한국어 가격 문장인식을 위한 FSN의 개선된 문법적 구조
An Improved Grammatical Structure of the FSN
for the Recognition of Korean Price Sentences

김동주, 홍광석
성균관대학교 정보통신공학부 휴먼컴퓨터연구실

Dong-Ju Kim, Kwang-Seok Hong
HCI Lab, School of Information & Communications Engineering, Sungkyunkwan University

요 약

본 논문에서는 실용적인 한국어 가격 문장인식기를 구성하기 위한 유한 상태 망(FSN)의 개선된 문법적인 구조를 제안한다. 기존에 제안된 가격 문장의 문법 제한은 실용적인 한국어 가격 문장의 인식을 위하여 문법 제한과 문법적인 구조를 수정해야 할 필요가 있다. 문법 제한은 실용적 관점에서 한국어 가격 문장의 세 번째와 네 번째 문법 제한을 개선하였다. 본 논문에서는 기존의 문법 제한을 개선하였고, FSN[1]의 문법적인 구조의 단점을 보완하였다. 개선된 문법적 구조의 성능평가를 위하여 FSN0, FSN-1, FSN-2의 세 가지 실험을 하여 평가하였다. 가격 문장의 인식결과, 단어 인식률은 FSN0, FSN1, FSN2의 경우에 각각 81.37%, 83.92%, 85.49%의 성능을 보였다. 또한 문장 인식률은 FSN0, FSN1, FSN2의 경우에 각각 35%, 45%, 52%의 성능을 보였다.

Abstract

In this paper, we present an improved grammatical structure of the finite state network(FSN) for constructing useful recognizer of practical Korean price sentences. The grammar constraints of Korean price sentences in the existing method are necessary to modify grammar constraint and grammatical structure for the recognition of practical Korean price sentences. The grammar constraints are improved in the third and the fourth grammar constraint of Korean price sentences for the practical point. In this paper, we improve the grammar constraints and make up for the weak point in the grammatical structure of the FSN[1]. Three kinds of experiments were performed to evaluate the improved grammatical structures; FSN0, FSN-1, FSN-2. As the recognition results for price sentences, the word recognition rates were 81.37%, 83.92%, and 85.49%, for FSN0, FSN-1, and FSN-2. Also, the sentence recognition rates were 35%, 45%, and 52%, respectively.

key words : price sentence recognition, grammatical structure, finite state network(FSN)

I . Introduction

The price sentences as a recognition target task are one of the useful tasks in lots of practical areas. In restricted domain sentence recognition, the FSN is the most popular method for grammatical structures. J.-P.

won(Korean monetary unit) amount recognition. They implemented a grammatical structure of the FSN for the recognition of Korean price sentences. The grammatical structure used the grammar constraints of Korean price sentences. However, it is

not well suited for practical Korean price sentences, especially, the third and the fourth grammar constraint of Korean price sentences in [1].

All Korean digits are composed of a syllable. Ten Korean digits are such as; 1-/il/, 2-/i:/, 3-/sam/, 4-/sa/, 5-/o/, 6-/yuk/, 7-/chil/, 8-/pal/, 9-/ku/, 0-/kong/. Frequent error patterns in a Korean connected digit recognizer are founded that two digits ('i:' and 'o') are involved in most insertion and deletion errors [2]. Also, Korean digit recognizer has many inexplicit boundaries of the digit and may happen to pronounce the same phone differently by coarticulation effects[3]. In [1], several strict grammar constraints of Korean price sentences are applied. In this reason, the grammatical structure of J.-P. Ham et al. will be showed performance better than practical Korean price sentences.

In this paper, we focus on the grammar constraints of practical Korean price sentences. We have arranged the nodes and the arcs of the FSN properly for practical Korean price sentences. Also, we have make up for the weak points in the grammatical structure of the FSN[1]. Two improved grammatical structures, FSN-1 and FSN-2, are presented and the recognition performances are compared with those of non- grammatical structure FSN0.

II. Improved Grammar Constraints of Korean Price Sentences

The grammar constraints of Korean price sentences are summarized in [1] as follows;

- The basic group is composed of 4 digits.
- The units of the basic group are /sip/(10^1), /baek/(10^2), and /chun/(10^3) : $BG \equiv d_1$ /chun/ d_2 /baek/ d_3 /sip/ d_4, where d_n, $n=1, \cdot \cdot \cdot, 4$, means digit.
- The meta-units are /man/(10^4), /uk/(10^8), and etc.: $\cdot \cdot \cdot BG$ /uk/ BG /man/ BG.
- When '1' is followed by the basic group units, it is not pronounced. Only the unit after it is pronounced.
- Any conjunction such as and is not used.
- The variations of the digits such as -ty and

-teen are not appeared.
- It ends with /won/.

However, for the practical use, it is necessary to make up for the practical point in the third and the fourth grammar constraint of Korean price sentences. There are two differences;

- The meta-units are /man/(10^4), /uk/(10^8), and etc.: $\cdot \cdot \cdot BG$ /uk/ BG /man/ BG. However, in case 10000, it is pronounced /man/ as well as /il/ /man/.
- When '1' is followed by the basic group units, it is pronounced or not. It is pronounced or only the unit after it is pronounced.

In the first difference, an exceptional case 10000 is considered. Also, in the second difference, a practice of utterance for '1' in front of basic group units is considered.

III. Improved Grammatical Structure of FSN

Non-grammatically structured FSN noted as FSN0 is depicted in Fig.1. It is equal to the one pass algorithm without grammar. The arcs of FSN are HMM and represent words to be recognized. The nodes correspond to boundaries of words. Silence model is added to allow silence period within a sentence.
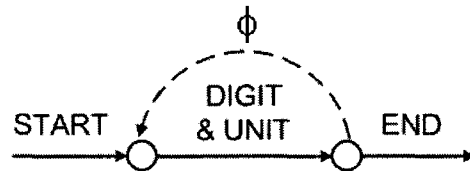

Figure 1. The structure of FSN0 in [1].

For practice Korean price sentence, sub-FSN noted as Digit is designed and substituted it for Digit1 in Fig.2 and Digit2 in Fig.3. Two sub-FSN's noted as Digit1 and Digit2 were designed to cover special rules as mentioned in [1]. Another sub-FSN noted as Unit for the units of the basic group is designed. Two sub-FSN's, noted as Digit and Unit, are shown in Fig.4 and Fig.5.
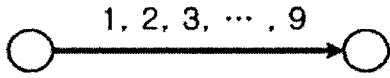
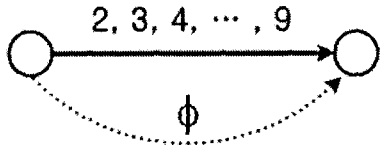Figure 2. The structure of sub-FSN Digit1 in [1].

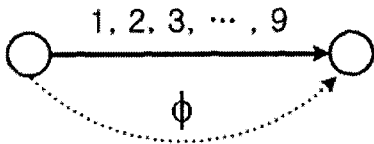Figure 3. The structure of sub-FSN Digit2 in [1].

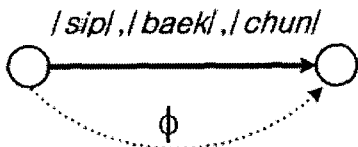Figure 4. The structure of sub-FSN Digit.

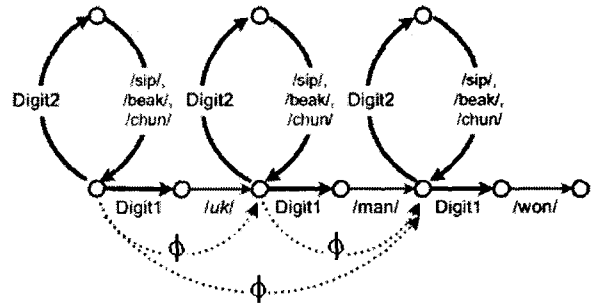Figure 5. The structure of sub-FSN Unit.

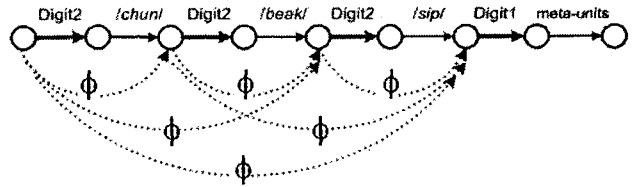Figure 6. The structure of FSN1 in [1].

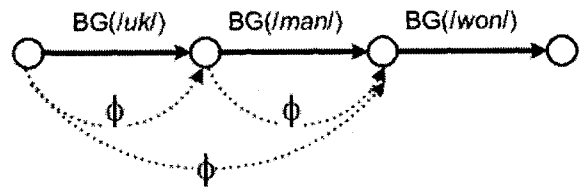Figure 7. The structure of sub-FSN BG (meta-unit) in [1].

Figure 8. The structure of FSN-2(the same as FSN2 in [1]).

Two grammatically structured FSN's noted as FSN1 and FSN2 were designed to cover grammar constraints[1]. The structure of FSN1 in Fig.6 has weak points. FSN1 does not connect /sip/ with /uk/, /baek/ with /uk/, /chun/ with /uk/, /sip/ with /man/, /baek/ with /man/ and /chun/ with /man/, and does not finish with /sip won/, /baek won/ and /chun won/. Also, the structure of FSN2 in Fig.7 and Fig.8 has the same weak points in FSN1. We have to correct these weak points for common use.

In this paper, two kinds of improved grammatical structure of FSN are presented, which will be denoted by FSN-1 and FSN-2. An improved grammatically structured FSN-1 with sub-FSN's is designed as given in Fig.9. FSN-1 has made up for the weak points in the grammatical structures of the FSN1[1]. FSN1 and FSN-1 are not allow that lots of grammatically illegal sentences contain a combination of 'digit+digit', 'unit+unit' or reverse ordered meta-units. However, FSN1 and FSN-1 are allow grammatically illegal sentences contain a combination of reverse ordered units of the basic group. Another improved grammatically structured FSN-2 with sub-FSN's is designed as given in Fig.8 and Fig.10. FSN-2 has made up for the weak points in the

grammatical structures of the FSN2[1]. In FSN-2, another sub-FSN denoted by BG (meta-unit) is used. The improved structure of BG (meta-unit) is depicted in Fig.10. FSN-2 covers all of the improved grammar constraints in practical Korean price sentences as mentioned before.
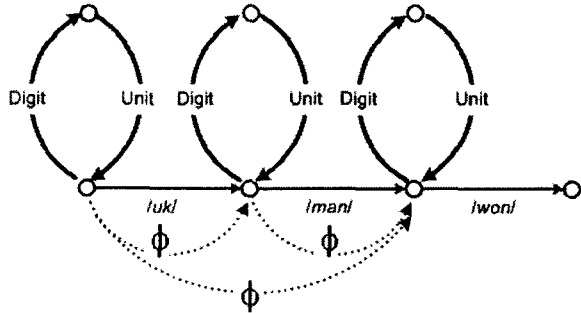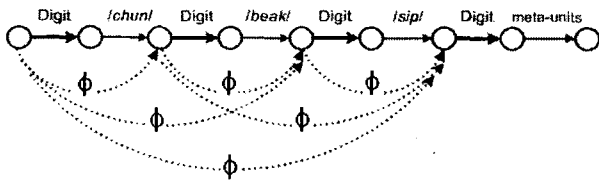

Figure 9. The structure of FSN-1.

Figure 10. The structure of sub-FSN BG (meta-unit).

Ⅳ. Experiments and Results

Three kinds of experiments were performed to evaluate the improved grammatical structures; FSN0, FSN-1, FSN-2. The recognizer of practical Korean price sentences was implemented using vocabulary independent speech recognition system[4]. In [4], recognition system uses CV(consonant vowel), VCCV, VC recognition unit that was trained using very large speech corpus including semi-syllable, continuous digit, name, PBW(phonetically balanced word), etc.

To evaluate the recognition performance of practical Korean price sentences, the speaker-independent price sentence database for test consisted of 10 talkers (5 male, 5 female). Each talker recorded 20 price sentences, which are phonetically balanced and in which all the possible combination of digits and units are included. The

speech signal was sampled at 11.025kHz and applied pre-emphasis with the filter $1 - 0.95\,z^{-1}$ to complement

the energy decrease in the high frequencies of human speech and analyzed every 10ms with Hamming window of 20ms. A 16th order MFCC analysis was performed.

Table 1. Recognition results [%]

	FSN0	FSN-1	FSN-2
Word accuracy	81.37	83.92	85.49
Sentence accuracy	35	45	52

The recognition results for non-grammatically structured FSN0 and two improved grammatically structured FSN-1 and FSN-2 are shown in Table 1. For FSN0, FSN-1, and FSN-2, the word recognition rates were 81.37%, 83.92%, and 85.49%, respectively. Also, the sentence recognition rates were 35%, 45%, and 52%, respectively. In FSN0, a lot of digits and units were substituted in inappropriate places. In FSN-1, a few combination of reverse ordered units of the basic group were appeared. In FSN-2, all of the recognized sentences are grammatically valid sentences. Therefore, as showed in table 1, as the more strict grammar was applied, the better performance was achieved.

Ⅴ. Conclusion

In this paper, an improved grammatical structure of the FSN is proposed for constructing an useful recognizer of practical Korean price sentences. The improved grammatical structures were designed according to the grammar of practical Korean price sentences. Also, it is implemented by arranging the nodes and the arcs of the FSN.

The system was tested in three modes, namely, FSN0, FSN-1, FSN-2, and speaker independent. As the recognition results for Korean price sentences, the word recognition rates were 81.37%-85.49%, and the sentence recognition rates were 35%-52%. These results are encouraging because no special speech

data for training HMM used for price sentence recognizer. With training data for price sentences we expect better performance. Therefore prospective future research areas include training model using speech data for practical Korean price sentences.

References

[1] J.-P. Ham, T.-Y. Yang, C. Lee, and D.-H. Youn, "A Grammatical Structure of the FSN for the Recognition of Korean Price Sentences," IEICE Trans. Inf. & Syst., vol.E84-D, no.11, pp.1577-1579, Nov. 2001.

[2] O.W. Kwon and C.K. Un, "Context-dependent word duration modelling for Korean connected digit recognition," Electron. Lett., vol.31, no.19, pp1630-1631, Sep. 1995.

[3] J.-S. Youn, K.-W. Chung, D.-S. Lee and K.-S. Hong, "Korean Speech Recognition using Semisyllable HMM," Int. Conf. on Signal Processing Application and Technology, vol.2, pp1339-1343, Sep. 1998.

[4] J.S. Youn and K.S. Hong, "An Implementation of the Vocabulary Independent Speech Recognition System Using VCCV Unit." The Journal of the Acoustical Society of Korea, vol.21, no.2, pp160-166, Feb. 2002.

김 동 주(Dong-Ju Kim)
準會員
1998 충북대학교 전파공학과 학사
2000 충북대학교 전파공학과 석사
2001-현재 성균관대학교 정보통신
        공학부 박사과정
관심분야 : 음성인식, 음성 코딩, 신호처리

홍 광 석(Kwang-Seok Hong)
正會員
1985 성균관대학교 전자공학과 학사
1988 성균관대학교 전자공학과 석사
1992 성균관대학교 전자공학과 박사
1990-1993 서울보건전문대학 전산정
        보처리과 전임강사
1993-1995 제주대학교 정보공학과 전임강사
1995-현재 성균관대학교 정보통신 공학부 부교수
관심분야 : 휴먼-컴퓨터 인터페이스, 음성인식 및 합성
E-mail : kshong@skku.ac.kr