

Information State Based Speech Recognition

Rebecca Jonson

Department of Philosophy, Linguistics and Theory of Science

Academic dissertation in linguistics, to be publicly defended, by due permission of the Faculty of Humanities at University of Gothenburg, on May 22, at 1 p.m., in T307, Olof Wijksgatan 6 (Gamla Hovrätten), Humanisten, Göteborg.

Abstract

Ph.D dissertation in Linguistics at University of Gothenburg, Sweden, 2010

Title: *Information State Based Speech Recognition*

Author: Rebecca Jonson

Language: English

Department: Department of Philosophy, Linguistics and Theory of Sciences, University of Gothenburg, Box 200, SE-405-30 Göteborg

Series: Gothenburg Monographs in Linguistics 41

Published at: <http://hdl.handle.net/2077/22169>

One of the pitfalls in spoken dialogue systems is the brittleness of automatic speech recognition (ASR). ASR systems often misrecognize user input and they are unreliable when it comes to judging their own performance. Recognition failures and deficient confidence estimation affect the performance of a dialogue system as a whole and the impression it makes on a user. Humans outperform ASR systems on most tasks related to speech understanding. One of the reasons is that humans make use of much more knowledge. For example humans appear to take a variety of knowledge-based aspects of the current dialogue into account when processing speech. The main purpose of this thesis is to investigate whether speech recognition also can benefit from the use of higher level knowledge sources and dialogue context when used in spoken dialogue systems.

One of the major contributions of this thesis is to provide more insight into what type of knowledge sources in spoken dialogue systems would be potential contributors to the task of ASR and how such knowledge can be represented computationally. In the framework of *information state based* dialogue management we have an important source of semantic and pragmatic knowledge represented in the *information state*. We will investigate if the knowledge in the information state can help to alleviate the search problem and reliability estimation in speech recognition. We call this knowledge and context aware approach to speech recognition *information state based speech recognition*.

The first part of this thesis investigates approaches to obtaining better initial language models more rapidly for spoken dialogue systems and ways of dynamically selecting the most appropriate models based on the dialogue context.

The second part of this thesis concerns the use of the speech recognition output and investigates how additional knowledge sources can enhance a dialogue system's decision-making on how to proceed and make use of speech recognition hypotheses.

The thesis presents several experimental studies addressing the issues described above and proposes an integration of the explored techniques into the GODiS dialogue system.

Keywords: dialogue systems, speech recognition, language modelling, dialogue move, dialogue context, ASR, higher level knowledge, linguistic knowledge, N-Best re-ranking, confidence scoring, confidence annotation, information state, ISU approach.