odt2braille brings Braille to your Office

Bert Frees, Christophe Strobbe & Jan Engelen*
Katholieke Universiteit Leuven
Kasteelpark Arenberg 10
3001 Heverlee-Leuven
Belgium

Abstract
OpenOffice.org, the free and open-source productivity suite, is one step closer to becoming the ultimate authoring environment for accessible and high-standard documents. Thanks to odt2braille, an OpenOffice.org extension developed at the Katholieke Universiteit Leuven (Belgium), Braille is now available to anyone who knows how to use a word processor. odt2braille is powered by the open-source Braille translation engine Liblouis. Embosser communication and conversion to various Braille file formats is handled by the open-source Braille Utils API. The work described in this paper has been performed in the framework of the European R&D project ÆGIS (ÆGIS, 2011).

1 Introduction

Braille production is an expensive and time-consuming business. Generally, dedicated Braille tools are used to convert regular print to Braille. These tools have grown from primitive Braille editors to specialised Braille authoring environments that provide automatic formatting and translation services. Still, often a lot of manual work is required.

The typical workflow of a Braille production system is a multi-stage process. This is illustrated in illustration 1 below. If a document is not available in an electronic text format, it has to be scanned first, then converted with optical character recognition (OCR), verified, edited and possibly tagged to increase its machine readability. Next, the text file is automatically translated using a separate Braille tool. If the file was previously enriched with markup, automatic Braille formatting can be performed as well.

The consequence of this multi-stage approach is that any changes to the master file require recreation of the derivative edition(s). This is especially disadvantageous if a lot of manual work is involved. Furthermore, any translation or layout issues arising from poor choices made in the original cannot simply be fixed in the derivative edition, or they will be lost with the next source update.
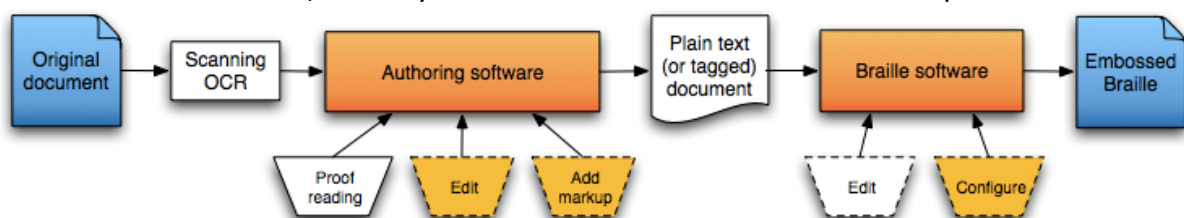
Illustration 1: Typical Braille production workflow: a multi-stage process

The next section discusses the state of the art in Braille production and sums up some remarkable open-source projects that are making efforts to reduce the cost of Braille production, to provide open translation systems and to make Braille production more open and available to individuals. The remainder of the paper focuses on odt2braille, an open-source project that tries to break down barriers by integrating the entire production workflow in a major office suite.

2 State of the Art and Related Work

2.1 Duxbury

Duxbury is the market leader in Braille editors. The commercial Duxbury Braille Translator (DBT) for Windows has been the standard for many years. DBT provides translation and formatting facilities to automate the process of conversion from regular print to Braille (and vice versa), and also provides word-processing facilities for working directly in the Braille as well as the print. Furthermore, it can import files from MS Word & WordPerfect, and it integrates nicely in MS Word. DBT supports almost all languages, Braille standards and embossers, and it is constantly updated. Of course all this comes with a relatively high price.

2.2 AutoBraille/NorBraille

AutoBraille and NorBraille are the result of collaborations between Sensus ApS, Synscenter Refsnæs and the national libraries for the blind in Denmark and Norway. The purpose of these projects is to automate the production of well-formatted, multi-volume Braille books, based on well-structured DAISY Talking Books. The Braille formatter processes the document based on options in a settings file and manages calls to the underlying RoboBraille/SB4 Braille translation and hyphenation engines.

The rationale behind these projects is: "while significant efforts are being invested in scanning, proofing, tagging and validating documents for production of DAISY Talking Books, the resulting source files are rarely used for other purposes. However, documents that have been tagged for DTB production may well contain all the information required for producing material in other alternative formats including Braille" (Christensen, 2010).

2.3 BrailleBlaster

A similar project is BrailleBlaster: a free, open-source, multi-platform and fully accessible Braille transcription software package (Boyer, 2011). BrailleBlaster is sponsored by ViewPlus Technologies and Abilitiessoft. While the main goal is to provide automatic transcription of DAISY files, several other XML files will be supported as well. Like DBT, BrailleBlaster will have a dual-window view: a "DAISY view" for regular text, with an interface similar to common word processors, and a WYSIWYG "Braille view" for working directly in the Braille. BrailleBlaster's native file format is Unified Tactile Document Markup Language (UTDML), which is DAISY 3.0 XML enhanced to include formatted Braille and tactile graphics.

2.4 Dotify

Dotify is an open-source Braille formatter/translator written in Java (Håkansson, 2011). The project is an initiative of the Swedish Library of Talking Books and Braille (TPB) and has its origin in the Braille In DAISY project[1].

The main objectives of Dotify are "to provide an open-source Braille translation system for Java, to define components for Braille translation systems sharing and

---

[1] The Braille In DAISY project was an attempt to include Braille in the DAISY Consortium family of formats (Håkansson, 2006).

to provide a proof of concept implementation for these components". The intention is not to compete with systems such as Liblouis but to agree on some Braille software components and/or system files and formats that can be shared between Braille implementations. Design-wise, Dotify wants to establish a more modular approach to Braille translation and formatting.

Related is the development of Open Braille Formatting Language (OBFL), a language inspired by XSL-FO designed to describe "how a Braille formatter should render a text, including margins, indentation, lists, tables and so on, but excluding any additional file format processing".

3 odt2braille

odt2braille builds on the idea of single-source publishing based on DAISY XML, but takes it a level higher. In searching for a more common and more general-purpose format that can be used as a truly accessible master format for deriving alternative formats including Braille and DAISY[2], the Open Document Format (ODF) has been selected. ODF is a widespread and well-established open format that can be edited by anyone using one of the many authoring applications that support it (while editing or creating DAISY books requires specialised tools with which the general public is not familiar). Moreover, anyone can make ODF documents accessible thanks to the emergence of accessibility tools.

odt2braille provides automatic Braille transcription of OpenDocument Text (ODT) files. Documents can be printed with a Braille embosser, or they can be exported to various Braille file formats. The user has full control over the Braille layout, just like he has over the normal print layout. odt2braille is an extension for Writer, the word-processing part of the office suites OpenOffice.org and LibreOffice. All functionality is embedded in the authoring environment itself. The goal of this approach is to make Braille available to anyone who knows how to use a word processor.

Illustration 2 shows the functionality that odt2braille adds to OpenOffice.org and illustrates how similar extensions can be used to create other alternative formats such as DAISY, large print, Bliss, etc.
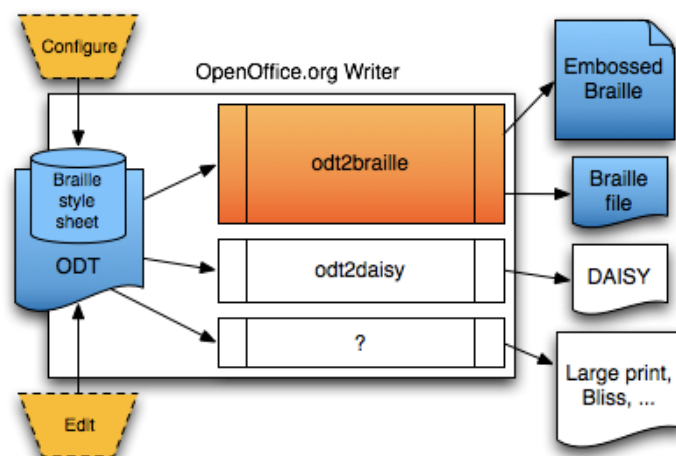
Illustration 2: Braille production workflow in OpenOffice.org

---

[2] Since the release of odt2daisy, ODF can also be used as a source for DAISY books (Spiewak, 2009).

3.1 Free and open-source

odt2braille is freely available on the web[3] as an extension for OpenOffice.org Writer and the more recent alternative (fork) LibreOffice Writer. Together they are the leading free and open-source office suites. odt2braille being freeware was a very important requirement. The goal is to reduce as much as possible the cost of Braille production, especially for teachers and other individuals.

3.2 Integration in a major office suite

A smooth integration into a major office suite is the main feature that makes odt2braille different from traditional Braille editors. Instead of creating a new and dedicated user interface that mimics some of the features of a word processor, odt2braille simply integrates in an existing mainstream word processor. As a result, users will generally have very little trouble learning to work with this tool, whether or not they are experienced in Braille. The combination of authoring and Braille transcription into a single stage also greatly simplifies the Braille production workflow, which is typically a multi-stage process.

3.3 Separation of content and style

odt2braille is designed with the principle of separate presentation and content in mind. A document can be presented in various forms: on paper (ink), as a web page, in Braille, as a DAISY book, etc. Because it is essentially the same document in each case, it makes sense to keep the content separate, and to provide different presentations on an on-demand basis using stylesheets, rather than duplicating the document for each alternative format.
In addition to the rich layout for ink-printing (for sighted users), a Writer document can now have basic Braille formatting. The Braille stylesheet is stored together with the content and other stylesheets in the ODT file (as metadata).

3.4 Large scope and flexibility

As a European initiative, odt2braille targets all of Europe and by extension the whole world. This is very ambitious because different countries have different Braille alphabets, rules and conventions on layout and mathematical Braille. Sometimes there are even several competing conventions within the same country. It makes flexibility one of the key objectives of the project.
odt2braille supports an impressive number of languages and allows a high degree of customisation of the Braille layout, which makes the software highly suitable for professionals. In order not to overwhelm the (inexperienced) user with preferences, documents can also be easily transcribed according to existing Braille formatting guidelines, requiring a minimum of user interaction. odt2braille has become a very powerful tool, but at the same time a lot of effort has been made to keep it understandable and user-friendly.

4 odt2braille in Detail

The overall architecture and the major components of odt2braille are presented in illustration 3. The conversion of an ODT file starts with a preprocessing step

---

[3] http://odt2braille.sourceforge.net

that transforms the original content and stylesheets to an alternative representation that is suitable as an input for the Braille formatter[4]. The formatter produces a Braille file in PEF-format, which can then be verified through a Braille preview and subsequently converted to other file formats, or printed on a Braille embosser. The Braille formatting is powered by Liblouisxml, and for the embosser communication and file conversion, odt2braille relies on the Braille Utils library. Apart from a graphical user interface in OpenOffice.org, odt2braille also provides an Ant task for automating bulk conversions of ODT files.
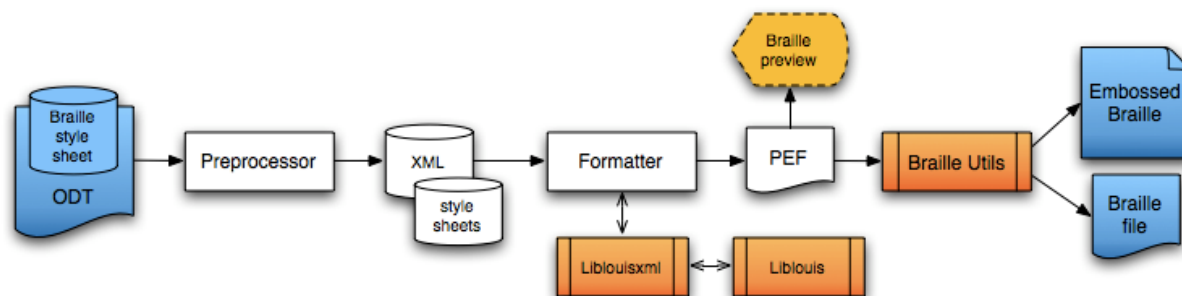
Illustration 3: odt2braille architecture and major components

4.1   Liblouisxml

Liblouisxml is the heart of the Braille conversion. It is an open-source library intended to provide complete Braille transcription services for XML documents (Liblouisxml, 2011). Liblouisxml is built on top of Liblouis, an open-source Braille translation and hyphenation engine (Liblouis, 2011). Liblouis is based on the translation routines in the BRLTTY screen reader for Linux. It supports contracted and uncontracted translation for over 40 languages. The translation is driven through translation tables written in a syntax that supports a rule- or dictionary-based approach. The Braille formatting in Liblouisxml is managed by semantic mappings that define how a specific XML element is to be rendered in Braille. Liblouisxml can also handle mathematical expressions (MathML). Both Liblouis and Liblouisxml were developed by John Boyer.

4.2   Braille Utils

Braille Utils provides a cross-platform API for embossing and converting Braille in Portable Embosser Format (PEF) (Braille Utils, 2011; Håkansson, 2009). PEF is a format for representing Braille documents in digital form, accurately and unambiguously. It's a digital Braille hardcopy - the PDF of Braille. It can be used for Braille embossing and archiving anywhere in the world. PEF uses markup (XML) rather than control characters to define structure, and uses Unicode Braille patterns rather than ASCII characters, because Unicode is locale-independent. PEF can also contain metadata.
Braille Utils provides conversion from PEF to a variety of other Braille file formats, and it handles communication with all kinds of embossers. Embosser types from Index Braille, Braillo, Interpoint, ViewPlus, CIDAT, Enabling Technologies

---

[4] The formatter is somewhat limited in the complexity of input formats it can handle. Also, it requires stylesheets to be prepared in a special syntax. The preprocessing step includes numbering of pages, headings and list items, linking of captions with tables and illustrations, and dividing the document into multiple volumes, each with their own table of content, front matter, note section, etc.

and Mountbatten are supported. This software is the work of Joel Håkansson and Magnus Karlströms from TPB, and it originated from the DAISY Pipeline[5].

5 Accessibility Checker

In order to create a valid and usable Braille document, authors first need to create an accessible source document. Users would greatly benefit from a tool that would support them in making their documents accessible and preparing them for Braille conversion, a tool that would highlight possible issues that could lead to unwanted artefacts in the Braille output.

The accessibility checker for OpenOffice.org and LibreOffice Writer developed in the framework of the ÆGIS project provides this kind of guidance. The checker is fully integrated in the office suite and makes checking the accessibility as straightforward as checking the spelling. Authors are being alerted to accessibility issues through errors and warnings. Upon request, issues are highlighted and repair suggestions are put forward. The tool also offers a few automatic and semi-automatic repair functions.

6 Future Work

There are a few things on our wish list: more extensive embosser support, more accurate Braille translations, and availability on more platforms.

We ask users to inform us about embossers that lack support, or give us feedback when support needs improving. The Braille Utils catalogue contains some embossers that have not been tested yet. Furthermore, Braille experts can help us improve the quality and accuracy of the Braille translations. Liblouis relies on contributions from volunteers and feedback from users. The library supports a large number of languages, but there is still room for improvement with regard to quality. We also wish to support more official Braille formatting guidelines. Finally, odt2braille is available on Windows and Debian, but porting to Mac OS X turned out to be more problematic than expected.

Acknowledgments

References

- ÆGIS (2011): *ÆGIS - Open Accessibility Everywhere: Groundwork, Infrastructure, Standards*. [2011.08.15: http://www.aegis-project.eu]
- Boyer, John (2011): *BrailleBlaster Specifications Draft*. [2011.08.15, BrailleBlaster: http://www.brailleblaster.org/bbspec01.html]
- Braille Utils (2011). [2011.08.15, Google: http://code.google.com/p/brailleutils]
- Christensen, Lars Ballieu (2010): *Exploiting the Daisy Format for Automated, Single-Source Braille Publishing*. In: Computers Helping People

---

[5] The DAISY Pipeline is a cross-platform, open source framework for DTB-related document transformations (DAISY Consortium, 2011).

with Special Needs, Lecture Notes in Computer Science, 2010, Volume 6179/2010, pages 57-61.

- DAISY Consortium (2011): *DAISY Pipeline: DTB-related Document Transformations*. [2011.08.15: http://www.daisy.org/pipeline]
- Håkansson, Joel (2006): *Braille In DAISY Project Charter*. [2011.08.15, DAISY Consortium: http://www.daisy.org/project/braille-daisy-project-charter]
- Håkansson, Joel (2009): *Introduction to Portable Embosser Format*. [2011.08.15, DAISY Consortium: http://www.daisy.org/projects/braille/braille_workarea/pef.html]
- Håkansson, Joel (2011): *Introduction to the design of Dotify*. [2011.08.15, Google: http://code.google.com/p/dotify/wiki/DotifyOverview]
- Liblouis (2011). [2011.08.15, Google: http://code.google.com/p/liblouis]
- Liblouisxml (2011). [2011.08.15, Google: http://code.google.com/p/liblouisxml]
- Spiewak, Vincent / Strobbe, Christophe / Engelen, Jan (2009): *Odt2DAISY: Authoring Full DAISY Books with OpenOffice.org*. [2011.08.15, slideshare: http://www.slideshare.net/vspiewak/odt2daisy-authoring-full-daisy-books-with-openofficeorg]