

<b>Institution: University of Birmingham</b>
<b>Unit of Assessment: UoA 11 - Computer Science and Informatics</b>
<b>Title of case study: Providing Accessibility to Scientific Documents to Visually Impaired Readers via Mathematical Formula Recognition</b>
<p><b>1. Summary of the impact</b> (indicative maximum 100 words)</p> <p>Our research is concerned with enabling access to mathematical literature to users with visual impairments (i.e., blind or partially sighted users) or print impairments (i.e., users with specific learning disabilities like dyslexia or dysgraphia). Therefore the impact is primarily of a societal nature: we enable visual and print impaired learners' access to scientific and mathematical knowledge from which they were previously excluded, thereby furthering an inclusive teaching and learning environment. With the number of people with learning disabilities being over 1 million (<a href="http://www.learningdisabilities.org.uk/help-information/Learning-Disability-Statistics-/">http://www.learningdisabilities.org.uk/help-information/Learning-Disability-Statistics-/</a>) and the number of visually impaired people predicted to rise over 2 million by 2020 in the UK alone (<a href="http://www.rnib.org.uk/aboutus/Research/statistics/Pages/statistics.aspx">http://www.rnib.org.uk/aboutus/Research/statistics/Pages/statistics.aspx</a>), the work is significant in providing equal opportunities to learners in the STEM subjects.</p> <p>Our research has led to work with Google Inc. to enhance mathematics accessibility on the Web via the screen-reader ChromeVox. It enables the full text to speech translation of mathematics on the Web <b>for all users of the Chrome browser and Android platforms</b> and has been included in ChromeVox since version 27, released 9/5/2013, which has 31,518 downloads from the Chrome store as of 11/10/2013.</p> <p>Our work has also resulted in an assistive technology tool, called MaxTract, for providing access to teaching and learning material. It has been deployed within digital mathematics libraries to enhance accessibility to online material. Through direct feedback we are aware of a number of visually impaired users that have actively used our tool.</p>
<p><b>2. Underpinning research</b> (indicative maximum 500 words)</p> <p>The research for this case study is in scientific document analysis, especially in mathematical formula recognition. Its aim is to find and analyse mathematical expressions in documents in order to transform them into a format that can be machine analysed and processed. While the recognition of regular text, e.g. via optical character recognition from scanned documents, is fairly routine, the recognition of formulae is a highly non-trivial task due to the complex two dimensional layout of mathematical expressions, the importance of special attributes, like accents and fonts, and the lack of dictionaries usable for correction of recognition results.</p> <p>We have developed a linear grammar that processes streams of characters with associated spatial information — in the form of character bounding boxes and their absolute positioning on a page — and parses it into a two dimensional parse tree representation reflecting the two dimensional structure of formulas. The grammar employs constraints on the relative positioning and size of neighbouring characters and sets of characters in order to construct two dimensional formula layout. This makes the recognition process entirely independent of any information on the actual characters and purely a heuristic function of spatial character distribution. The resulting parse trees are an effective data structure for translation into diverse mathematical markup formats, such as LaTeX and (presentation) MathML, but also allow for a description in natural language [1].</p> <p>As a further refinement we have improved the interpretation of the parse tree into two dimensional formulas by introducing some semantic interpretation of its components. Specifically we are taking information on separation distance of character sequences and character fonts into account when</p>

**Impact case study (REF3b)**

constructing formulas and clustering component subformulas [2].

Our work has led to the development of a novel tool, MaxTract, which enables the extraction of content from PDF documents. Unlike comparable tools that perform PDF re-engineering, MaxTract does not treat mathematical content as noise and omit it. Instead, mathematical expressions are analysed using our grammatical approach and translated into processable formats using modular output drivers, which allows easy customisation of the system for different application domains. In particular, MaxTract can generate standard mathematical markup formats (LaTeX, presentation MathML) and as description in natural language, which can be given directly to speech engines. In comparison with a state-of-the-art approach to analysing mathematical PDF documents [3], MaxTract has been shown to be the most reliable tool to re-engineer PDF documents.

The work on formula recognition and accessibility led to an invitation to Sorge by TV Raman, the head of the accessibility engineering group at Google, to apply for a visiting faculty project with Google for a sabbatical in the 2012/13 academic year. This project was accepted and Sorge worked with the ChromeVox team at Google to provide text-to-speech translation of mathematics in a fully-fledged web-based screen-reader. In his year at Google Sorge developed a rule based approach to speaking mathematical expressions on the web that are either given as MathML markup or rendered via MathJax. He implemented a speech rule engine that is now embedded as a core feature of ChromeVox, exposing a public API for easy customisation of speech rules by authors of websites. In addition his framework enables the speech translation of image-based maths expressions with alternative text markup, making maths on websites like Wikipedia or Mathworld fully accessible.

Furthermore, Sorge developed and implemented an approach to allow visually impaired readers to interactively engage with mathematical content by systematically exploring subexpressions, as well as a novel transformation of presentation MathML markup into semantic interpretation, which was continuation of work in [2,4] that is geared towards a more natural reading experience for K-12 and undergraduate mathematics. These were included in the ChromeVox version 30 and 31 releases.

The Birmingham research team led by Sorge (Senior Lecturer) includes Josef Baker (2006-13), PhD student and then RA, and Alan Sexton (2006-12), Lecturer.

**3. References to the research** (indicative maximum of six references)

The key publication for our research is [1], which established the grammatical approach leading to a flexible representation of mathematical expressions that was central to the rest of the research. [2] and [4] contain the initial steps towards semantic analysis of formulas and layout reconstruction, which were significantly expanded by Sorge's work at Google and which were released with ChromeVox version 31.

**Publications**

[1] J. Baker, A. Sexton, and V. Sorge. A linear grammar approach to mathematical formula recognition from PDF. In J. Carette, et al (eds), Intelligent Computer Mathematics — Joint Proceedings of Calculemus 2009 and MKM 2009, volume 5625 of LNAI, Springer Verlag, Berlin, Germany. (MKM Best Paper Award) doi: 10.1007/978-3-642-02614-0\_19

[2] J. Baker, A. Sexton, and V. Sorge. Faithful mathematical formula recognition from pdf documents. In 9th IAPR International Workshop on Document Analysis Systems, pages 485–492, Boston, USA, June 9–11 2010. ACM Press. doi: 10.1145/1815330.1815393

[3] J. Baker, A. Sexton, V. Sorge, and Masakazu Suzuki. Comparing approaches to mathematical document analysis from PDF. In International Conference in Document Analysis and Recognition, pages 463–467, Beijing, China, 2011. IEEE Computer Society Press, Los Alamitos, CA, USA. doi: <http://dx.doi.org/10.1109/ICDAR.2011.99>

[4] J. B. Baker, A. P. Sexton, and V. Sorge. Towards reverse engineering of PDF documents. In P.

## Impact case study (REF3b)

Sojka and T. Bouche, editors, Towards a Digital Mathematics Library, DML 2011, pages 65–75, Bertinoro, Italy, July 2011. Masaryk University Press. ISBN 978-80-210-5542-1. available from: <http://www.cs.bham.ac.uk/~jbb/documents/dml11.pdf>

**Projects**

[5] European digital mathematics library. EU CIP-ICT Grant, 1<sup>st</sup> February 2009 – 31<sup>st</sup> January 2013. Euros 373,160. PI: Volker Sorge, CIs: Alan Sexton, Mark Lee, RA: Josef Baker

[6] Improving accessibility to mathematical teaching resources. JISC OER Rapid Innovation Grant, 1<sup>st</sup> April – 30<sup>th</sup> September 2012. £22,460. PI: Volker Sorge, RA: Josef Baker

**4. Details of the impact** (indicative maximum 750 words)

The impact of our research is of a societal nature, by enabling visually and print impaired users' access to scientific and mathematical literature from which they were previously excluded. Making scientific and teaching material accessible for visually impaired students is essential in an inclusive teaching and learning environment. However, it is still very difficult and expensive to make mathematical documents accessible, making this a major obstacle for visually impaired learners wanting to pursue subjects such as Mathematics, Physics, or Computer Science in both further and higher education. As a consequence, our research provides a novel means to further inclusion of the visually impaired in the mathematical sciences.

**Text to Speech for Mathematics, Google, Inc., Mountain View, USA.** From September 2012 to September 2013 Sorge spent a sabbatical year working with the accessibility engineering group at Google, Inc. to enable full text to speech translation of mathematics with the ChromeVox screen-reader. The initial release of the work has been launched at Google IO in May 2013 [7] and a further release in September 2013 provides enhanced semantic interpretation and an API for customisation of the rule engine [8]. ChromeVox is open source [9] and a web-based screen-reader for Chrome on all platforms. As of 11/10/2013 it has been downloaded 31,518 times from the ChromeStore [10]. ChromeVox is also the screen reading solution for web content on Android devices via WebViews. While there have been over 1 billion Android activations, the exact number of users of accessibility services on these devices is confidential.

Furthermore, the WebViews API is also used by ePub3 readers to make content accessible. An example is the Ideal Group Reader [11] which concentrates on STEM material that has been installed in beta around 35,000 times as of 9<sup>th</sup> July 2013.

**From PDF to Accessible Mathematics via MaxTract.** MaxTract is available as a free web service at <http://www.cs.bham.ac.uk/research/groupings/reasoning/sdag/submit.php>. It allows users to upload PDF documents and get the document in an accessible format back. We have also worked directly with users via email support and at specialist workshops on e-inclusion for sciences and mathematics. To obtain feedback to both enhance the system and assess its impact we have conducted a case study with a number of blind students and researchers at Johannes Kepler University, Linz, Austria and the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, where two of the foremost European Centres for Integrated Studies are located. The following are two quotes from questionnaire replies:

*“Before MaxTract, if I wanted to access a PDF document with mathematical contents I needed to hire a person who scanned the document and wrote the mathematical formulae into it by hand. This is now eliminated - I have digital access, although not yet optimal one, which eliminates the expense of hiring a person to do the work mentioned.”*

*“It is faster. As I already said, without MaxTract other people are involved for me to receive accessible material. [...] Using MaxTract does convert when I submit a document within a few minutes.*

*It is cheaper - not for me in person but for the university. [...] Using MaxTract is free. And it is*

## Impact case study (REF3b)

*easier. I submit a document and after a few minutes I get a converted one back. No-one has to be asked, no checking for errors, no waiting.” [sources 12-16]*

In addition, we are deploying the results of our research in knowledge transfer projects, which demonstrate the impact of our work outside the immediate scientific community.

**European Digital Mathematics Library [5]** This is a project under the CIP-ICT programme (a programme which is explicit about being for knowledge transfer only, not allowing for new research in a project) with the aim of assembling the European Digital Mathematics Library (EuDML) that offers working mathematicians a central portal to access a large collection of mathematical literature as well as provides functionality that is specific to mathematics. We were a technology provider in the project, integrating MaxTract into the EuDML workflow for enhancement of over 5000 documents to enable advanced search features on mathematical formulas to support research mathematicians and engineers and provide accessibility for some of the content for visually impaired users [4].

**Improving Accessibility to Mathematical Teaching Resources [6]** was a JISC OER Rapid Innovation grant, classified as teaching only, with the aim to develop MaxTract from the research prototype into an assistive technology tool that meets the identifiable needs of accessibility support practitioners for providing access to mathematical teaching material in higher and further education. The resulting assistive technology tool has been successfully used in the OERPUB project (<http://kefletcher.blogspot.com/>).

##### 5. Sources to corroborate the impact (indicative maximum of 10 references)

[7] Google I/O 2013 - Advancing Web Accessibility with ChromeVox, showing Sorge explaining ChromeVox along with Google engineers Charles Chen and David Tseng. Accessed on 11/10/2013. <http://www.youtube.com/watch?v=YyWu9HB9QtU>

[8] Google Developers Live - Spoken Mathematics on the Web, showing Sorge explaining the ChromeVox API for Mathematics along with Google engineer David Tseng. Accessed on 11/10/2013. <https://developers.google.com/live/shows/5881057312243712>.

[9] Google open source repository of ChromeVox exhibits Sorge's authorship behind the mathematics part. Accessed on 11/10/2013. E.g., [https://code.google.com/p/google-axe-chrome/source/browse/trunk/chromevox/speech\\_rules/speech\\_rule\\_engine.js](https://code.google.com/p/google-axe-chrome/source/browse/trunk/chromevox/speech_rules/speech_rule_engine.js)

[10] <https://chrome.google.com/webstore/detail/chromevox/kgejglhpjiefppelpmljglcjbhoiplfn>

[11] IDEAL Group Reader® Beta II, <https://play.google.com/store/apps/details?id=org.easyaccess.epubreader> (Accessed on 31/10/2013)

[12] ██████████. Answers to Questionnaire, 9 March 2013.

[13] ██████████. Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany. Answers to Questionnaire, 26 March 2013.

[14] ██████████. Vienna University of Technology, Vienna, Austria. Answers to Questionnaire, 15 March 2013.

[15] ██████████. Johannes Kepler University, Linz, Austria. Answers to Questionnaire, 29 March 2013.

[16] ██████████. Johannes Kepler University, Linz, Austria. Answers to Questionnaire, 18 March 2013.