

Genetic-based approach for cue phrase selection in dialogue act recognition

ABSTRACT

Automatic cue phrase selection is a crucial step for designing a dialogue act recognition model using machine learning techniques. The approaches, currently used, are based on specific type of feature selection approaches, called ranking approaches. Despite their computational efficiency for high dimensional domains, they are not optimal with respect to relevance and redundancy. In this paper we propose a genetic-based approach for cue phrase selection which is, essentially, a variable length genetic algorithm developed to cope with the high dimensionality of the domain. We evaluate the performance of the proposed approach against several ranking approaches. Additionally, we assess its performance for the selection of cue phrases enriched by phrase's type and phrase's position. The results provide experimental evidences on the ability of the genetic-based approach to handle the drawbacks of the ranking approaches and to exploit cue's type and cue's position information to improve the selection. Furthermore, we validate the use of the genetic-based approach for machine learning applications. We use selected sets of cue phrases for building a dynamic Bayesian networks model for dialogue act recognition. The results show its usefulness for machine learning applications.

Keyword: Genetic algorithm; Feature selection; Cue phrase selection; Ranking feature selection; Dialogue act recognition