Information Retrieval

Gábor Rédey, Attila Neumann, Zoltán Sütő

Abstract

This paper is just a thoughtful initiative to reconsider information retrieval languages and thereby the whole information retrieval process.

Human thinking is mostly language independent, however, human information exchange is massively language dependent, therefore the process of information retrieval is obviously language dependent, too. Natural languages are still not commonly used as information retrieval languages, as despite the idea itself seems to be promising, it poses several serious problems as well. Though, there have been scientific researches and achievements targeting this domain since the early 90's.

Traditional information retrieval languages have rather poor syntactic structure compared to that of natural languages. Many times vaguely represented queries can be replied only by "noisy" irrelevant answers loaded by imperfect information. In this case progress can be expected from information retrieval languages with more versatile syntax, which can model not only discrete concepts (or their conjuncts), but also their natural language relations as well. The paper is going to outline such a new type of information retrieval language with a better expressive power.

The paper shows that there are knowledge representation languages which can faithfully represent the logical fine structure of natural languages, and to which both natural language texts and background knowledge (ontologies, thesauri) can easily (almost automatically) be translated, while the query itself can also be represented by the same language. A full explanation can be given by a step by step tracing the generated answer.