7th International Conference on DEVELOPMENT AND APPLICATION SYSTEMS
Suceava, Romania, May 27 – 29, 2004

BUILDING SYSTEMS FOR RAPIDLY ANALYSIS AND MODELING OF SPEECH FOR CREATION OF LANGUAGE RESOURCES

Nicolae OBJELEAN
State Agrarian University of Modova.
Str. Mircesti 44, Moldova, MD-2049, Moldova Chisinau
n.objelean@uasm.md

Abstract. *We proposed to build systems for rapidly creation of language recourses, which need to resolve two problems of current speech recognition and speech synthesis of natural language. We deal with the problem adaptation of automatic speech recognition systems for new vocabulary. For that aims we proposed building systems for creation linguistics knowledge. The ended goal of this research is to design a speech recognition and synthesis systems that can be rapidly configured for any vocabulary adaptation for new task.*
Keywords: *Language modeling, Speech Recognition, linguistics knowledge.*

Introduction

On of the difficult process in configuring a speech recognizer is the generation of a spelling-to-pronunciation dictionary. This problem is partly alleviated by the use of a very large dictionary and creation of speech language recourses. But in most application, a large dictionary cannot cover all words in the vocabulary. For example, proper names (such as those for a voice-dialing telephone) and applications-specific terms (medical, legal, and scientific) may be absent. A more serious problem in the addition of new words after a system has been configured. This is problematic because users cannot be expected to have sufficient phonetic knowledge to produce reliable phonetic representations.

The Automatic generation systems of pronunciation dictionary

We propose a system that automatically generated the phonetic pronunciation of a new word and given the spelling plus one or two utterance of words. Such a systems could also be used to check existing pronunciation dictionaries. There are three promising strategies for this task:
If *phonetic recognition* were sufficient accurate, we could simply request the user to speak a new word once, and use the recognized phone sequence as the pronunciation for that world. In practice, however, the best phonetic recognizers [1] are only capable of about 60% recognition accuracy. Yet, the same algorithms and models in these mediocre phonetic recognizers can perform extremely well in word or sentence recognition, because of the additional constraints applied at the lexical and grammatical level. Thus, if sufficient constraints were applied to the phonetic recognizer, accurate recognition would be possible [2].
We propose the architecture of systems in which two knowledge sources generate constraints that guide a phonetic recognizer. In particular, we plan to:

- Using of phonetics information for produces the most likely phone sequence.
- Modify a text-to-speech system to produce a network of possible pronunciations for a word given spelling of the word for creation of speech synthesizes systems. This network must contain the right pronunciation of all words in the time.
- Create a probabilistic model that maps spelling to phone sequence. This module will assign probabilities to transitions in the pronunciation network.

Is very important in this approach to pronunciation generation is tight integration of various knowledge sources that are relatively independent. For instance, our best phone recognizer may perform 60% accuracy. However, using syntactical and a finite state pronunciation network drastically restrict the number of substitutions, insertions, and deletions [3]. The addition of probabilities to the network further reduces the complexity of the recognition task, and we expect to achieve perfectly recognition. Moreover, we expect these automatically-generated pronunciations to be more consistent than manually-generated ones, which is especially important to for adding new words. Figure 1 illustrates the *System for generation of pronunciation vocabulary* and the process of generating the network, assigning the probabilities, as well converting a spoken word and it's spelling into a phone sequence.

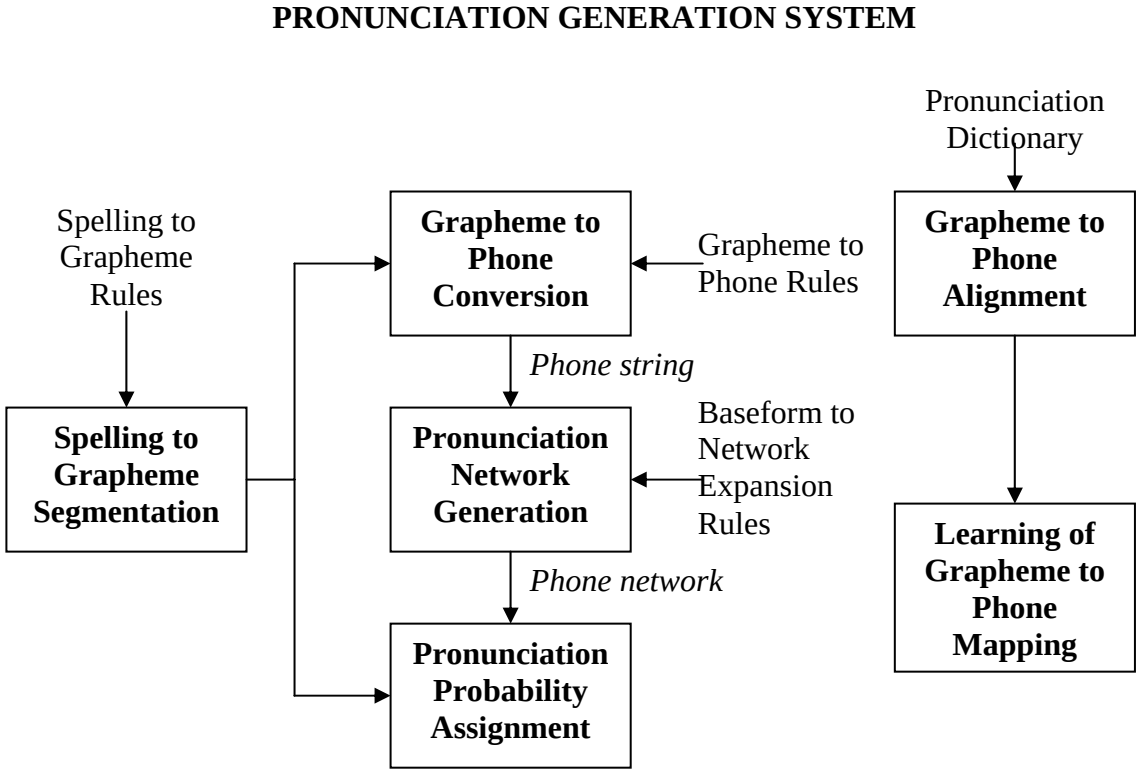

Figure 1. System for generation of pronunciation vocabulary

We will evaluate the resulting pronunciations for new words in several ways. We will first evaluate them by comparing against "correct" pronunciations, and measuring accuracies at phone and word levels. Also, we will use these pronunciations as the new dictionary in a word recognition task to obtain an actual *recognition accuracy*. Both sets of results will be compared against those of a manually derived dictionary in order to determine the feasibility of our techniques.

We believe that can obtain even better performance on the task of pronunciation dictionary generation. First, the above results were obtained with 28450 sentences of training. We expect that we will have an order of magnitude more training data within the thirst year of the proposed research. This will substantially improve the accuracy. Moreover,

we will modify the phonetic recognizer to use *generalized triphones* and *generalized allophones*. Finally, isolated words will be used for pronunciation generation, which makes the recognition task easier, since the speech will be much more carefully spoken. In finale we can to construct a language knowledge and phone speech resources by using this system, and by using new subword models. We believe that system aided creates new probabilistic state grammar and can accurately map words into corresponding phone sequence.

Conclusions

This system proved phonetically word analysis of the text, which examines the structure of each words and aiding the generation of vocabularies for new tasks and give the language knowledge. Also the software developed for collection of language resources, creation phonetically models of the words and using its for generation pronunciations dictionary.

Acknowledgements

I wish to thank of colleagues from Carnegie Mellon University Pittsburgh and especially Dr. Alex Rudniky.

References

[1] Objelean, N. C. *Estimation of complexity recognition of language of speech dialogue in the automated systems*. Engineering Problems. Moscow, Science, p. 74-79, 1981.

[2] Trunin-Donskoi, V.N. *The Automatic recognition of the alphabet as a means of comparison of technical opportunities of speech recognition systems*. Problems of elaboration the speech understanding systems. Moscow, Science, p. 30-46, 1980.

[3] Lee, K. F. *A connectionist Approach to Continuous Speech Recognition*. IEEE International Conference on Acoustics, Speech and Signal Processing. April, 1998.