Proceedings of the asmeconf
International Examples Congress and Exposition
AIECE21
January 20, 2021, Cambridge, MA

AIECE2021-0002

LANGUAGE SUPPORT IN ASMECONF: NON-LATIN ALPHABETS, LUALATEX, AND FONTSPEC

John H. Lienhard V[1,*]

[1]Department of Mechanical Engineering
Massachusetts Institute of Technology, Cambridge, MA

ABSTRACT

This note describes the use of asmeconf *to format multilingual documents in Latin or non-Latin alphabets. Font support encompasses the Arabic, Chinese, Greek, Hindi, Japanese, Korean, Marathi, Russian, and Tamil languages, among others. For asian alphabets,* LuaLaTeX *and* fontspec *are employed. The system fonts that must be installed for* fontspec *are listed, and examples of simple abstracts are shown in twenty-five languages.*

Keywords: asmeconf, language support, non-Latin alphabets, fontspec, LuaLaTeX

1. INTRODUCTION: WHY HAVE THIS?

The asmeconf class [1] provides a template for formatting conference papers submitted to the American Society of Mechanical Engineers. The goal of adding language support to asmeconf is to enable authors to include translations of a paper's abstract or brief quotations in languages other than English. Although the entire asmeconf template may, in principle, be switched to another language without modifying the class file, I have not explored this option in much detail. These language capabilities are experimental, and their future development will be guided by the feedback that I receive from authors.

2. THE BABEL PACKAGE

The typesetting of languages is handled by the babel package [2], which is called by the asmeconf class. For many languages, babel includes language definition files (.ldf) that provide information about section or caption titles, hyphenation rules, and so on. When an .ldf exists, babel will recognize the language as a global option that can be passed as an option to asmeconf, assuming that an appropriate font is available. A list of the many languages with .ldf files is given in the babel documentation.

For languages in Latin scripts, it's usually safe to assume that the font is present, and many such languages have .ldf files. For other scripts, additional steps are needed. The asmeconf class handles this differently under pdfLaTeX and LuaLaTeX.

*Corresponding author: lienhard@mit.edu
Version 1.0, January 18, 2021

3. NON-LATIN SCRIPTS UNDER PDFLATEX

When using pdfLaTeX, asmeconfwill load appropriate fonts for Greek, Vietnamese, and certain cyrillic-script languages (see Table 1). The user can give the corresponding class option and then call for a change of language as described in Section 6. No additional work is required.

To access a broader range of fonts, asmeconf can be used under LuaLaTeX, with fontspec. In that case, asmeconf will employ fonts that are installed in the user's operating system, rather than LaTeX fonts.

4. SYSTEM FONTS

The fontspec package [3] allows LuaLaTeX to access fonts that are installed on the user's system. Today, these fonts are normally in unicode, a 16-bit format that allows a font to contain a vast number of glyphs—up to 2^{16}. Multiple languages can be contained within a single font. Specialized unicode fonts are dedicated to particular languages, especially those such as Japanese that have many thousands of characters.

When processed in pdfLaTeX, asmeconf uses the newtxtext and inconsolata fonts, a collection of eight-bit fonts, for Latin script. To use fontspec, we must replace those fonts with corresponding unicode fonts (the math fonts, from newtxmath, are unchanged). Specialized fonts are needed for some additional scripts. Thus, the user will need to install several unicode fonts onto their own system in order to use asmeconf with fontspec. Fortunately, all these fonts are all free and easily downloaded.

The needed fonts are listed in Table 2.

5. USING FONTSPEC WITH ASMECONF

When running LuaLaTeX, the [fontspec] option should be called, to load the appropriate fonts. With fontspec, babel will use .ldf files (if available) and separate initialization files (.ini). If a language option is called for which there is no .ldf file, an

TABLE 1: Languages in non-Latin scripts for which asmeconf is known to provide font support. Class options that must be called are shown.

Language	Option	pdfLaTeX	LuaLaTeX
Arabic	bidi=basic		✓
Belarusian	belarusian	✓	✓
Bengali	—		✓
Bulgarian	bulgarian	✓	✓
Chinese	—		✓
Greek	greek	✓	✓
Hindi	—		✓
Japanese	japanese		✓
Korean	—		✓
Macedonian	macedonian	✓	✓
Marathi	—		✓
Russian	russian	✓	✓
Serbian	serbianc*	✓	✓
Tamil	—		✓
Ukrainian	ukrainian	✓	✓
Vietnamese	vietnamese	✓	✓

* Serbian option [serbianc], for both engines, uses cyrillic. In pdfLaTeX, use \selectlanguage{serbianc}. In LuaLaTeX, instead select "serbian-cyrillic".

error will result. However, such languages may still have an .ini that provides necessary information. For example, Chinese and Korean do not have .ldf files, but they do have .ini files. These languages can be accessed as described in Section 6.

Japanese typesetting is a little more complicated. When [japanese] is given as an option to the class, asmeconf calls the luatexja-fontspec package [4], which is a specialized module for typesetting Japanese.

When captions and the like are not needed (as for short passages), babel can load many languages "on the fly", with only a basic call in the .tex file (see Section 6), if an appropriate font is available.

What about support for scripts not shown in Table 2? Macros from babel for adding fonts can be placed into the preamble of your document. The babel package supports roughly 250 languages, and asmeconf has been tried with only about thirty.

6. HOW TO CALL A LANGUAGE

A language is called by \begin{selectlanguage}{<lang>}, where <lang> is the lower-case name of the language. For example, suppose that a Spanish language abstract is desired. The user puts [spanish] as a global option (this language has an .ldf file), and then writes:

```
\begin{selectlanguage}{spanish}
\begin{abstract*}
Este es el resumen del artículo...
\end{abstract*}
\end{selectlanguage}
```

Nota Bene: **1)** Your .tex file *must* be saved in utf-8 encoding. Some operating systems default to a different encoding that will garble unicode characters. **2)** The features used to provide language support under fontspec require an up-to-date LaTeX distribution (2020 or later). **3)** The features described here require asmeconf version 1.22 or later (2021).

7. ABSTRACTS

Examples of abstracts in various languages now follow. Reading the source .tex file for this document may clarify the syntax.

摘要

这是文章的摘要。我们用中文书写, 描述了问题, 方法和结果, 还包括了参考文献。

摘 要

這是文章的摘要。我們用中文書寫, 描述了問題, 方法和結果, 還包括了參考文獻。

摘 要

係文嘅摘要。我哋用中文書寫, 描述了問題, 方法同結果, 仲包括埋參考文獻。

RESUMEN

Este es el resumen del artículo. Escribimos en español. Se describen el problema, los métodos y los resultados. También se incluyen referencias.

ABSTRACT

This is the summary of the article. We write in English. The problem, methods, and results are described. References are also included.

सारांश

यह हिंदी में लिखे गए एक लेख का सारांश है। समस्या, विधियों और परिणामों का वर्णन किया गया है। संदर्भ भी शामिल हैं।

সারসংক্ষেপ

এটি নিবন্ধের সংক্ষিপ্তসার। আমরা বাংলা ভাষায় লিখি। সমস্যা, পদ্ধতি এবং ফলাফল বর্ণনা করা হয়। উল্লেখগুলিও অন্তর্ভুক্ত রয়েছে।

ملخص

هذا هو ملخص المقال. نكتب بالعربية. يتم وصف المشكلة والطرق والنتائج. يتم تضمين المراجع أيضاً.

RESUMO

Este é o resumo do artigo. Escrevemos em português. O problema, métodos e resultados são descritos. Referências também estão incluídas.

АННОТАЦИЯ

Это резюме статьи. Пишем по русски. Описаны проблема, методы и результаты. Библиография также включена.

概要

この論文の日本語での要約は以下のとおりです。問題、方法、および結果が説明されています。参考資料も添付してあります。

TABLE 2: System fonts used by `asmeconf` with `fontspec`. For all fonts, load regular and bold face. For Latin, Cyrillic, and Greek, also load italic and bold italic. For Noto Sans Arabic, install semibold instead of bold.

Script	Language	Fonts	Where to get the font
Latin*	most European languages	Tex Gyre Termes, Tex Gyre Heros	http://www.gust.org.pl/projects/e-foundry/tex-gyre
Arabic	Arabic, Punjabi, Urdu, others	Amiri, Noto Sans Arabic	https://github.com/alif-type/amiri https://github.com/googlefonts/noto-fonts
Bengali	Assamese, Bengali, others	Noto Serif Bengali, Noto Sans Bengali	https://github.com/googlefonts/noto-fonts
Cyrillic	Belarusian, Bulgarian, Macedonian, Russian, Serbian, Ukrainian, others	Noto Serif, Noto Sans, Noto Sans Mono	https://github.com/googlefonts/noto-fonts
Devanagari	Hindi, Kashmiri, Marathi, Nepali, Sanskrit, others	Noto Serif Devanagari, Noto Sans Devanagari	https://github.com/googlefonts/noto-fonts
Greek	Greek	Noto Serif, Noto Sans, Noto Sans Mono	https://github.com/googlefonts/noto-fonts
Hangul	Korean	Noto Serif CJK KR, Noto Sans CJK KR, Noto Sans Mono CJK KR	https://github.com/googlefonts/noto-fonts
Japanese	Japanese	Noto Serif CJK JP, Noto Sans CJK JP, Noto Sans Mono CJK JP	https://github.com/googlefonts/noto-fonts
Simplified Chinese	Mandarin	Noto Serif CJK SC, Noto Sans CJK SC, Noto Sans Mono CJK SC	https://github.com/googlefonts/noto-fonts
Tamil	Tamil, others	Noto Serif Tamil, Noto Sans Tamil	https://github.com/googlefonts/noto-fonts
Traditional Chinese	Traditional Mandarin, Cantonese	Noto Serif CJK TC, Noto Sans CJK TC, Noto Sans Mono CJK TC	https://github.com/googlefonts/noto-fonts

* The Latin fonts are *required* with `asmeconf+fontspec`, even if English is the only language called. Install others as needed.

सारांश

हा लेखाचा सारांश आहे. आपण मराठीत लिहित आहो. ह्यात समस्या , पद्धती आणि परिणामाचे वर्णन केले आहेत. संदर्भ देखील समविष्ट आहेत.

ÖZET

Bu, makalenin özetidir. Türkçe yazıyoruz. Sorun, yöntemler ve sonuçlar açıklanmaktadır. Referanslar da dahildir.

초록

이것은 한국어로 쓰인 논문의 초록입니다. 문제, 방법 및 결과가 설명되어 있습니다. 참조도 포함됩니다.

RÉSUMÉ

Ceci est le résumé de l'article. Il est écrit en français. Le problème, les méthodes et les résultats sont décrits. Des références sont également incluses.

ZUSAMMENFASSUNG

Hier ist die Zusammenfassung des Beitrags. Wir schreiben auf Deutsch. Die Fragestellung, die Methoden und die Ergebnisse werden beschrieben. Der Beitrag enthält auch ein Literaturverzeichnis.

சாராம்சம்

இது கட்டுரையின் சுருக்கம். நாங்கள் தமிழில் எழுதுகிறோம். சிக்கல், முறைகள் மற்றும் முடிவுகள் விவரிக்கப்பட்டுள்ளன. குறிப்புகளும் சேர்க்கப்பட்டுள்ளன.

TÓM TẮT NỘI DUNG

Đây là phần tóm tắt của bài báo khoa học. Chúng tôi viết bằng tiếng Việt. Vấn đề, các phương pháp và các kết quả được mô tả trong phần này. Tài liệu tham khảo cũng được bao gồm.

SOMMARIO

Questo è il riassunto dell'articolo. Scriviamo in italiano. Vengono descritti il problema, i metodi e i risultati. Sono inclusi anche i riferimenti.

RINGKASAN

Ini adalah ringkasan dari artikel tersebut. Kami menulis dalam bahasa Indonesia. Masalah, metode, dan hasil dijelaskan. Referensi juga disertakan.

STRESZCZENIE

To jest podsumowanie artykułu. Piszemy po polsku. Opisano problem, metody i wyniki. Literatura źródłowa zostanie udostępniona.

MUHTASARI

Huu ni Muhtasari wa makala. Tunaandika kwa Kiswahili. Matatizo, Mbinu na matokeo yameelezewa. Marejeleo pia yamejumuishwa.

АНОТАЦІЯ

Це короткий зміст статті. Пишемо українською мовою. Описана проблема, методи та результати. Стаття включає список використаної літератури.

ΠΕΡΙΛΗΨΗ

Αυτή είναι η περίληψη του άρθρου. Χρησιμοποιούμε την ελληνική γλώσσα. Περιγράφεται το πρόβλημα, οι μέθοδοι και τα αποτελέσματα. Περιλαμβάνονται επίσης αναφορές.

САЖЕТАК

Ово је резиме чланка. Пишемо на српском. Описани су проблем, методе и резултати. Укључене су и референце.

SUMMARIUM

Hoc argumentum in articulum. Latine scribere nobis. Quaestio est, modi, et describit eventus qui. Alia sunt opera citatis.

8. CHANGING THE LANGUAGE OF THE ENTIRE DOCUMENT

The main language of the document is English by default. This means that section headings, captions, and other words chosen by macros will be in English.

To chose a different main language, give the option `[main=..]`, e.g., `[main=french]`. The standard caption and section names will follow `babel`'s dictionary for the language chosen. Users may additionally change "Keywords", "Nomenclature", "Corresponding author", and "Joint first authors" by renewing the commands `\keywordname`, `\nomname`, `\CAwords`, and `\JAwords`. Changes to the page footer are described in the `asmeconf` documentation [1]. The pdf bookmark for "Appendices" may be changed by renewing `\appendicesname`. (See the `babel` documentation [2, §1.15] if the dictionary for the main language lacks the correct word for, say, `abstractname`).

Note that some languages make characters active, which can have unexpected effects. For example, Spanish makes "." an active character, which clashes with the default `dcolumn` decimal separator, `d{n.m}` [5]. `\spanishdeactivate{.}` stops this behavior. Alternatively, the decimal separator can be changed to a comma by putting `\newcolumntype{d}[1]{D{,}{,}{#1}}` in the preamble.

I have not explored this usage in detail, and I would welcome comments from authors who attempt such conversions.

9. CONCLUSION

`asmeconf` has multilingual capabilities under both pdfLaTeX and LuaLaTeX. In particular, with `fontspec`, `asmeconf` can support a wide range of scripts and languages within a LaTeX environment. These capabilities are largely experimental, and their future development will be guided by feedback that I receive from those using these features.

10. AKNOWLEDGMENTS

Thanks to everyone who helped me sort out the translated abstracts: Leon Awerbuch, Debanjan Banerjee, George Barbastathis, Svetlana Boriskina, Jacopo Buongiorno, Hyung Won Chung, Aslı Demir, Vibhawari Deshmukh, Zi Hao Foo, Turga Ganapathy, Steven Gerasimoff, Christine Gervais, Igo Krebs, Omar Labban, Nguyen Le, Bora Mikić, Judith Mmari, Aarón Montoya-Moraga, Thao Nguyen, Yoshiki Okamoto, Danyal Rehman, Yagnaseni Roy, Jaichander Swaminathan, Gregory Thiel, and Keiji Yano. All were generous with their time.

REFERENCES

[1] John H. Lienhard, V. "`asmeconf`: A template for ASME conference papers." Version 1.22. Comprehensive TeX Archive Network (2021). Accessed January 18, 2021, URL https://ctan.org/pkg/asmeconf.

[2] Braams, Johannes L. and Bezos, Javier. "`babel`: Localization and internationalization." Version 3.52. Comprehensive TeX Archive Network (2020). Accessed January 3, 2021, URL https://ctan.org/pkg/babel.

[3] Robertson, Will. "The fontspec package: Font selection for LuaLaTeX and XeLaTeX." Version 2.7i. Comprehensive TeX Archive Network (2020). Accessed January 5, 2021, URL https://ctan.org/pkg/fontspec.

[4] LuaTeX-ja project team. "The LuaLaTeX-ja package." Version 20210103.0. Comprehensive TeX Archive Network (2021). Accessed January 5, 2021, URL https://ctan.org/pkg/luatexja.

[5] Carlisle, David. "The `dcolumn` package." Version 1.06. Comprehensive TeX Archive Network (2014). Accessed January 8, 2021, URL https://ctan.org/pkg/dcolumn.