Natural Language Processing with Mathematical Fuzzy Logic

Matthias Hofer
Vienna University of Technology
(Supported by FWF projects I1897-N25 and W1255-N23)

The world wide web became part of our daily lives, and as such many people consider it a global service contractor. While, as a matter of fact, information on the web is not always edited in a systematic way, it is widely agreed upon, that logic, regarded as a scientific discipline, is the tool of choice to ease the assessment of complex information. Classical FOL is not always the adequate theory to express the content of potentially huge data sets. As an example, consider a statement taken from the webpage of the national library of Austria:

"The department's holdings of old printed books, about 500.000 in total, covers...".

Expressing the same content in a similarly well perceivable way in classical FOL, requires the exact number of books kept in the library, while a deviation of just one item would render the statement completely wrong. This issue can neatly be resolved in MFL, by adding vague quantifiers, like "about n", or "about $x\%$", and also allowing for intermediate truth values in the real unit interval. This leads to an increase of robustness, and to a representation of content much closer to human communication than in classical FOL. In this regard, two complementary procedures for applying MFL to huge data sets on the web can be explored:

- querying
- data summarization

While querying is a quite well investigated field, data summarization becomes more and more important, as world wide operating companies gather more information from the web than their staff could ever summarize by hand. Hence, expressive and robust theories, as available within the frame of MFL, could have great potential impact on existing endeavours in this direction. Therefore, the evaluation of quantitative natural language statements, as well as the extraction of true such statements from a given database, is a challenging and promising field of activity, worth exploring in much detail. We can understand this as a semantic foundation for existing statistical industrial applications. Following this line of thought, we can also stipulate that MFL might itself benefit from focusing more on NLP (natural language processing) applications, since this is where a large number of people would indeed be directly affected. Even though this seems clear to many researches of MFL, the theory itself is neither designed nor well adjusted to that fact yet. Since Microsoft, Facebook, Google and many more try to achieve exactly this kind of applications by whatever means works, the MFL community should explore the potential of its theory into this direction in the least marginal manner.

References

Cintula, P. and Hájek, P. and Noguera, C. (2011). *Handbook of Mathematical Fuzzy Logic*. College Publications.

Cintula, P. and Fermüller, C.G.. and Noguera, C. (2015). *Handbook of Mathematical Fuzzy Logic - Volume 3*. College Publications.

Glöckner, I. (2006). *Fuzzy Quantifiers - A computational theory*. Springer.