Grapheme-to-Phoneme Conversion Based on a Fast TBL Algorithm in Mandarin TTS Systems

Min Zheng[1], Qin Shi[2], Wei Zhang[2], and Lianhong Cai[1]

[1] Computer Science Department in Tsinghua University
100084, Beijing, China
kristy99@mails.tsinghua.edu.cn, clh-dcs@tsinghua.edu.cn
[2] IBM China Research Lab, 100085, Beijing, China
{shiqin, zhangzw}@cn.ibm.com

Abstract. Grapheme-to-phoneme (G2P) conversion is an important subcomponent in many speech processing systems. The difficulty in Chinese G2P conversion is to pick out one correct pronunciation from several candidates according to the context information such as part-of-speech, lexical words, length of the word, or position of the polyphone in a word or a sentence. By evaluating the distribution of polyphones in a large text corpus with correct pinyin transcriptions, this paper points out that correct G2P conversion for 78 key polyphones greatly decrease the overall error rate. This paper proposed a fast Transformation-based error-driven learning (TBL) algorithm to solve G2P conversion. The correct rates of polyphones, which originally have high accuracy or low accuracy, are both improved. After compared with Decision Tree algorithm, TBL algorithm shows better performance to solve the polyphone problem.

1 Introduction

The ability to predict the pronunciation of a written word accurately is an important sub-component within many speech processing systems. This task is typically accomplished through explicit pronunciation dictionaries or Grapheme-to-Phoneme (G2P) rule sets. In most of the alphabetic languages such as English, the main problem G2P module is to generate correct pronunciations for words that are out of vocabulary (OOV). Many approaches for letter-to- phoneme conversion have been proposed [1][2]. However, unlike the OOV problem, the difficulty in Chinese G2P conversion is to pick out one correct pronunciation from several candidates according to the context information such as Part-Of-Speech, lexical words or position of the polyphone in a word or sentence. Traditionally, the commonly used method in most Mandarin TTS systems is to list as many as possible the words with polyphonic characters and their pronunciations into a dictionary. But such dictionary can not solve all the problems about polyphones, summarizing pronunciation rules according to the context is needed to handle more complicated problem. Recently, various data-driven methods including neutral network[3], decision trees[4][5], pronunciation-by-analogy models[6] and extended stochastic complexity methods[7] have been tried to solve this G2P problem.

L. Wang and Y. Jin (Eds.): FSKD 2005, LNAI 3614, pp. 600 – 609, 2005.

In this paper, the TBL algorithm is proposed to solve the G2P problem in mandarin TTS system. As an automatic rule learning methods, it is proved to be efficient for short distance prediction. TBL is widely used in numerous tasks, including learning rules for part-of-speech tagging (Brill, 1995)[8], prepositional phrase attachment (Yeh & Vilain, 1998)[9] and grammatical relation extraction (Ferro, Vilain, & Yeh, 1999)[10] etc; Now we leverage it to solve the polyphone problem and receive great improvements. Besides, a comparative experiment based-on decision tree is done using the same features and corpus that are used in TBL experiment. The decision tree is successfully dealt with some similar problems like parsing (David M. 1995)[11], prosody labeling (Xijun Ma, 2003)[12] and phrase break prediction (Byeong chang Kim, 2000)[13] etc. Compared the two results, the TBL algorithm is shown better performance to solve this polyphone problem.

The paper is organized as follows: The introduction of the TBL algorithm is explained in Section 2. In this section, a fast TBL algorithm for rules learning is also described. By using the algorithm, the training time of a transformation-based learner is speeded up without sacrificing performance. Section 3 shows the analysis and evaluation process of polyphones, including selection of polyphone candidates and corpus preparation. Section 4 describes the experiment, including features selection, template design and algorithm realization. Section 5 gives a comparative experiment based-on decision tree. Final conclusions are given in Section 6.

2 Transformation-Based Learning Algorithm

2.1 Introduction of TBL algorithm

Transformation-based learning (TBL) (Brill,1995) is one of the most successful rule-based machine learning algorithms. The central idea of transformation-based learning (TBL) is to learn an ordered list of rules which progressively improve upon the current state of the training set. The algorithm is illustrated in Figure 1.

1) ***Initial State:*** The training corpus is first standardized into initial state. This initial state is annotated with tagging which has been derived from statistical analysis of the training corpus or limited domain knowledge;
2) ***Template:*** It is composed of several features. The rules search space is limited by the templates. Templates describe valid rules and must describe properties that will reliably indicate when a rule is applicable.
3) ***Rules Generating:*** The error driven model uses each erroneous tag in the training corpus to propose a set of rules by template instantiation.
4) ***Learning Process:*** The TBL algorithm attempts to duplicate the training corpus from the initial state by iteratively learning rules which patch errors in the current state. In each iteration, shown highlighted in Figure 1, the highest scoring rule is appended to the learnt sequence and applied to the current state. Iterations continue until no more improvement can be made.