Abstract

Anaphora occurs very frequently in written texts and spoken dialogues. Almost all NLP applications such as machine translation, information extraction, automatic summarization, question answering system, natural language generation, etc., require successful identification and resolution of anaphora. Though the significant amount of work has been done in English and other European languages, the computational work, in reference to Hindi, is lagging far behind. In this paper, we present a review of work done in the field of anaphora resolution in Hindi. We also cover different issues and challenges in developing computational models for Hindi. The paper covers issues related to syntactic/semantic structure of Hindi and influence of cases on pronouns, mainly personal pronoun. The work is mainly dedicated to pronominal anaphora. As the pronouns in Hindi do not differentiate gender, it is the verb and case markers that signify different genders. Influence of case markers in deciding the antecedent is explored. Pronoun resolution in context of EHMT (English-Hindi Machine Translation) systems is demonstrated to substantiate the need of anaphora resolution for NLP application.

Index Terms

Computer Science Artificial Intelligence

Keywords

Anaphora Anaphora Resolution Case Markers Natural Language Processing Pronominal
Anaphora