RAVI: Reading Assistant for Visually Impaired

Akashdeep Bansal, Indian Institute of Technology Delhi, India akashdeep@cse.iitd.ac.in
Himanshu Garg, Indian Institute of Technology Delhi, India himanshu.garg@cse.iitd.ac.in
Neha Jadhav, Indian Institute of Technology Delhi, India neha.manik.jadhav@cse.iitd.ac.in
Sunil Kumar, Indian Institute of Technology Delhi, India grukul.sunil@gmail.com
M Balakrishnan, Indian Institute of Technology Delhi, India mbala@cse.iitd.ac.in

There are over 285 million persons with visual impairments in the world [1]. Technologies like Screen reader and Refreshable braille display provide them significant access to the world of digital content. Unfortunately, PDFs, one of the most common formats used for digital publishing, poses significant challenges in accessing the eBooks. Non-unicode encoding and absence of the markup or metadata for various structural elements are the key reasons behind the inaccessibility of the PDF document. In STEM content, the challenge gets worse due to the absence of syntactic and semantic information related to equations and tables. Similarly, Diagrams that are important for understanding the underlying concept are inaccessible due to the absence of alt-text, sonification and/or an interactive interface [2, 3, 4, 5], etc. Therefore, Inaccessibility of the study material becomes a hindrance in pursuing STEM (Science, Technology, Engineering, and Mathematics) career options. These challenges increase manifold in Indian context with diverse languages, associated fonts and multiple authorities who prepare and approve textbooks for education.

In RAVI, a project at AssisTech lab, Indian Institute of Technology Delhi, India, we are working towards developing tools for automatically converting inaccessible digital books into standard compliant EPUB. RAVI aims to provide solutions in the following four areas -

1. Automated tool for converting inaccessible STEM documents into a fully accessible format such as EPUB

In this module, we are aiming at developing a tool which can automatically convert a PDF document into a standard compliant EPUB. The key reasons behind choosing the PDF document are: most of the document formats can be easily converted into PDF document and PDF is one of the most common formats used for digital publishing. There are several tools [6,7,8,9] that have been developed for making study material accessible. We didn't find any of them working comprehensively for various elements of STEM eBook, implying still a manual intervention is required.

In our approach, we are parsing PDF to first convert digitally generated PDF into raw HTML. The HTML file is then analyzed to remove non-linearity due to different mathematical constructs such as superscript, subscript, etc. by linearizing them. The resultant file is passed through a

classifier to classify the content into text, equations, and text inside equations. This enables us to treat text and equations separately and adapt the document structure analysis algorithms accordingly. Further, the document is processed to extract various structural elements like paragraph, list, table, caption, header, footer, page number, heading, and diagram. The analysis is performed on the basis of the visual parameters such as left and right margin of line, line spacing, character position, font size, color, and style. The rules are similar to how a sighted person identifies them visually. Finally, the extracted information is encoded according to the EPUB standards.

2. Automatic/Semi-automatic approaches for making diagrams accessible with the screen reader

In this module, we are aiming at making STEM diagrams accessible through a non-visual and non-tactile approach. In the past, several approaches are proposed to make STEM diagrams accessible by providing alt-text, hierarchical alt-text [2, 10], sonification [11, 12], haptic [13, 14, 15], interactive [2, 3, 4, 5], etc. We didn't find any tool which can take raster image of STEM diagrams as input and output it with any of the above-proposed formats as output.

In our approach, we are classifying the images into various categories like experimental setup, geometry, histogram, pie-chart, ray-diagram, biological, etc. Further, we are working towards developing category specific modules, which will generate the interface which works best for that particular category of diagrams. To understand the best interface for a specific category, we are analyzing the different kinds of interfaces suggested in the literature for each and every category of diagrams.

We have also developed a web application for crowdsourcing of image description. In this approach, each image is passed through three independent volunteers followed by a reviewer. To capture the context, we are also showing relevant pages of the document to the volunteers.

3. Improving the user experience of reading equations with screen readers

In this module, we are aiming to improve the user experience of reading equations with screen reader by optimizing the verbosity and naturalness and minimizing the ambiguity and cognitive load. These are accomplished on the basis of semantics and complexity level of the equation, users' age profile and familiarity level, etc.

4. Scripts for converting Indian language text from non-unicode encodings to unicode encoding

In this module, we aim to improve accessibility of eBooks in Indian languages. Many Indian print publishers have continued to use traditional true type fonts that don't support unicode encoding. eBooks that are created out of these texts are inaccessible due to use of glyph based

encodings. To address this, we have developed scripts that convert the text in non-unicode encoded fonts to unicode encoded fonts within InDesign. Using these scripts more than 120 textbooks of India's National Council of Educational Research and Training (NCERT) have been successfully converted.

The scope of this work is not limited to making STEM eBooks accessible for persons with visual impairment. It also has importance in enhancing the performance of search engines and improving the reading experience of the sighted people. It enhances the search engine performance by providing access to syntactic and semantic information for the equations, the semantics of the diagrams, and regional language text in non-unicode encodings. It Improves the reading experience of sighted persons by reflowing the content of the eBook, the need for which is increasing due to increasing use of small screen devices like mobile, kindle, and tablet, etc. Further, it also enhances the reusability of the content.

Talk detail

In our talk, we will discuss our approach and the reasons behind its adoption. We will also discuss various challenges which we have encountered. Finally, we will discuss how we have overcome some of those challenges and our plans for others.

Acknowledgment

We would like to thanks our funding agencies Ministry of Human Resource and Development (MHRD) and Department of Atomic Energy (DAE), Government of India.

References

1. “Blindness and Visual Impairment Fact Sheet N 282,” August 2014, World Health Organization, http://www.who.int/en/news-room/fact-sheets/detail/blindness-and-visual-impairment
2. Sorge, Volker, Mark Lee, and Sandy Wilkinson. "End-to-end solution for accessible chemical diagrams." *Proceedings of the 12th Web for All Conference*. ACM, 2015.
3. Fitzpatrick, Donal, A. Jonathan R. Godfrey, and Volker Sorge. "Producing Accessible Statistics Diagrams in R." *Proceedings of the 14th Web for All Conference on The Future of Accessible Work*. ACM, 2017.
4. Bahram, Sina. "Multimodal eyes-free exploration of maps: TIKISI for maps." *ACM SIGACCESS Accessibility and Computing* 106 (2013): 3-11.
5. Siddhartha Gupta, Manshul V Belani, Dinesh Kaushal, and M. Balakrishnan, Microsoft Excel ChartsTM accessibility: An Affordable and Effective Solution, In Proceedings of the 3rd International Workshop on Digitization and E-Inclusion in Mathematics and Science, 2016.
6. Baker, Josef B., Alan P. Sexton, and Volker Sorge. "MaxTract: Converting PDF to LaTeX, MathML and Text." *International Conference on Intelligent Computer Mathematics*. Springer, Berlin, Heidelberg, 2012.

7. Suzuki, Masakazu, et al. "An integrated OCR software for mathematical documents and its output with accessibility." *International Conference on Computers for Handicapped Persons*. Springer, Berlin, Heidelberg, 2004.
8. ABBYY. 2017. ABBYY FineReader 14.0. https://www.abbyy.com/en-eu/finereader/
9. MathPix. 2018. MathPix: Converting Images to L A TEX. https://mathpix.com/
10. Vítor Carvalho and Diamantino Freitas. 2015. Automatic description of SVG images for the visually impaired: a Gestaltic approach. Procedia Computer Science 67 (2015), 2–11.
11. NASA. 2008. MathTrax. https://prime.jsc.nasa.gov/mathtrax/
12. John A Gardner and Vladimir Bulatov. 2006. Scientific diagrams made easy with IVEO T M . In International Conference on Computers for Handicapped Persons. Springer, 1243–1250.
13. Ford Europe. 2018. Feel The View – A smart window for blind passengers. Video. http://fordeurope.blogspot.com/2018/04/feel-view-smart-window-for-blind.html
14. Michael Curran. [n. d.]. AudioScreen – An add-on for the NVDA. https://github.com/nvaccess/audioScreen
15. Matt Calder, Robert F Cohen, Jessica Lanzoni, and Yun Xu. 2006. PLUMB:: an interface for users who are blind to display, create, and modify graphs. In Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility. ACM, 263–264.