by formula typesetting, especially for web-based materials, are not deemed a priority.

These factors lead to a search for off-the-shelf, easy solutions with a low cost and an even lower impact on the learning process, attacking the needs of users with very varying degrees of technological and mathematical savvy, but especially addressing those with a lower profile in both fields [2].

As the internet has become a medium for most activities, not only web-based teaching and learning [3], the Web 2.0 [4] and the future expectations for the Semantic Web have an effect on almost all of them. Essentially, we consider the web as a platform where software applications, rather than documents, live; where these software applications are designed to harness "collective intelligence" and effectively move from a developer-centric point of view to a user-centered one. According to that user-centered model, the user's data is the most important element in every transaction, and developers allow users total or almost total control of their assets. One very meaningful cause/effect of that is the extreme lowering of the cost of access to very sophisticated resources to have a strong presence on the web. The advent of lightweight content management software allows the average web user or organization to publish quality information in a comfortable and efficient way. This leads naturally to a digitalization of contents and, in the case of mathematical e-learning, to the building of a digital mathematics library, containing resources of all kinds, both for research and educational purposes, and covering all ranges of mathematical sophistication.

For mathematical educational resources, though, the situation is not as advantageous [5]. Firstly, mathematical language requires a set of tools which, at the moment, are not sufficiently widespread. Thus, publishing web pages containing formulae marked in the MathML presentation standard doesn't automatically mean that most users will be able to read those web pages, and a number of alternative methods must be put in place. Secondly, authors and editors of educational resources—as we have already mentioned for users—are not necessarily tech proficient, so content creation tools not aiming for the base line, in our experience, run a big risk of not being adopted, leading to a lack of digital resources in the library, creating a sort of digital divide for mathematics educational resources.

In our particular case, moreover, educational materials are produced through a process that has been developed for all subjects at the University and then delivered through the institutional Learning Management System (LMS). None of the two phases is particularly suited to mathematical content. This content, though, must be delivered in an adequate way to students without breaking neither the technological environment nor the pedagogical model. Thus, we have looked for a number of solutions, that, while far from optimal in many aspects, are currently working and allow for acceptable academic performance.

We will now present the solutions we have developed and used grouped in two main categories: those for mathematical content delivery on the web and those for the verbalization of mathematical formulae: how to write, and then how to read, mathematics on the web.

2 Writing Mathematical Content on the Web

Mathematical content delivery has gone in the last thirty years from "chalk on blackboard" and manual typesetting to TEX and LATEX typesetting—which, with the automatization and precision it brought about, meant a revolution—and is currently moving to a web based model [6]. In recent years, research has focused on the semantic aspects of content, as exemplified by the OMDoc markup document for mathematical documents [7] or the ActiveMath web based learning environment [8]. Semantic information associated to digital content is a necessary step for added value services. In the field of mathematics education, one added value is the verbalization of formulas, not only for visually disabled people but also for long-life learners who have disconnected the visual representation of a formula and its verbalization. Other added value services may include graphical representation of a math formula or web searching in semantic-based databases, as exemplified by the NIST Digital Library of Mathematical Functions [9].

MathML, at present, seems to be the only viable alternative for content delivery if no means are available to develop a custom solution. However, it is not without its problems, especially in our context. Being a markup-dense language, it is not as easy to write or understand without the use of computer software as LATEX, for example, can be. When producing MathML content, that means that most of it will not be authored directly in MathML: the author will either use some kind of WYSIWYG editor, provide LATEX content to be translated, or provide the output of a computer algebra system as input. The first two cases can be the cause of semantically erroneous but "visually correct" markup. Also, when dealing with an editorial process that is under a severe workload and for which mathematical content is a very small fraction of the work, asking editors to use new technologies can be a source of friction and costly errors in the process.

2.1 Web Based Mathematical Resources

In this section we present our experience creating an interactive web based learning material [10] designed for a calculus course in the computer engineering studies.

This material uses the Wiris software [11], an on-line Computer Algebra System, which allows mathematical calculations on-line. This software was chosen because it offered some features that made it more suitable to our necessities than other existent commercial software, in particular its availability of both on-line and local versions (so students should be able to work from any computer connected to Internet, which is one of the main features of UOC. It is also a multilingual tool that allows mathematics computing to be done in the different languages (Spanish or Catalan in our case), and it allows us construction of interactive learning exercises. The learning material we are going to present, although impregnated with the use of Wiris, was designed and created to be used with any other software of similar features.

When that material was designed in 2006, perhaps the most widespread tool to export mathematical notation to the web was the one used in Wikipedia, Texvc [12], transforming standard mathematical LaTeX notation into PNG graphic files. Other available tools were TeX4ht [13], Hermes [14], TtM [15] or blahTeX [16]. They all work in approximately the same way: they take the standard text and convert it into (X)HTML, and mathematical formulae are converted into either graphic files or MathML mark-up. Thinking about "forward-compatibility", and worrying about the inevitable loss of information when converting mathematical mark-up into images, we chose the tex4moz LaTeX package to generate XHTML plus MathML. At the time Firefox was in its 1.5 release, the first one with native MathML support (depending on the operating system, some additional typography downloads were necessary), and support for Internet Explorer was in the form of a freely available plug-in [17].

The conversion process was not trivial: tex4moz was still in development, and while it worked remarkably well, it had some troubles, if non-standard packages were used or if LaTeX markup was not correct. Errors in the generated XML document were hard to correct, because of the working of XHTML+MathML, which won't render non-valid documents and providing no feedback about the nature of the error, and extensive use of the equation editor in Amaya [18] or other such tools was required.

From a developer's point of view, Wiris takes the form of a Java applet taking its parameters in the form of MathML commands. Neither the developer nor the students need to have any knowledge of MathML, as the applet provides an easy to use editor and menu based tools.

As a result we obtained a simple interface, with acceptable mathematical notation and easy navigation. Furthermore: where in a traditional material we would have an exercise, we were able to provide a link to some interactive material developed with Wiris providing an infinite number of exercises. Where in a traditional book we would have a figure and a written explanation, here we were able to offer the possibility of a dynamic experiment to practice the concepts.

As a result we obtained a simple interface, with acceptable mathematical notation and easy navigation. Furthermore: where in a traditional material we would have an exercise, we were able to provide a link to some interactive material developed with Wiris providing an infinite number of exercises. Where in a traditional book we would have a figure and a written explanation, here we were able to offer the possibility of a dynamic experiment to practice the concepts.

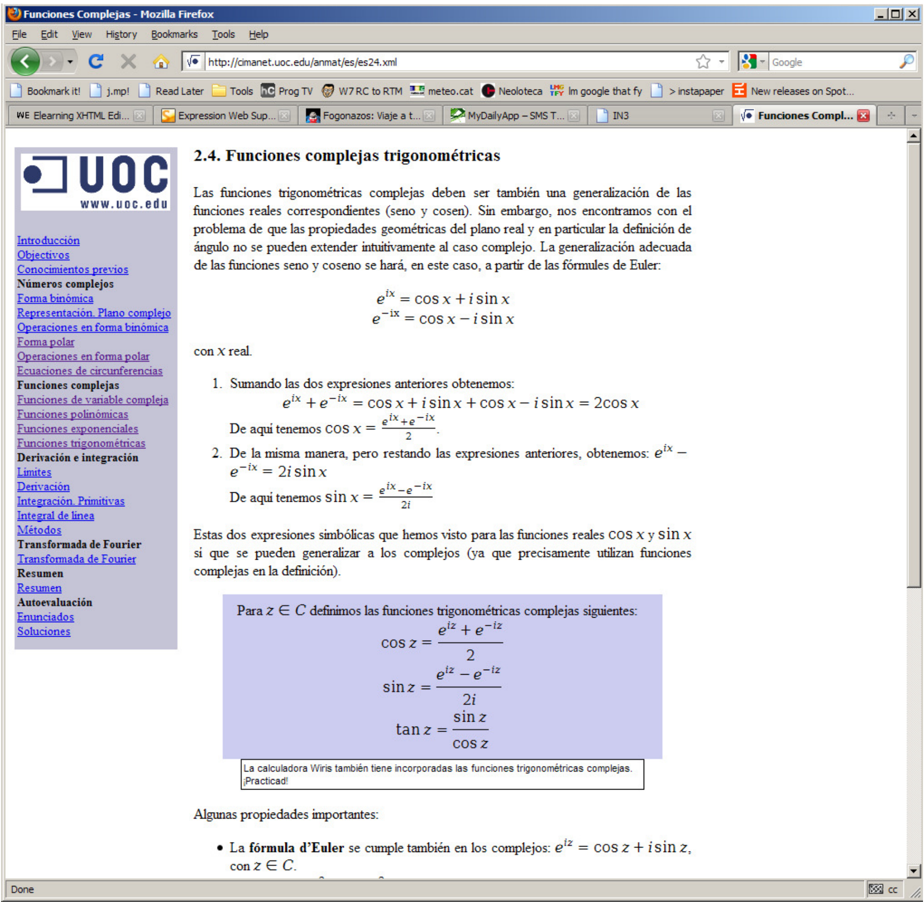

Fig. 1. Example of the UOC-Spanish language version of the web based material

2.2 Mathematical Content and Email

Of secondary interest for the creation of a digital mathematics library, but essential for mathematical e-learning, is the use of mathematical formulae in online communication. As we have mentioned, we are bound by the institutional LMS, with its own HTML based mail solution, which doesn't allow the use of MathML presentation markup.

That had severely limited the ability of students and faculty to use mathematical formulae in communication, forcing users to resort to verbalization of formulae or the attachment of scanned handwritten materials or documents produced by tools with WYSIWYG formula editors, mainly Word and OpenOffice Writer. These solutions were slow and awkward, and represented an obstacle to communication. In informal inquires we found that those barriers were keeping students from presenting doubts they had in the Virtual Learning Environment.

Once the need was detected we decided that the best solution would be a plug-in for the webmail platform, which would be made available progressively to those students enrolled in mathematical courses. Another design constraint that had to be taken into account was that it should be as transparent as possible to those users who didn't want to adopt the solution. The chosen solution was a pseudo-LaTeX language which could be used with minimal training and was readable even without the plug-in rendering the formulae.

We observed from the very first pilots that the amount of communication in the classrooms involving formulae increased significantly.

3 Reading Formulae on the Web: Verbalization and Speech Tools

As previously stated, the reading of mathematical formulae, while trivial for students coming out straight from secondary education and even those who don't but are learning in a traditional blackboard environment represents a problem for those who took their last mathematics course long ago and now don't have an instructor reading aloud the formulae as she writes them. Communication, understanding and memorization efforts are sensibly hampered, thus, for a sizable amount of students.

While we were especially worried about the impact of not being able to verbalize formulae on learning performance, when talking about a future digital mathematics library the focus would be on the accessibility of the contents in the library, both for people with sight disabilities and those whose disability comes from lack of an adequate and current education.

Two different approaches have been taken to solve the problem. On one hand we have Rodolfo, a small "formula repository" containing formulae and human read MP3 audio snippets for them and, on the other hand, a project developed by Maths for More to take MathML (presentation) and LaTeX formulae embedded in XHTML pages and automatically produce a transcription for each of them.

Rodolfo is a small, basic and simple web-based database storing formulae, encoded either in LaTeX or in MathML presentation, plus their different possible

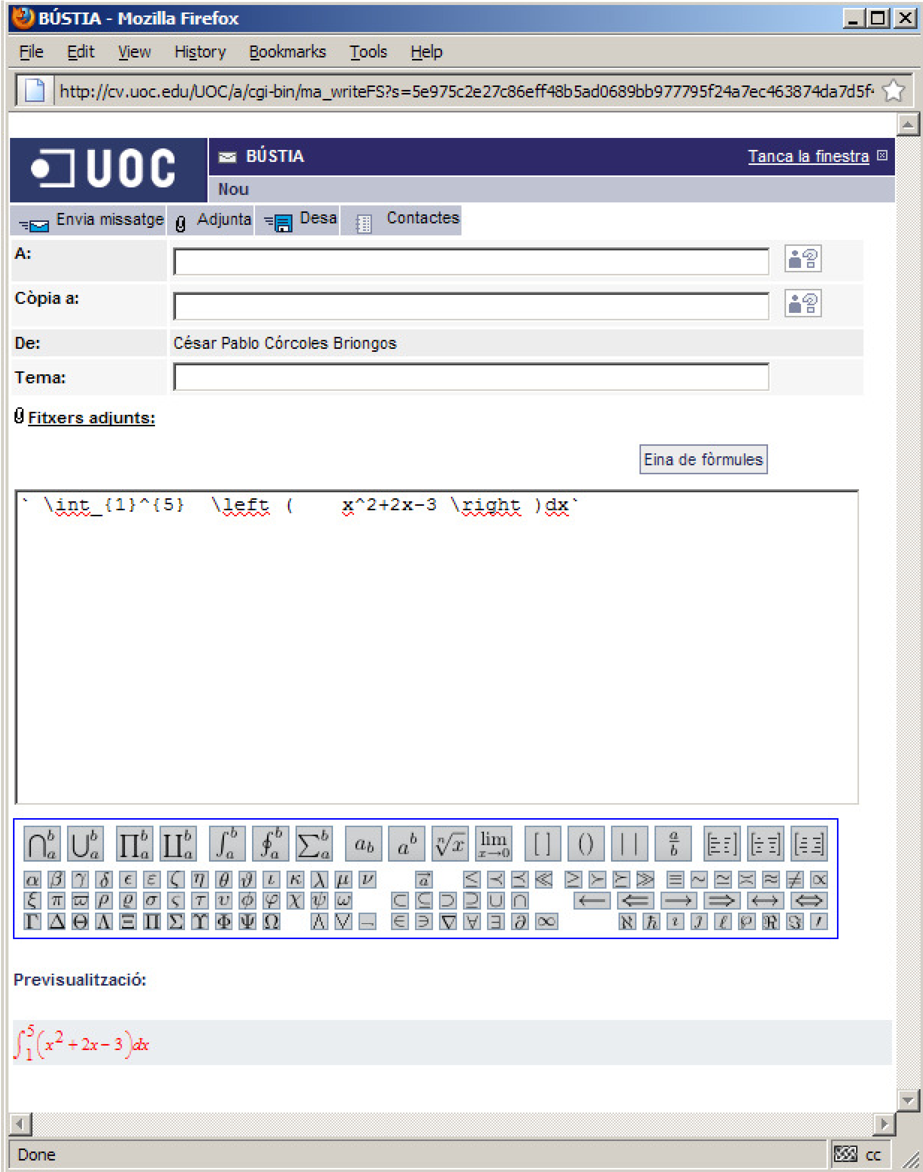

Fig. 2. Example of the pseudo-LaTeX language and visualization in the UOC mail platform

readings in any language and MP3 files of those readings. The project was born of the need to embed those audio readings in web based materials to improve academic performance, after detecting a number of students having trouble with them. Thinking about reuse, a very simple framework was created so files

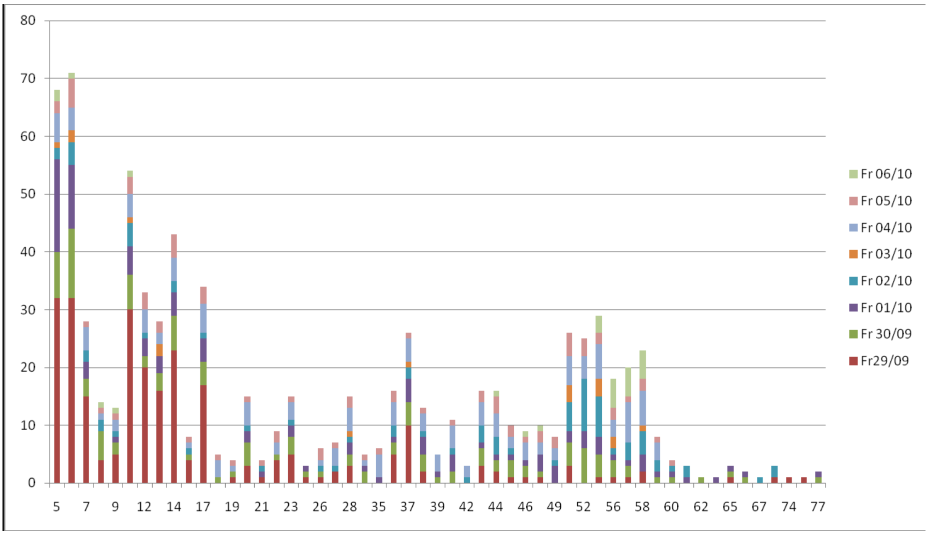

Fig. 3. Information from server logs regarding MP3 use for the different formulae in a learning material. Spikes correspond with new terms appearing in the particular formula, proving the need students have of a verbalization tool.

could be retrieved with moderate ease. At the moment, only LaTeX search has been implemented, not semantic search.

Besides being a useful tool as it is, Rodolfo seems to be in line with efforts such as the "notation census" hosted by the Math-Bridge wiki. Also, it should not be extremely difficult to build it into a "formula repository" [19] which could also link to related definitions or content in a semantically meaningful way.

A completely different approach to the same problem is presented by the (currently nameless) verbalization tool developed by Maths for More: it is an automatic tool (server-based, developed in Java) scanning web pages for LaTeX and/or MathML presentation formulae and generating automatically a text transcription [20] (which could then be sent to a Text-To-Speech engine). Obviously, trying to extract meaning from presentational markup such as the one available is bound to present ambiguous situation. Luckily, the scope of every single webpage is limited, so we can resort to adding a meta tag to the page indicating the domain it covers so that the tool can use that knowledge for disambiguation.

4 Final Remarks and Future Work

Advances in web-based mathematics education depend heavily on the technologies to communicate the language of mathematics on the web and the specific contexts and audiences we are targeting. We have presented experiences taken place in an exclusively online university such as the UOC where the difficulties in the development and deployment of these technologies have a deep effect on the technical studies which need to use mathematical content.

A few years ago mathematics courses at the UOC were delivered as handbooks, taking no profit of the advantages presented by hyperlinking, multimedia and interactivity. While we cannot say that we are anywhere near building a digital library of mathematics, steps have been taken towards the digitalization of every mathematics related course at the institution. Some of those steps have allowed the use of resources such as a simple Computer Algebra System, allowing us to focus more on the meaning and the understanding, and less on the memorization of algorithms by students. We have also seen how digitalization—and keeping or adding as much semantic information as possible—has allowed for verbalization, leading to better accessibility and communication among students and teachers, which was also helped by adding a simple, transparent tool to help students with formula writing. In the future, linking to resources such as the NIST Digital Library of Mathematical Functions should allow for a much richer and more productive learning experience by students.

There is still a lot of work to be done so that we can say that the transmission and consumption of mathematical content is not a problem at an online University such as UOC and that our online content can be called a *real* digital library. In particular, besides going further into each of the lines we have recently started looking into available solutions for the input of formulae using handwriting, using digitizing tablets or tablet PCs as input devices and software to record that input and, if at all possible, convert it into semantic content.

As a final remark, we state our strong opinion that, much in the same way widespread LaTeX adoption depended on the availability of easy, robust authoring tools, we predict the same thing will happen with web based mathematics. In our case, we are at a point where the web as a platform is finally reaching enough maturity as to represent a viable publishing platform with the adoption of HTML5 (plus technologies such as MathML but also SVG or Canvas, for example). The mathematics community has always been at the forefront in web specification writing, with MathML being the first the first XML language recommended by the World Wide Web Consortium in 1998, with continuing work today to improve it and adapt to the times. And thus we believe that, in order to see that standard take the current mathematical library effectively into the digital era, a bigger stress must be put in the development of tools that will allow both current LaTeX users and newcomers to use them efficiently. Also, a strong effort must be made so that potential users know available tools and can locate them easily and find their corresponding documentation and tutorials.

Acknowledgments. This work has been partially supported by three UOC educational innovation projects the years 2007, 2008 and 2010, respectively, and by the eLearn Center.

References

1. Juan, A., Huertas, M., Steegmann, C., Córcoles, C. and Serrat, C.: Mathematical E-Learning: state of the art and experiences at the Open University of Catalonia. International Journal of Mathematical Education in Science and Technology, vol. 39, no. 4, pp. 455–471 (2008).
2. Kahn, P. and Kyle, J. (eds.): Effective learning & teaching in Mathematics & its applications Kogan Page (2002).
3. Alexander, B.: Web 2.0: A new wave of innovation for teaching and learning? *EDUCAUSE Review*, 41(2), 32–44 (2006). `http://www.educause.edu/apps/er/erm06/erm0621.asp?bhcp=1`
4. O'Reilly, T.: What is Web 2.0. O'Reilly's [blog] (2005). `http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html`
5. Miner, T. and Topping, P.: Math on the Web: A Status Report – Focus: Distance Learning, (2001). `http://www.dessci.com/en/reference/webmath/status`
6. Pajo, K. and Wallace, C.: Barriers to uptake of web based technology by university teachers, J. Dist. Edu. 16, pp. 70–84 (2001).
7. Kohlhase, M.: OMDoc: Towards an Internet Standard for the Administration, Distributions and Teaching of Mathematical Knowledge. Proceedings of Artificial Intelligence and Symbolic Computation. Springer LNAI (2000).
8. Melis, E., Andrès, E., Büdenbender, J., Frischauf, A., Goguadze, G. Libbrecht, P., Pollet, M and Ullrich, C.: ActiveMath: A Generic and Adaptive Web-Based Learning Environment. International Journal of Artificial Intelligence in Education, no. 12, pp. 385–407 (2001).
9. NIST Digital Library of Mathematical Functions. `http://dlmf.nist.gov`
10. Bringslid, O. Mathematical e-learning using interactive mathematics on the Web, Eur. J. Eng. Edu. 27, pp. 249–255 (2002).
11. 'Maths for More, `http://www.mathsformore.com/`
12. Texvc, `http://en.wikipedia.org/wiki/Texvc`
13. TeX4ht: LaTeX and TeX for Hypertext, `http://www.cse.ohio-state.edu/~gurari/TeX4ht/`
14. Hermes – a semantic XML e-publishing tool for LaTeX authored scientific articles, `http://hermes.roua.org/`
15. TtM, a TeX to MathML translator, `http://hutchinson.belmont.ma.us/tth/mml/`
16. Blahtex, `http://www.blahtex.org/index.php?page=home`
17. MathPlayer, `http://www.dessci.com/en/products/mathplayer/`
18. Amaya Home Page, `http://www.w3.org/Amaya/`
19. Minguillón, J., Huertas, M. A., Juan, A. A., Sancho, T., Cavaller, V. (2008). "Using learning object repositories for teaching Statistics". Proceedings of the First Workshop on Methods and Cases in Computing Education. Salamanca: Pàg. 53–61. ISBN: 978-84-691-8558-2.
20. Sancho, T., Córcoles, C., Huertas, M. A., Pérez, A., Marquès, D., Villalonga, J. (2008). "Automatic Verbalization of Mathematical Formulae for web-Based Learning Resources". A: REMENYI, D. The Proceedings of the 7th European Conference on e-Learning. Academic Publishing Limited. Pàg. 405–414. ISBN: 978-1-906638-23-8 Cd.