Using GF in multimodal assistants for mathematics

Dominique Archambault
EA 4004 – CHArt-THIM
Université Paris 8
dominique.archambault@univ-paris8.fr

Olga Caprotti
Chalmers University
Gothenburg, Sweden
caprotti@chalmers.se

Aarne Ranta
Chalmers University
Gothenburg, Sweden
aarne@chalmers.se

Jordi Saludes
Universitat Politècnica de Catalunya
Barcelona, Spain
jordi.saludes@upc.edu

Abstract

In this paper we present an ongoing effort in the computational linguistic field to develop multilingual assistive technologies for mathematics. The approach is based on the Grammatical Framework used in combination with UMCL and semantic encodings of mathematics like OpenMath and MathML3. We review the formats for supporting alternative modalities of presentation of mathematics in order to evaluate the effectiveness of the proposed approach. Finally we describe a very initial prototype built using off-the-shelf speech synthesis to interface to the Sage suite of computer algebra systems.

1 Introduction

Mathematics education has recognized strategic importance in today's technologically-oriented society. The necessity to widen scientific education to a broader audience, preferably with fewer resources, is an equation that cannot be solved by mathematics alone. Cognitive science, educational expertise, computer science and linguistics are also needed. In this paper, we discuss how to equip with multimodal accessible interfaces tools, services, and sample content which automate parts of mathematics instruction. Proper use of multimodality will make the delivery of education much more effective without compromising the quality.

The core enabling technologies for multimodality in mathematical learning assistants combine advances in computational linguistics with rich semantic markup formats for mathematics such as MathML and OpenMath. When used to represent mathematical fragments within digital resources, they allow the automatic generation of multimodal alternative presentation of the mathematics content. These include onscreen display with highlighting that is synchronized with synthetic speech, thereby adapting to the learning requirements of students needing multimodal visual and aural presentations. The real issue is that ICT has not yet been effectively employed and exploited to support mathematical activity. Still today there is a noticeable mismatch between mainstream software for mathematics and existing screen readers when it comes to multimodal presentations of formulae. In many cases still (prominent ones such as Wikipedia and mathworld.wolfram.com) formulae are printed on the screen using graphical glyphs. Hence, learning environments are basically unusable for mathematics by people with disabilities such as dyslexia, dysgraphia, and any visual impairment.

When interactivity is required, for instance to enter a solution in a input field, alternative means of navigation, for instance by vocal control or by keyboard, provide an easy solution usable by everyone that can be achieved by supplying a rich source document to the browser.

In [16], the US National Mathematics Advisory Panel highlighted the importance of mathematics in future and that "high-quality computer-assisted instruction (CAI) drill and practice, implemented with fidelity, be considered as a useful tool in developing students' automaticity (i.e., fast, accurate, and

K. Yamaguchi, M. Suzuki (eds.): DEIMS 2012, volume 1, issue: 1, pp. 1-10

effortless performance on computation), freeing working memory so that attention can be directed to the more complicated aspects of complex tasks." The US Panel also noted that more high quality studies regarding proper use of technology in instruction are needed.

Learning to master certain routine tasks, like computing with fractions, solving equations, computing limits, derivatives and integrals, forms a large part of mathematics education in high schools and in the beginning of tertiary education. Training these skills produces the computational fluency and execution of procedures required, together with conceptual understanding, to support efficient problem solving. The delivery of drills and of formative assessments can be done especially well by automatic learning systems. However, further research is needed to devise cost-effective ways in which to enhance mainstream assessment systems by verbal and multilingual presentations, thus personalizing the delivery of the material according to the students' preferences. Multimodal representation and interfaces have shown to increase the learning outcomes in primary school ICT-supported mathematics instruction.

In the field of Assistive Technologies, *i.e.* technologies helping people with disabilities, ICT can provide extra assistance in overcoming "print disabilities", namely low vision, blindness, and dyslexia. It is worth reminding that the European Commission adopted, the eAccessibility Act [1] in an effort to develop, implement and promote standards for eAccessibility, disseminate education on designing resources and methods that are fit for all, and foremost to research new ICT technologies that address the needs of people with disabilities.

2 Digital formats for Mathematics

When mathematical content is transmitted electronically or using any other media, it must necessarily be encoded using some format, which is rich enough to carry the informational content suited for the end user or application. The choice of such a format depends on many factors and bears consequences in terms of how well the content travels in space (e.g. across borders) and in time (e.g. can it be archived and processed by a computer?). In this paper we only review those formats that impact multimodality in terms of aural, visual and multilingual display of mathematics expression. When dealing with larger mathematical contexts, such as a lecture or a scientific paper, the formats might become more complicated, yet they probably still relate to the basic ones.

The Mathematical Markup Language MathML is the XML language recommended by the World Wide Web Consortium for electronic representation of mathematics on the internet, intended to facilitate the use and re-use of scientific content on the Web. For other applications such as computer algebra systems, and print typesetting, MathML-Presentation has been primarily used to encode mathematical notation for high-quality visual display. Its semantic counterpart, MathML-Content is designed for applications where semantics plays a key role such as scientific software or voice synthesis. However, only few tools currently produce MathML-Content and semantic markup of mathematics has been mostly encoded using OpenMath [6]. With the new MathML3.0 specification [4], the XML encoding of MathML-Content has aligned itself with the OpenMath XML encoding, *de facto* unifying the two semantics-based representations.

MathML has been used in the Digital Talking Book format promoted by the DAISY Consortium [2] in their specification extension since 2007. Yet, according to the latest information by the US National Center on Accessible Instructional Materials (NIMAS[3]), many DAISY books still display mathematical content by images with alternative texts, due to the relatively poor support of MathML in software tools.

[1] `http://eur-lex.europa.eu/smartapi/cgi/sga_doc?smartapi!celexplus!prod!DocNumber&lg=en&type_doc=COMfinal&an_doc=2005&nu_doc=425`

[2] `http://www.daisy.org/project/mathml`

[3] `http://aim.cast.org`

To remedy this, the Web has offered already a portable solution in MathJax[4], a cross-browser display engine for mathematics encoded in LaTeXand in MathML. MathJax is based on JavaScript and can be easily integrated in web sites, for instance in the Moodle learning management system, thus advancing the usage of formats that are more amenable to assistive software. Open source tools for the production of DTB books are collected via the Daisy pipeline, an Eclipse-based platform. It is a framework that allows to apply a number of transformers converting one format into another. A quick look into the documentation for the converter of DTB to LaTeX, `se_tpb_dtbook2latex`[5], lists the support for MathML as future work. This a rather straightforward task that could be achieved by calling the GF web service, in a manner similar to what is explained later.

The focus of DAISY and NIMAS is on accessible books and is therefore complementary to the approach described here which addresses interactive software and tools used to better mathematics instruction. We have not been able to further investigate the availability of DTB players that support speech rendering of mathematics. However we believe that the prototype described in this paper can serve as model to enhance a number of existing tools such as those described on the DAISY wiki[6]. The work relies on two key enabling technologies: OpenMath, used as a language-independent representation for decorating mathematical text; and the Grammatical Framework, for the automatic generation of natural language presentations.

The approach builds and extend the results of the eContent EU project WebALT (EDC-22253-WEBALT). In [7] advanced language technologies are employed to encode language-independent mathematical text so that the resulting representation is amenable to automated generation of e-content in a number of natural languages. The technology developed during the WebALT project handles only a small fraction of the full domain of applicability, namely that of short sentences which state a mathematical problem to be solved. For this restricted domain, which however covers a large amount of interactive educational content, it is possible to edit a multilingual version and to produce automatically a version in a natural language, including Spanish, English, French, Swedish, Italian, Finnish, and Catalan. By adapting the presentation of the mathematical content to the native language of the student, the system adapts to the learning abilities of its user and presents the questions in a natural interface. For the tutor, it is possible to write exercises in one language, to store them using a language independent representation, and to make them available to various lecturers in their own preferred language. The major software tool used by the WebALT mathematical grammar library to carry out advanced language processing is GF, the Grammatical Framework [18, 19].

The Grammatical Framework[7] is a grammar formalism based on type theory and consisting of a special-purpose programming language, a compiler of the language, and a generic grammar processor. The compiler reads GF grammars from user-provided files, and the generic grammar processor performs various tasks with the grammars: generation, parsing, translation, type checking, computation, paraphrasing, random and exhaustive generation, and syntax editing. GF is particularly interesting for the modular design which promotes grammar engineering, re-usage of grammars as software components, and multilinguality via the GF Resource Grammar Library (RGL). This library is a general resource that has been used in many projects involving language processing, including WebALT and currently MOLTO[8], Multilingual Online Translation. In the current version (3.1), it covers 26 languages.

The WebALT mathematical grammar libraries (MGL) are GF grammars that handle the language fragment consisting of problem statements that are brief and technical, of the style "Compute the sum ... where ...". The MOLTO project extend these libraries by enlarging the mathematical language fragments

[4] http://www.mathjax.org
[5] http://data.daisy.org/projects/pipeline/doc/transformers/se_tpb_dtbook2latex.html
[6] http://www.daisy.org/daisypedia/mathml-in-daisy-resources
[7] http://www.grammaticalframework.org
[8] http://www.molto-project.eu

that can be encoded in a language-independent way and by adding more languages, including LaTeX. The modular design of the MGL makes is easy to add new "linearizations" to the library, namely to program how to generate multimodal presentations of the content suitable for speech synthesizing, or read by peripherals designed to enhance accessibility.

The target corpus includes basic mathematics, high-school mathematics and statistics, and introductory university courses such as calculus, linear algebra, logic and analytic geometry. In particular, it also tackles the formal mathematical vernacular used to introduce a new concept in a definition, to state a property or a theorem and possibly a proof. We envision that in a near future this will allow one to represent, in a language independent way, educational materials (including books in DTB) for a lesson, but most importantly, interactive quizzes and examinations with detailed multimodal dialog feedback.

MOLTO is also investigating the possibility of tackling more applied problem statements, in the style *"The annual production of a gold mine was measured to be..."* and including typical relational word problems, which require students to elaborate an algebraic model in order to find a solution: *"Paul is forty years old. He has two sons, Mat and Kevin, and one daughter, Emily. Mat is one quarter of Paul's age. Kevin is two years younger than Mat. How old is Kevin?"*

W3C has put a lot of effort in designing recommandations for voice enabled web applications. The W3C Speech Interface Framework ("Voice Browser" Activity) is a suite of markup specifications aimed at enabling the creation of web applications that can be accessed via any telephone, and allowing people to interact with these applications via speech and telephone keypads. It covers voice dialogs (VoiceXML), speech synthesis (SSML), speech recognition (SRGS, SISR), pronunciation lexicon (PLS), call control (CCXML, SCXML) and other requirements for interactive voice response applications, including use by people with hearing or speaking impairments.

The Voice Extensible Markup Language (VoiceXML) is designed for creating audio dialogs that feature synthesized speech, digitized audio, recognition of spoken and dual tone multi-frequency key input, recording of spoken input, telephony, and mixed initiative conversations. We note that GF has already been used in connection with VoiceXML in the TALK project (TALK, 2004-2006) and is equipped with a library to export the VoiceXML format.

Very recently, tactile feedback is also being addressed by the new Device APIs Working Group who has published the First Public Working Draft of the Vibration API.

Finally, the widely spread Cascading Style Sheets (CSS) is a language that describes the rendering of HTML and XML documents not only on screen, but also on paper (embossing), and in speech. In fact, CSS defines aural properties that give control over rendering XML to speech by the media types 'aural' and 'speech'. The *CSS for Speech* module [24] describes the text to speech properties proposed for CSS Level 3. They are designed for matching the model described in the *Speech Synthesis Markup Language* (SSML) Version 1.0 and allow producing an aural presentation associated to the structure of a document in a way similar to how visual presentations can be associated for printed media or screening rendition. When using voice properties, the canvas consists of a two channel stereo space and a temporal space in which to specify audio cues before and after synthetic speech. The CSS properties also allow authors to vary the characteristics of synthetic speech (voice type, frequency, inflection, etc.).

3 Alternative modalities to access Mathematics

Multimodal presentations and interfaces have shown to increase the learning outcomes in primary school ICT-supported mathematics instruction, in particular where the content was represented both verbally (printed or spoken words and formulas) and non-verbally (illustrations, photos, videos and animations) [14]. Modalities correspond to codings systems for transmitting information, through a give human sense. For instance speech, music and non speech earcons are 3 audio modalities. The main modalities used to interact with a computer are visual, while blind people use audio and tactile modalities. Traditional mathematics education is mainly visual whereas ICT-supported instruction can introduce alternative modalities and modes.

There are two key schemes for representing mathematical content in order to render its structure aurally: lexical and prosodic cues. Lexical markup explicitly denotes structural information such as the delimiters of a radical expression or a fraction so that for instance one expresses with *"begin-frac a over b end-frac"* the fraction $\frac{a}{b}$. The drawback of this approach, widely used in many tools, is the cognitive overload on the listener's working memory. Prosodic cues are pauses and modifications in the pitch, tempo, rhythm and tone which are introduced to help correlate the aural rendering to the structure of the expression. The presentation of syntactically complex material conveys the structure, and grouping of an expression using pausing and alterations in the speaking rate adopting the paradigm that it is possible to convert a sequence of juxtaposed symbols, delimited by both white space and other visual cues (such as parentheses) into a serially transmitted linguistic approximation. Any phonological model of verbal mathematics must be able to represent ambiguities that arise from the inherent 2-dimensional nature of the written expressions. Consider for instance a^{b+c} and $a^b + c$, where exactly the same symbols are used. Published work on prosodic presentation of mathematical content include [22, 20, 15] and more recently [10, 11].

Many of the modern design features of current speech systems have been pioneered in the PhD Thesis by Raman [17]. Here an Audio System for Technical Readings (AsTeR) allows reformatting of electronic documents in LaTeX (a very popular typesetting system for mathematical documents) and to produce audio documents. AsTeR introduces a rule-based language, called Audio Formatting Language (AFL), in order to render the internal document representation to audio thus effectively distinguishing the semantic structure of the document from one of its presentations, e.g. audio. Statements from AFL are grouped into rendering rules for each object of the document tree, and groups of rules can be themselves grouped into a rendering style in a manner which is similar to the notion of a style-sheet that produces MathML presentation from content markup. Multimodality is achieved by activating and deactivating specific rules. Audio rendering uses a mixed approach of lexical and prosodic cues for instance higher and lower pitch to reflect superscripts and subscripts and grouping with variable substitution, like in *"fraction x over y where x is ... and y is ..."* when the complexity of the sub-expressions is sufficiently large.

AudioMath [12] is a project whose main aim is to provide a tool, to work either standalone or integrated with a text-to-speech engine, that converts expressions presently not "understandable" by a regular text-to-speech engine, both in text and in mark-up elements (MathML Presentation). Currently it only supports European Portuguese. The features of the mathematics audio rendering are parsing, interpretation and conversion of MathML into plain text format, generation of the appropriate prosodic contour for reading of the math formula's text, and intra-formula browsing device.

More recently [5] has proposed to use a limited set of spatialized earcons and spearcons (which they define as audio icons based on very short accelerated speech), to disambiguate the structure of mathematical formulae, in an intuitive and unambiguous manner. They try to reduce the cognitive demand due to the use of multiple cues, spatial audio, complex hierarchies of non-speech sounds to represent the structure and scope of equations. Spearcons lie somewhere between a clearly comprehensible spoken utterance and an abstract non-speech sound, easier to learn and to recognize, which makes them good candidates to expose structural elements such as fractions, superscripts and subscripts in an intuitive way.

During the VICKIE project[9] [1], the main commitment was to provide a technological environment facilitating the inclusion of a visually impaired student in the mainstream education. As a result of that work, the first prototype version of the Universal Maths Conversion Library (UMCL) [2, 3] was

[9] Visually Impaired Children Kit for Inclusive Education (EU funded project IST-2001-32678)

produced to provide an abstract programming interface that would translate among various formats for mathematics, including those more relevant to accessibility such as the national Braille languages (French and Italian), to and from MathML and LaTeX. The universal representation chosen for the mathematical formulas is MathML presentation. In particular, any MathML expression is pre-processed and reduced to a canonical form in order to capture similarities that would translate to the same Braille representation. The UMCL enables applications to provide Braille experienced users with the notation they have been taught, so they do not have to learn another one. While the choice of Presentation MathML can be easily motivated by the availability of tools to produce presentation markup, it has been argued extensively in the literature how, when possible, the semantic representation, e.g. in MathML-Content or OpenMath, is more flexible and more amenable to manipulation. We are considering extending the MOLTO MGL by UMCL modules which handle input and output in the formats suitable for aural rendering and for processing multimodal mathematics material. Notice that MGL already supports content markup such as OpenMath, which is used as the language-independent representation and LaTeX.

4 Mathematics editors and readers

Ordinary screen readers, those that are distributed with operating systems, are usually able to read the interface of a computer program and often are also able to read textual documents. They work by converting a document to plain text and simply reading all the characters on the screen. This often leads to a less effective presentation than would be the case if the document structure were retained. It is however impossible for them to tackle mathematical expressions which are usually treated as images or interactive fields in a web page.

Reading and editing mathematical text is central to any enhancement of software or ICT for teaching and learning. This includes the synchronicity of multimodal display, both with ready content and dynamically while editing.

Reading mathematics is not a linear process, but requires the focus of attention to shift between parts of current expressions and previous expressions. Hence mathematics editors and associated displays usually allow navigation, selection, zooming, and enlargement. While the primary editing tool traditionally employs a mouse/keyboard combination, voice-driven input can be the better choice for a motor impaired user. Multimodal rendering of mathematics could thus be defined to include the visual/aural display of the formulae while at the same time connecting to a Braille tablet for those users that need a transcription in linear form. Used in a classroom, the multimodal synchronicity will ensure that all students, together with the teacher, may set their own preferred mode of display and participate in the manner most effective to their needs.

Fire Vox (Fire Vox: A Screen REading Extension for Firefox) is a screen reader plugin for the web browser Firefox designed to accommodate different needs. Commands are keyboard activated and a highlighting feature allows following the part of the screen that is being read. Most important is the support of MathML using Nemeth's MATHSPEAK rules[10]. Fire Vox also supports the CSS3 Speech Module and interactivity in Live Regions marked up in WAI-ARIA, Accessible Rich Internet Applications [8], defining roles, states and properties of a web interface so that assistive technologies can make use of better semantics to provide alternative modalities of access. Associated with aural presentation, it can define, for instance, vocal input cues whenever an input field is marked to have role input.

We are evaluating a number of editors in order to choose the most flexible formats and design a solution which is portable. We list a few systems here below, without the claim of being exhaustive.

[10] `http://people.rit.edu/easi/easisem/talkmath.htm`

LAMBDA[11] [9] is a mathematical editor designed for the educational needs of both students and teachers, especially in the school and university environment. It is based on the Lambda linear code, a specific code which was designed by the LAMBDA project team. This code has localizations to several European languages, which tend to use the same Braille symbols than Braille national codes, while its structure is specific. Expressions in LAMBDA code are displayed on a Braille display, while they are displayed on the screen with a special font, which symbols are easy to memorize. It is convertible into MathML, aynchronously, in a separate window. Voice synthesis can either speak the name of an element or read an entire formula in a natural language. The LAMBDA software could use GF to allow natural language presentation of the mathematical formulas and to add localized language versions that are not yet included.

The WebALT project produced its own editor, the TextMathEditor. The natural extension of this work is the addition of aural input and output plus those formats needed for Braille peripherals. The editor uses 2-dimensional palette based formula editing, has predictive typing for the semantically correct expressions and run-time display of the natural language presentation in any of the languages or formats supported.

MaWEn, Mathematical Working Environment for the Blind, [21] is a prototype application developed as part of the MICOLE project[12]. It is a comprehensive, collaborative, bi-modal software solution designed to address the issues of working on documents of mixed content - textual and mathematical. It uses MathML presentation together with the UMCL conversion library to enable the localized representation of formulas in the Braille mathematics code of the user's choice and, synchronously, in a natural visual rendering. GF here could be used to add speech rendering and multilingualism.

We believe that GF could provide web services to any major software system, from MathsGenie [13] to Infty [23], that would add multilinguality and multimodality. Of particular interest is the InftyEditor, as it is part of an integrated OCR software suite for mathematical documents. Its OCR functionality is a step forward in supporting scholars and students with visual disabilities to read and write scientific papers.

5 Mathematical environments

The solution we propose should not be limited to editors and readers/players of mathematical content but also be adoptable by leading educational mathematics software systems. We will be targeting in particular STACK[13], a computer aided assessment system for mathematics used for both formative and summative assessment. Currently STACK's multilingual capabilities are limited to the primitive substitution of fixed strings from a look-up table. English and Dutch are supported. Mathematical questions have to be translated individually to each new language: a prohibitively expensive task. Here a "question" signifies the entire item, including the question itself, potential feedback, and fully worked model solutions. The STACK system should be able to accept mathematical input in the OpenMath markup language, and provide output in this format; and to interface with the automatic translation tools developed elsewhere. This includes not only the front end, as seen by students, but also STACK question authoring tools for the teacher.

Mathematical assessment systems, such as STACK, are well suited for sciences in general. Almost any discipline needs to deal with mathematical formulae, and systems providing automatically graded

[11]LAMBDA: Linear Access to Mathematic for Braille Device and Audio synthesis (EU funded project IST-37139), `http://www.lambdaproject.org`

[12]MICOLE: Multimodal Environment for Inclusion of Visually Impaired Children (2004-2007, EU funded project), `http://micole.cs.uta.fi`

[13]`http://www.stack.bham.ac.uk`

homework, quizzes and examinations need to be able to deal with mathematics (to read it is a minimum requirement). Such systems, enhanced with multilingual aural rendering, will have an impact in all areas of education, not only eScience.

5.1 Voice-driven Computations in Sage

This is a work in progress prototype to show how the *Mathematical Grammar Library* developed in MOLTO[14] can be used to interact in natural language with Computer Algebra Systems (CAS) (specifically with Sage).

The system consists on a command-line interface written in the Haskell programming language that talks to the notebook server that is embedded in every Sage installation. This server accepts HTTP requests containing Sage commands as strings and returns a JSON-encoded block. We expect this block to be delivered immediately, with a status flag saying if the computation has already been carried out or, when the computation is lengthy, that the system is still working on it.

An abstract grammar generating all the dialog queries and answers sits at the center of the system. It takes the form of a GF[18] grammar and acts as an interlingua between a given natural language and the specific CAS language.

A user query in natural language is *parsed* by GF into the abstract representation and then converted (*linearized* in GF parlance) into a expression understandable by the the underlying CAS. This is evaluated by the CAS and the returned expression parsed in the reverse direction into an abstract representation again. At this moment we could convert formulas into their textual equivalents or paraphrase the answer to make it less stilted. For example, the system could convert something which would have been rendered as *"the second root of x"* into the more palatable: *"the square root of x"*.

The final step is to convert the abstract representation of the answer into a natural language sentence to be output to the user. Notice that we are not forced to use the same natural language in which the query was stated.

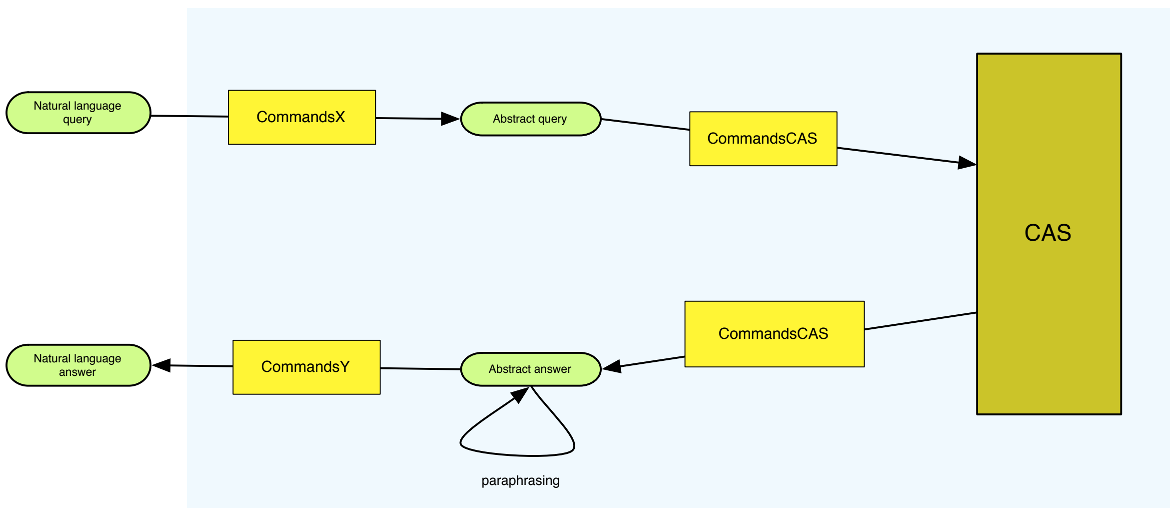

Figure 1: Basic query/answer interaction: `Commands`L denote the concrete grammar for language L. `CommandsCAS` is the concrete grammar for the CAS.

[14]MOLTO: Multilingual Online Translation (EU funded project FP7-ICT-247914), `http://www.molto-project.eu`

For all this to work one has to provide an *abstract* GF grammar as mentioned before and *concrete* grammars for each of the natural languages targeted and the CAS language. This is not so difficult task because the linguistic details of the implementation are hidden in the *resource library* of GF. The CAS concrete, being an artificial language, is much easier.

The present prototype can be built within several machine architectures and operating systems and for the MacOS X case, it uses the built-in text-to-speech facility to provide aural output of the calculation answer.

6 Final remarks

Traditionally, aural rendering often means text-to-speech generation (TTS). Many problems inherent to TTS can be avoided by the GF-based approach described in this paper since we can use grammar-to-speech (GTS). For instance, TTS usually has to perform some syntactic parsing to pronounce homographs correctly (*"Did you record that record?"*) and to establish the correct intonation. Sometimes these problems have no solution because the input text is genuinely ambiguous. In GTS, the starting point is a grammatical structure that contains enough information to solve both problems. How to implement this approach in GF was to some extent investigated in the TALK project (TALK, 2004-2006), in the task of generating system responses in dialogue systems. But much remains to be done – in particular, to enrich the RGL with different topical structures requiring different prosodic contours. To some extent, spoken-language versions of grammars must be written separately from written-language versions, which in GF means separate concrete syntax modules. Yet another task is to study the speech rendering of mathematical formulas, which in ordinary text are represented symbolically. The solutions created for a localized language in some of the tools we discussed should permit a generalization to language-independent GF solutions.

Acknowledgement

This paper was written with the support of the MOLTO project, co-funded by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n^o FP7-ICT-247914[15].

References

[1] D. Archambault and D. Burger. The Vickie Project. In K. Miesenberger, J. Klaus, and W. Zagler, editors, *Proceedings of ICCHP 2002 (International Conference on Computers Helping People with Special Needs)*, volume 2398 of *LNCS*, pages 90–97, Linz, Austria, 2002. Springer.

[2] D. Archambault, D. Fitzpatrick, G. Gupta, A. I. Karshmer, K. Miesenberger, and E. Pontelli. Towards a Universal Maths Conversion Library. In K. Miesenberger, J. Klaus, W. Zagler, and D. Burger, editors, *Proc. ICCHP 2004 (9th International Conference on Computers Helping People with Special Needs)*, volume 3118 of *LNCS*, pages 664–669, Berlin, 2004. Springer.

[3] D. Archambault and F. Guyon. UMCL Transcription Tools. In G. J. Gelderblom, M. Soede, L. Adriaens, and K. Miesenberger, editors, *Proc. of AAATE 2009 (Everyday Technology for Independence and Care - AAATE 2011)*, pages 416–423, Maastricht, The Netherlands, 2011.

[4] R. Ausbrooks, S. Buswell, D. Carlisle, G. Chavchanidze, S. Dalmas, S. Devitt, A. Diaz, S. Dooley, R. Hunter, P. Ion, M. Kohlhase, A. Lazrek, P. Libbrecht, B. Miller, R. Miner, C. Rowley, M. Sargent, B. Smith, N. Soiffer, R. Sutor, and S. Watt. Mathematical Markup Language (MathML) Version 3.0. Technical report, W3C Recommendation, 2010.

[15] The contents of this paper is the sole responsibility of the authors an in no way represents the views of the European Commission or its services

[5] E. Bates and D. Fitzpatrick. Spoken mathematics using prosody, earcons and spearcons. In K. Miesenberger, J. Klaus, W. Zagler, and A. Karshmer, editors, *Computers Helping People with Special Needs*, pages 407–414. Springer Berlin / Heidelberg, 2010.

[6] S. Buswell, O. Caprotti, D. Carlisle, M. Dewar, M. Gaetano, and M. Kohlhase. The OpenMath standard. Technical report, version 2.0. Technical report, The Open Math Society, 2004. http://www. openmath. org/standard/om20, 2004.

[7] O. Caprotti, W. Ng'ang'a, and M. Seppälä. Multilingual technology for teaching mathematics. *Proceeding of the International ...*, 2005.

[8] J. Craig and M. Cooper. Accessible Rich Internet Applications (WAI-ARIA) 1.0, 2011.

[9] A. D. N. Edwards, H. McCartney, and F. Fogarolo. Lambda:. In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility - Assets '06*, page 48, New York, New York, USA, Oct. 2006. ACM Press.

[10] D. Fitzpatrick. Speaking Technical Documents: Using Prosody to Convey Textual and Mathematical Material. In K. Miesenberger, J. Klaus, and W. Zagler, editors, *Proc. ICCHP 2002 (International Conference on Computers Helping People with Special Needs)*, volume 2398 of *LNCS*, pages 491–498, Linz, Austria, July 2002. Springer.

[11] D. Fitzpatrick. Mathematics: How and What to Speak. In K. Miesenberger, J. Klaus, W. Zagler, and A. I. Karshmer, editors, *Proc. ICCHP 2006 (10th International Conference on Computers Helping People with Special Needs)*, volume 4061 of *LNCS*, pages 1199–1206, Linz, Austria, July 2006. Springer.

[12] F. Helder and F. Diamantino. Audio Rendering of Mathematical Formulae using MathML and AudioMath. . In *UI4ALL2004*, Vienna, Austria, 2004.

[13] A. Karshmer, G. Guta, and D. Gillan. Architecting an Auditory Browser for Navigating Mathematical Expressions. In K. Miesenberger, J. Klaus, and W. Zagler, editors, *Proc. ICCHP 2002 (International Conference on Computers Helping People with Special Needs)*, volume 2398 of *LNCS*, pages 477–485, Linz, Austria, July 2002. Springer.

[14] R. Moreno. Who Learns Best with Multiple Representations? Cognitive Theory Predictions on Individual Differences in Multimedia Learning. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002*, number 1, pages 1380 – 1385, Denver, Colorado, 2002. Association for the Advancement of Computing in Education (AACE), P.O. Box 3728, Norfolk, VA 23514. Tel: 757-623-7588; e-mail: info@aace.org; Web site: http://www.aace.org/DL/.

[15] M. O'Malley, D. Kloker, and B. Dara-Abrams. Recovering parentheses from spoken algebraic expressions . *IEEE Transactions on Audio and Electroacoustics*, 21(3):217–220, 1973.

[16] N. M. A. Panel. Foundations for success: The final report of the National Mathematics Advisory Panel. Technical report, U.S. Department of Education, 2008.

[17] T. V. Raman. *Audio Systems for Technical Reading*. PhD thesis, Department of Computer Science, Cornell University, NY, USA, May 1994.

[18] A. Ranta. GF: A Multilingual Grammar Formalism. *Language and Linguistics Compass*, 3(5):1242–1265, Sept. 2009.

[19] A. Ranta. *Grammatical Framework: Programming with Multilingual Grammars*. Center for the Study of Language and Information - Studies in Computational Linguistics, 2011.

[20] R. Stevens and A. Edwards. A sound interface to algebra . In *IEE Colloquium on Special Needs and the Interface*, pages 511–513, 1993.

[21] B. Stöger, K. Miesenberger, and M. Batušić. Mathematical working environment for the blind motivation and basic ideas. *Computers Helping People with Special Needs*, pages 627–627, 2004.

[22] L. A. Streeter. Acoustic determinants of phrase boundary perception. *Journal of the Acoustical Society of America*, 64(6):1582–1592, 1978.

[23] M. Suzuki, T. Kanahori, N. Ohtake, and K. Yamaguchi. An Integrated OCR Software for Mathematical Documents and Its Output with Accessibility. In K. Miesenberger, J. Klaus, W. Zagler, and D. Burger, editors, *Proc. ICCHP 2004 (9th International Conference on Computers Helping People with Special Needs)*, volume 3118 of *LNCS*, pages 648–655, Berlin, July 2004. Springer.

[24] D. Weck. CSS Speech Module. Technical report, W3C, 2011.