

## Project Summary

**Intellectual merit.** Knowledge is distilled into the informal or natural language used in ordinary discourse. Much of it is computationally stored in texts. Computational understanding is critical to handling the large amounts of text that already exist; as well as, the geometrically growing volume that will come to exist. Knowledge cannot be readily summarized, accessed or understood unless inherent utterance ambiguity is detected and resolved. Recognizing *intentional* ambiguity in natural language is a vertical step to handling natural language. Humor often depends on intentional ambiguity; detecting it offers a clear, distinct research focus. Therefore, the *purpose* of this research is to detect intentional ambiguity that results in humor. The *objective of this research* is to computationally detect short jokes for young children that are dependent on lexical ambiguity or phonological similarity. It is a step toward attainment of the Principle Investigator's long-term career goals, which include imprecise ontology development, autonomous knowledge extraction, autonomous semantic tagging, and disambiguation of whole language dialogue. Pursuing two specific aims will attain the proposal's overall research objective: (1) Build a description logic-based imprecise ontology, containing a minimal number of concepts and relations between them for annotation of texts for young children. An ontology will be constructed from a children's dictionary and a collection of young children's texts. Restricting the domain to young children's jokes is expected to decrease complexity while retaining the core issues. The domain restriction reduces the background knowledge that needs to be captured. This, in turn, leads to a smaller, more manageable ontology with fewer inferences to be drawn from it. It allows a concentration on humor detection methods with a reduced emphasis on capturing and inferring knowledge. (2) Determine a method for recognition of script overlap and opposition that create jokes based on lexical or phonological ambiguity. Formal methods will be developed for the identification of scripts within a text that lead to humor. The development of such methods is a step closer to an understanding of the product of human cognitive processes and "more human" intelligent systems, which is a *rationale* for undertaking this research. It is expected that once an ontology-based children's joke detector is developed, the ontology can be expanded to include other knowledge. *The contribution is significant because formal computational methods will be developed for the restricted domain of intentional disambiguation represented by humor detection.* The results will fundamentally advance the field of computational humor. It is also expected that the detector can help children and second-language learners master the language, thus meeting one of the *educational objectives* of the research.

**Broader Impacts.** The proposed research will *aid in achieving sociable computing and advance discovery while promoting teaching and learning.* *Sociable computing* seeks to enable informal interaction with a computer. Many people, particularly the elderly and those with limited education have difficulty interacting with computers using today's precise and meticulous interface mechanisms; informal communication is needed. The PI heads his university's multi-disciplinary Data and Knowledge Management group. The group's Education Plan focuses on undergraduate participation in research activities; the PI's research is a magnet. The research will have an important impact on *underrepresented groups*, particularly women; one of the group's members is a woman who acts as a mentor to undergraduate women who wish to become involved with computing research. Similarly, the proposed research may become part of the University of Cincinnati Society of Women Engineers Middle School outreach program for girls grades 5 to 8. Advancement of such students is expected to have *broad societal impact*, which will be complemented by the positive effects that the research outcomes are expected to have on the society as a whole. The *educational infrastructure* of the university will be enhanced by incorporating the proposed research into the Semantic Web and Computational Intelligence courses. Outcomes of both the research and educational endeavors will be *broadly disseminated*. The results will be published in peer-reviewed journals and the proceedings of widely attended conferences.