

Automatic Detection Of New Words In A Large Vocabulary Continuous Speech Recognition System

Ayman Asadi[†]
Richard Schwartz[‡]
John Makhoul[‡]

[†] Northeastern University, Boston, MA 02115

[‡] BBN Systems and Technologies Corporation, Cambridge, MA 02138

ABSTRACT

In practical large vocabulary speech recognition systems, it is nearly impossible for a speaker to remember which words are in the vocabulary. The probability of the speaker using words outside the vocabulary can be quite high. For the case when a speaker uses a new word, current systems will always recognize other words within the vocabulary in place of the new word, and the speaker wouldn't know what the problem is.

In this paper, we describe a preliminary investigation of techniques that automatically detect when the speaker has used a word that is not in the vocabulary. We developed a technique that uses a general model for the acoustics of any word to recognize the existence of new words. Using this general word model, we measure the correct detection of new words versus the false alarm rate.

Experiments were run using the DARPA 1000-word Resource Management Database for continuous speech recognition. The recognition system used is the BBN BYBLOS continuous speech recognition system (Chow *et al.*, 1987). The preliminary results indicate a detection rate of 74% with a false alarm rate of 3.4%.

1 THE NEW WORD PROBLEM

The current continuous speech recognition systems are designed to recognize words within the vocabulary of the system. When a new word is spoken they recognize other words that are in the vocabulary in place of the new word. When this happens, the user does not know that the problem is that one of the words spoken is not in the vocabulary. He assumes that the system simply mis-recognized the word, and therefore he says the sentence

again and again. The current systems do not tell the user what the problem is, which could be very frustrating.

Adding the ability to detect new words automatically can be very efficient and will improve the performance of the system. Once a new word is detected it is possible to add the word to the vocabulary with some extra information from the user such as repeating the word with a carrier phrase and typing in the spelling of the word.

2 APPROACH

The obvious zero-order solution for this problem is to apply some rejection threshold on the word score. If the score reaches a level higher than the threshold then a new word is detected. However, when we examined the scores of words in a sentence, we found that the score of correct words varies widely, making it impossible to tell whether a word is correct or not. Therefore, this approach for detecting new words did not work well.

Our proposed solution is to develop an explicit model of new words that will be detected whenever a new word occurs. The word model should be general enough to represent any new word. It should score better than other words in the vocabulary in place of new words only. It should not appear in place of already existing words in the vocabulary. We tried two acoustic models of new words which are described below.

The first word model we tried was a new word model with a minimum of four phonemes long. It is a linear word model of 5 states and 4 identical phonemes with flat spectral distribution. The results were not encouraging due to the high false alarm rate and low detection rate.

The second word model that we tried was a word model that allows for any sequence of phonemes of at least two phonemes long. The model has 3 states, all

phonemes in parallel from the first state to the second state, all phonemes in parallel from the second state to the third state and all phonemes in parallel looping on the second state. All phonemes are context independent phonemes. Note that this is in contrast to the normal vocabulary of the system, which uses context dependent phoneme models.

We used a statistical class grammar to make the detection process more useful, and created a new word model for each open class. Open classes are the classes that accept new words (e.g. ship names, port names) as opposed to closed classes that do not accept new words (e.g. months, week-days, digits). By using separate new word models for the open classes we can make the distinction whether the new word was a ship name or a port name, etc. Also, it is easy to add the open class words to statistical class grammars and to Natural Language syntax and semantics.

3 EXPERIMENTS AND RESULTS

We presented new words to the system, simply by removing words, that occur in the test sentences, from the vocabulary. We give results for experiments that used the three state acoustic model for new words. The experiments were run on 7 speakers, 25 test sentences per speaker, from the May 88 test which are BEF, CMR, DMS, DTB, DTD, JWS and PGH. We varied the perplexity of the statistical class grammars simply by changing the number of training sentences. A bias against new words was implemented to reduce the false alarm rate.

Our first experiment was detecting new ship names or new ship names possessive. The perplexity of the grammar was 100. We had a detection rate of 83% and a false alarm rate of 1.7%. In the second experiment we changed the perplexity of the grammar to 60 to measure the effect of the perplexity on the detection rate and the false alarm rate. There was no significant difference in the detection rate (84%) but the false alarm rate dropped to 1.1%.

Our third experiment was detecting new port names with grammar of perplexity 100. We had a detection rate of 64% and a false alarm rate of 0.6%.

In the fourth experiment we tried to detect new words from 7 different classes which are ship name, ship name possessive, port name, water name, land name, capability and track name. The grammar had a perplexity of 100. The detection rate was 74% and the false alarm rate was 3.4%.

We measured the detection rate versus the false alarm rate for each experiment. The results are tabulated in table 1. The columns in the table are described then followed by an example as an illustration:

- new words: the new words were allowed in the following classes.
- perp: the perplexity of the grammar.
- cr: exact detection rate as a percentage of number of new words.

SENTENCE (1237)
REF: how many LAMPS
HYP: how many NEW-CAPABILITY

REF: cruisers are in
HYP: cruisers are in

REF: MOZAMBIQUE channel
HYP: NEW-WATER-NAME channel

- cc: close call or close detection rate. That is, the new word was detected but there was an insertion or deletion in its vicinity.

SENTENCE (0464)
REF: when+s HAWKBILL DUE in
HYP: when+s NEW-SHIP-NAME * in**

REF: port
HYP: port

- sw: switch between classes, i.e. the new word was detected, but was assigned to the wrong class.

SENTENCE (1006)
REF: when was PEORIA last
HYP: when was NEW-SHIP-NAME+S last

REF: in the atlantic ocean
HYP: in the atlantic ocean

- dr: total detection rate. Sum of cr, cc and sw.
- far: false alarm rate, percentage of number of false alarms to the total number of test sentences. A false alarm is a new word detected where there was no new word in that part of the test sentence.

SENTENCE (1025)

REF: WHEN DID sherman last
HYP: WHEN+S THE sherman last

REF: downgrade for asuw
HYP: downgrade for asuw

REF: mission AREA
HYP: mission NEW-TRACK-NAME

References

Y.L. Chow, M.O. Dunham, O.A. Kimball, M.A. Krasner, G.F. Kubala, J. Makhoul, P.J. Price, S. Roucos, and R.M. Schwartz. "BYBLOS: The BBN Continuous Speech Recognition System". In *IEEE Int. Conf. Acoust., Speech, Signal Processing*, pages 89-92. Dallas, TX, April 1987. Paper No. 3.7.

new words	perp	cr	cc	sw	dr	far
shipname(+s)	100	42	36	5	83	1.7
shipname(+s)	60	49	30	5	84	1.1
portname	100	27	37	-	64	0.6
7 classes	100	44	6	24	74	3.4

Table 1: Detection of new words results

While we would like the system to detect the exact location and the class of the new words, it is also useful to simply detect that a new word occurred. Thus we say that 74% of the time the system was able to inform the user that a new word had been used.

4 CONCLUSION

From the above results we conclude that the problem of detecting new words can be solved by selecting an appropriate word model for new words. In these experiments we have proved that this approach is viable and results in a detection rate of 74% and a false alarm rate of 3.4%. These results also suggest that a better word model can be used to enhance the detection rate. The use of a bias helps reduce the false alarm rate without affecting the detection rate significantly. Changing the perplexity of the class grammar did not affect the detection rate significantly but it reduced 35% of the false alarm rate.

Acknowledgements

The work reported here was supported by the Advanced Research Projects Agency and was monitored by the Office of Naval Research under Contract No. 00014-89-C-0008.