

# Automatic Simplification of Numerical Expressions using Fuzzy Logic

Susana Bautista<sup>1</sup>, Pablo Gervás<sup>2</sup>, Raquel Hervás<sup>1</sup>

<sup>1</sup> *Facultad de Informática, Universidad Complutense de Madrid*  
subautis@fdi.ucm.es, raquelhb@fdi.ucm.es

<sup>2</sup> *Instituto de Tecnología del Conocimiento, Universidad Complutense de Madrid*  
pgervas@sip.ucm.es

We live in an “Information Technology Society”, where there is a tendency to digitalize all kinds of information. The way in which information is written or presented can exclude many people whose level of reading skills makes them have problems in reading comprehension. There are several factors by which these skills can be affected, such as having had limited access to training, having social problems or having some cognitive disability. In addition, there are specific groups such as the deaf, the autistic, people with language disorders such as aphasia or dyslexia, people who are learning another language or the elderly, who have specific problems with reading.

A first solution to this problem is the manual simplification of information manually to adapt to the difficulties of target users to whom it is directed. There are several initiatives designed to develop the manual processing of text simplification following the European guidelines established by the IFLA [4]. However, manual simplification is too slow and tedious to be efficient in producing the desired material. Therefore, various attempts to automate part of this simplification process have been launched, focusing on the different transformations that can be applied in the process of text simplification.

The aim of simplifying texts is to transform a text to make it easier to understand for certain target users. These tasks mainly aimed at syntactic and lexical constructions that can be applied to the original text to generate a simplified version. In order to perform these tasks researchers must identify what causes this difficulty in specific readers [3], [5].

We focus on a case of information that can cause problems of understanding for many people and creates difficulties for readers: numerical information [1], [2]. Many times, we access information that is represented in the form of numerical expressions such as economic data, statistical data or demographic data. In addition we can find numerical information on a recipe, a news article or a report.

In our work we consider the expressions which represent quantities to be *numerical expressions*; *52.3%*, *0.48* or *3489*. These expressions might be simplified as *more than 50%*, *around 1/2* or *almost 3500*. In the process of simplification there is a loss of precision to achieve the simplified version of numerical expressions. The modifiers *more than* or *around* are used in these cases to solve this loss of precision. Furthermore, we consider different strategies such as changing the mathematical representation of the expression or rounding the original quantity. For example, *25%* can be transformed by *1/4* or *74.7%* by *75%*.

---

This paper has been partially supported by the project Prosecco (600653), funded by the European Commission under FP7, the ICT theme, and the Future Emerging Technologies (FET) program.

In this kind of strategies there is a decision to choose a modifier depending on the loss of numerical precision. We need to translate the numerical information lost into linguistic representation using words. We work with a range of values of the original numerical expressions and we want to obtain discrete values represented by linguistic modifiers to simplify the expressions. We want to study applying techniques of fuzzy logic in order to achieve the simplified versions of the original numerical expressions where there is a loss of precision.

The decision to use a modifier or not and what kind of modifier to select is based on the difference between the value of the original quantity of the numerical expression and the value of the simplified expression in each strategy. The kind of modifier depend on the rounded value of the original expression and the distance calculated by the loss of precision. These features should be considered in a set of rules in fuzzy logic in order to define what kind of modifiers could be used in order to help to generate the simplified version or the original numerical expression.

With this goal, the main of this work is to carry out the automatic simplification of numerical expressions present in the text using fuzzy logic to improve the current approaches. Our work is based on the conclusions achieved by empirical studies [1], [2] developed with experts. We consider real important to include fuzzy logic to redefine the rules of simplification applied in our systems and evaluate the output in real experiments with users. The way in which information is presented can cause reading and comprehension problems for many people. The adaptation of information is not an easy process but clearly necessary.

## References

- [1] S. Bautista, R. Hervás, P. Gervás, R. Power, and S. Williams. A System for the Simplification of Numerical Expressions at Different Levels of Understandability. In *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility (NLP4ITA)*, 2013.
- [2] S. Bautista and H. Saggion. Making Numerical Information more Accessible: Implementation of a Numerical Expressions Simplification Component for Spanish. *ITL-International Journal of Applied Linguistics. Special Issue on Readability and Text Simplification*. Peeters Publishers, Belgium, 165(2):299–323, 2014.
- [3] J. Carroll, G. Minnen, Y. Canning, S. Devlin, and J. Tait. Practical Simplification of English Newspaper Text to Assist Aphasic Readers. In *Proceedings of the Workshop on Integrating Artificial Intelligence and Assistive Technology (AAAI)*, pages 7–10, Madison, Wisconsin, 1998.
- [4] G. Freyhoff, G. Hess, L. Kerr, E. Menzel, B. Tronbacke, and K. V. Veken. European Guidelines for the Production of Easy-to-Read Information for People with Learning Disability. Technical Report ISLMH, 1998.
- [5] A. Siddharthan and M. Angrosh. Hybrid text simplification using synchronous dependency grammars with hand-written and automatically harvested rules. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2014)*, Gothenburg, Sweden, 2014.