

Fuzzy Matching of Strings

Lyubomir Filipov, Zlatko Varbanov

*Department of Information Technologies
University of V. Tarnovo “St. Cyril and St. Methodius”
Veliko Tarnovo, Bulgaria
lubomir_g_1991@abv.bg*

Keywords: fuzzy matching, approximate string matching, record linkage

Fuzzy matching is a technique used in computer-assisted translation as a special case of record linkage. It is implemented in most of the database engines and it is widely used in autocomplete of data.

In this paper fuzzy matching as part of approximate string matching is investigated. Basic algorithms like Soundex, Bitap, Boyer-Moore [1,2] are covered. Using the results about those algorithms, several database engines are compared and new way of handling fuzzy matching is offered.

References

- [1] G. Myers. A fast bit-vector algorithm for approximate string matching based on dynamic programming. *Journal of the ACM* 46 (3), May 1999, 395–415.
- [2] Robert S. Boyer, J. Strother Moore, A fast string searching algorithm, *Communications of the ACM*, Volume 20, Oct. 1977, 762–772.