Abstract

Logic-based models have been already proposed for information retrieval purpose. However, there is a need for new formalisms providing more generic frameworks. For this purpose, an information retrieval axiomatic theory is proposed in this paper, independently of any model. Our proposal which mainly relies on many-sorted logic allows to consider various sets in the domain of discourse that provides us a rich framework to model the different items such as documents, index terms, queries. The theory relies on a sound set of axioms driving the retrieval process as proof of theorems. As such the genericity consists of a main motivation; it will be proved that three classical information retrieval models, namely the Boolean model; the fuzzy-set-based extension of the Boolean model; and the vector space model, satisfy the proposed theory, establishing then its consistency. Beyond the genericity, the proposed approach may face concrete problems. Indeed, it is well known that the use of the classical settings of formal concept analysis theory for information retrieval does not allow disjunctions and negations in queries. For this purpose, this paper gives a characterization of these queries forms using appropriates theorems of the theory. Useful algebraic properties (i.e., isomorphisms) are then established for this end.