

**Baker, Josef B.; Sexton, Alan P.; Sorge, Volker**

**MaxTract: converting PDF to LaTeX, MathML and text.** (English) [Zbl 1360.68882](#)

Jeuring, Johan (ed.) et al., Intelligent computer mathematics. 11th international conference, AISC 2012, 19th symposium, Calculemus 2012, 5th international workshop, DML 2012, 11th international conference, MKM 2012, systems and projects, held as part of CICM 2012, Bremen, Germany, July 8–13, 2012. Proceedings. Berlin: Springer (ISBN 978-3-642-31373-8/pbk). Lecture Notes in Computer Science 7362. Lecture Notes in Artificial Intelligence, 422-426 (2012).

**Introduction:** In this paper we present the first public, online demonstration of MaxTract; a tool that converts PDF files containing mathematics into multiple formats including LaTeX, HTML with embedded MathML, and plain text. Using a bespoke PDF parser and image analyser, we directly extract character and font information to use as input for a linear grammar which, in conjunction with specialised drivers, can accurately recognise and reproduce both the two dimensional relationships between symbols in mathematical formulae and the one dimensional relationships present in standard text.

The main goals of MaxTract are to provide translation services into standard mathematical markup languages and to add accessibility to mathematical documents on multiple levels. This includes both accessibility in the narrow sense of providing access to content for print impaired users, such as those with visual impairments, dyslexia or dyspraxia, as well as more generally to enable any user access to the mathematical content at more re-usable levels than merely visual. MaxTract produces output compatible with web browsers, screen readers, and tools such as copy and paste, which is achieved by enriching the regular text with mathematical markup. The output can also be used directly, within the limits of the presentation MathML produced, as machine readable mathematical input to software systems such as Mathematica or Maple.

For the entire collection see [\[Zbl 1245.68013\]](#).

**MSC:**

[68U15](#) Computing methodologies for text processing; mathematical typography

**Software:**

[Maple](#); [Mathematica](#); [LaTeX](#); [MaxTract](#); [OCGtools](#)

**Full Text:** [DOI](#)