Finite State Grammar Transduction from Distributed Collected Knowledge

Rakesh Gupta[1] and Ken Hennacy[2]

[1] Honda Research Institute USA, Inc.
800 California Street, Suite 300
Mountain View, CA 94041
rgupta@hra.com
[2] University of Maryland
Institute for Advanced Computer Studies
College Park, MD 20742
khennacy@cs.umd.edu

Abstract. In this paper, we discuss the use of Open Mind Indoor Common Sense (OMICS) project for the purpose of speech recognition of user requests. As part of OMICS data collection, we asked users to enter different ways of asking a robot to perform specific tasks. This paraphrasing data is processed using Natural Language techniques and lexical resources like WordNet to generate a Finite State Grammar Transducer (FSGT). This transducer captures the variations in user requests and captures their structure.
We compare the task recognition performance of this FSGT model with an n-gram Statistical Language Model (SLM). The SLM model is trained with the same data that was used to generate the FSGT. The FSGT model and SLM are combined in a two-pass system to optimize full and partial recognition for both in-grammar and out-of-grammar user requests. Our work validates the use of a web based knowledge capture system to harvest phrases to build grammar models. Work was performed using Nuance Speech Recognition system.

1 Introduction

Humans often wish to communicate with robots about what they like done. It is awkward to be constrained to specific set of commands. Therefore, a free-form interface that supports natural human robot interaction is desirable.

A finite state transducer is a finite automaton whose state transitions are labeled with both input and output labels. A path through the transducer encodes a mapping from an input symbol sequence to an output symbol sequence [1]. Grammar is a structure that defines a set of phrases that a person is expected to say. In this work, our goal is to automate the process of creating a Finite State Grammar Transducer (FSGT) to map utterances to task labels from text data contributed by volunteers over the web.

It is a challenge to develop a grammar that will recognize a large variety of phrases and achieve a high recognition accuracy. Manual creation of a set of