Abstract

Automatic Speech Recognition (ASR) systems have steadily improved and are now in widespread use whereas, their accuracy continues to lag behind human performance, particularly in adverse conditions. Over the recent years, speech recognition technology has been making steady and significant progress. Speech recognition has often been suggested as a key to the universal information access, since the speech mode is the natural way to interact, does not require literacy. However, success stories of speech interfaces in developing regions are few and far between. The goal of this research described is to determine in what way speech can be handled best in a Speech Recognition System (SRS) and the prime concern is to find effective Confidence measure (CM) so as to minimize the error rate. CM enables us to assess the output of SRS. CM gives us with an approximation of the probability that a word in the recognizer output is either correct or incorrect. CM is a quantitative estimate of a word’s correctness. CM is a number between 0 and 1 indicating our degree of belief that a unit output by a recognizer (phrase, word, phone etc..) is correct. The most important application of CM is in speech dialogue systems (e.g.,) call routing, information provision, ticket booking etc., Errors can be disastrous in a recognition system, but confirmation of each content word is tedious. System can use CM to decide which words are correct and which words need to be confirmed or corrected. Out-Of-Vocabulary (OOV) words are common occurrence in many speech recognition
applications, and are known source of recognition errors.

In this research work, experiences in detecting OOV, error detection and error correction are summarized. A novel approach that incorporates Fuzzy Inference System (FIS) has been designed and tested to detect OOV and In-Vocabulary (IV) words. This work emphasizes OOV detection, confidence scoring and modeling. Evaluating words of discriminating power in ASR system is also described. The notion of characterizing CM in terms of their discriminative power and bias are given importance. Two types of CM are discussed namely Acoustic model based measure and Language Model (LM) based measure. Linear combination is adopted to combine the CM. In addition acoustic CM is been used to detect OOV and IV. The experimental result shows, incorporating FIS with the output of the recognizer gives an acceptable performance and OOV problem is reduced drastically. This work discusses about a new algorithm named Confidence Confirmation Algorithm (CCA) in finding the correctness of the word and also incorporating different modal strategy in error correction. The algorithm has been coined as CCA, since from the output of the recognizer it is not always possible to predict the nature of the correctness of the word. In order to get the confidence, additional information is needed to have a better accuracy of the word. Information for all category words in training corpus is used to define a function that gives an approximation of OOV word emission probability for each word category. Here this information is integrated into LM. In this model, word error rate is better on speech data with better OOV detection rate is achieved.