TOPICS

Managing Ambiguity in Reference Generation: The Role of Surface Structure

Imtiaz H. Khan,[a] Kees van Deemter,[b] Graeme Ritchie[b]

[a]*Department of Computer Science, King Abdulaziz University*
[b]*Department of Computing Science, University of Aberdeen*

Abstract

This article explores the role of surface ambiguities in referring expressions, and how the risk of such ambiguities should be taken into account by an algorithm that generates referring expressions, if these expressions are to be optimally effective for a hearer. We focus on the ambiguities that arise when adjectives occur in coordinated structures. The central idea is to use statistical information about lexical co-occurrence to estimate which interpretation of a phrase is most likely for human readers, and to avoid generating phrases where misunderstandings are likely. Various aspects of the problem were explored in three experiments in which responses by human participants provided evidence about which reading was most likely for certain phrases, which phrases were deemed most suitable for particular referents, and the speed at which various phrases were read. We found a preference for "clear" expressions to "unclear" ones, but if several of the expressions are "clear," then brief expressions are preferred over non-brief ones even though the brief ones are syntactically ambiguous and the non-brief ones are not; the notion of clarity was made precise using Kilgarriff's Word Sketches. We outline an implemented algorithm that generates noun phrases conforming to our hypotheses.

Keywords: Natural language generation; Generation of referring expressions; Ambiguity management; Surface ambiguity in GRE; Surface realization

1. Introduction

When designing a computer system that can produce written text in natural language (e.g., English) from some underlying non-linguistic representation of information (Reiter & Dale, 2000), an important component is the *Generation of Referring Expressions* (GRE).

Correspondence should be sent to Imtiaz Hussain Khan, Department of Computer Science, Faculty of Computing & IT, King Abdulaziz University, Jeddah, P.O. Box 80200, Jeddah, Kingdom of Saudi Arabia 21589. E-mail: ihkhan@kau.edu.sa