TeXTALK

T.V.Raman
Center for Applied Mathematics
Cornell University
Ithaca NY 14853–6201

October 6, 1990

Abstract

Traditionally books have become accessible to the visually handicapped only after they have been either published in Braille or more recently after they have been recorded in Talking Book format. However not all publications can be made accessible in this manner, and this still means that the choice of reading material available, especially in the technical fields continues to be severely restricted. The fact that more and more publications are first written by the author using a computer along with the development of voice synthesizer technology presents us with an excellent opportunity of overcoming this problem.

Several advances have been made during the past few years using voice synthesizer technology and optical character recognition culminating in the development of reading machines. These machines can read out printed matter, provided the quality of printing is extremely good. In addition they suffer the added disadvantage that special characters such as those occurring in Mathematical texts continue to present a problem. For material published before the time when books and papers were written using the computer, this approach is perhaps the only way out. However the fact that more and more publications are available in electronic format needs to be exploited as there is no point in trying to reconvert something that is already available in computer readable format using optical character recognition.

This of course leads us to the problem of reading TeX source directly using a talking computer. TeX has the advantage of representing all special symbols, using special names that are formed using the letters of the alphabet along with a few extra special characters. Thus the Greek letters, which are ubiquitous in any Mathematical text and present severe problems to any reading machine are already present in the TeX source in a format that can be read out by a voice synthesizer. However we also need to take account of the fact that a TeX source file is adorned with a lot of special characters that can sound extremely confusing if read out by the synthesizer.

We of course have the option of stripping away all these extra characters from the TeX source, but by doing this we would actually lose valuable information, that is being given to us for free. There is enough information present in a TEX source file to enable a printer to lay out Mathematical formulae and so that information should certainly be useful in transforming the text to a format that would be comparable to what a person reading out the printed text aloud would generate.

Consider the following example:

$\frac{1+\sqrt{5}}{2}$

This is what the above expression sounds like when read out by a voice synthesizer:

backslash left paren backslash frac leftbrace one plus backslash sqrt leftbrace five rightbrace rightbrace leftbrace two rightbrace backslash right paren .

It certainly sounds extremely mysterious! Instead consider the following piece of transformed text:

The fraction with numerator 1 + square root of 5 and denominator 2 end of fraction.

This certainly compares very well with what a person reading out the printed text resulting from the TEX source would say.

Similarly $A \subseteq A \cup B$ is transformed to A is a subset of A union B. TEXTALK is a program that looks at the TEX source and carries out transformations like the above. The resulting text is much easier to understand when read out by a voice synthesizer. As becomes evident from the introductory paragraphs the potential presented by this kind of software is enormous. The program as it currently stands is eminently usable but there is certainly a lot more work to be done on it. When complete it will perhaps result in something I have always hoped to achieve namely have the same kind of access to published material as anyone else. [1]

The above paragraphs discuss the transformation of specific pieces of text that occur within mathematical expressions. The document being read itself can be considered to have special text units that require special processing. Thus a technical document could be comprised of *equations, arrays, figures, tables* and we will have to process each of them differently. Thus we need to process different environments in the document using transformations that are suitable to the specific environment. Thus we need to transform the character “”’ to “prime” within the math environment of TEX . Similarly in the array environment we could use the fact that the “&” character acts as a field separator . Figures appearing in the text require special processing resulting perhaps in a tactile output of the figure. Occurrences of figures in the text can be easily detected and the entire text corresponding to the picture can be replaced by “picture deleted”.

All of the above discussion has been centered around TEX . However we can now generalise all of the above in order to achieve the same results with any word processing language. All we need to identify is the grammar used by that language and code the necessary transformations.

[1] I am visually handicapped myself and use a talking terminal to interact with my computer.