문장음성 이해를 위한 확률모델에 관한 연구

노 용 완† · 홍 광 석††

요 약

본 논문에서는 사전과 시소러스를 이용하여 문장음성 이해를 위한 확률모델을 제안한다. 제안한 확률모델은 입력되는 음성과 텍스트 문장에서 단어를 추출한다. 컴퓨터가 선택한 카테고리의 사전 DB와 입력된 문장에서 추출된 단어와 비교하고 확률모델로부터 확률값을 얻는다. 이때 컴퓨터로부터 상위어 정보를 알아내고 상위어 사전을 검색하여 단어를 추출하고 입력된 단어와 확률 모델을 비교하여 결과값을 얻는다. 사전과 상위어 사전으로부터 얻은 두개의 확률값을 더하고 그 값을 미리 정해진 임계값과 비교하여 문장의 이해도를 측정한다. 이와 같은 이해 시스템을 스무고개 게임에 적용시켜 그 성능을 평가 하였다. 상위어 확률 값(α)이 0.9이고 임계값 (β)은 0.38일 때 문장음성 이해의 정확도는 79.8% 였다.

A study on the Stochastic Model for Sentence Speech Understanding

Yong-Wan Roh† · Kwang-Seok Hong††

ABSTRACT

In this paper, we propose a stochastic model for sentence speech understanding using dictionary and thesaurus. The proposed model extracts words from an input speech or text into a sentence. A computer is sellected category of dictionary database compared the word extracting from the input sentence calculating a probability value to the compare results from stochastic model. At this time, computer read out upper dictionary information from the upper dictionary searching and extracting word compared input sentence caluclating value to the compare results from stochastic model. We compare adding the first and second probability value from the dictionary searching and the upper dictionary searching with threshold probability that we measure the sentence understanding rate. We evaluated the performance of the sentence speech understanding system by applying twenty questions game. As the experiment results, we got sentence speech understanding accuracy of 79.8%. In this case, probability (α) of high level word is 0.9 and threshold probability (β) is 0.38.

키워드 : 음성 인식(Speech Recognition), 확률모델(Stochastic Model), 문장음성 이해(Sentence Speech Understanding), 시소러스(Thesaurus), 음성이해 시스템(Speech Understanding System)

1. 서 론

인간의 지적 능력을 모델화하여 지적 시스템을 구축하고자 하는 인공지능의 연구가 활발해 짐에 따라 자연언어 정보처리 분야가 크게 부각되고 있다[1].

인공지능이란 인간의 사고 능력을 컴퓨터로 하여금 실현시키려 하는 것이며 인간의 지적행위의 대상이 되는 것이 지식인 것처럼 인공지능의 주된 대상도 지식이다[2]. 그런데 인간의 지적행위나 사고는 자연언어에 귀착된다고 할 수 있으며, 따라서 인간이 컴퓨터를 사용함에 있어서 마우스나 키보드를 사용하는 것을 자연언어로 대체하고, 컴퓨터가 사용자의 수작업에 의한 프로그램에 의존하지 않고 스스로 생각하며 작동하는 인공지능화 하려는 것이 세계적인 추세이다[3]. 컴퓨터와 인간이 자연언어로 대화하고, 컴퓨터는 인간의 자연언어를 이해함으로써 주어진 문제를 이해하여 해결하도록 하는 인간의 지적인 능력을 컴퓨터에 이식시키려고 하는 것이다[2]. 자연언어 이해 분야는 응용시스템에서 연속음성 인식이 가능하게 된 요즘에 그 비중을 더해가면서 지적인 언어 이해 시스템에 대한 연구가 활발해지고 있다[4]. 인식된 문장을 이해하여 이해된 결과에 의해 판단 및 반응을 하는 것은 초보적 단계에 있으나 많은 기업과 연구소에서 이를 위한 연구가 활발히 진행되고 있다[1, 5].

최근에 사전을 기반으로 한 한국어 의미망 구축과 활용에 대해 사전이나 의미기반의 정보 검색에 대한 많은 연구

※ 본 연구는 한국과학재단 목적기초연구(R05-2002-001007-0)지원으로 수행되었음.
† 준 회 원 : 성균관대학교 대학원 정보통신공학부
†† 종신회원 : 성균관대학교 정보통신공학부 교수
논문접수 : 2003년 5월 19일, 심사완료 : 2003년 11월 4일

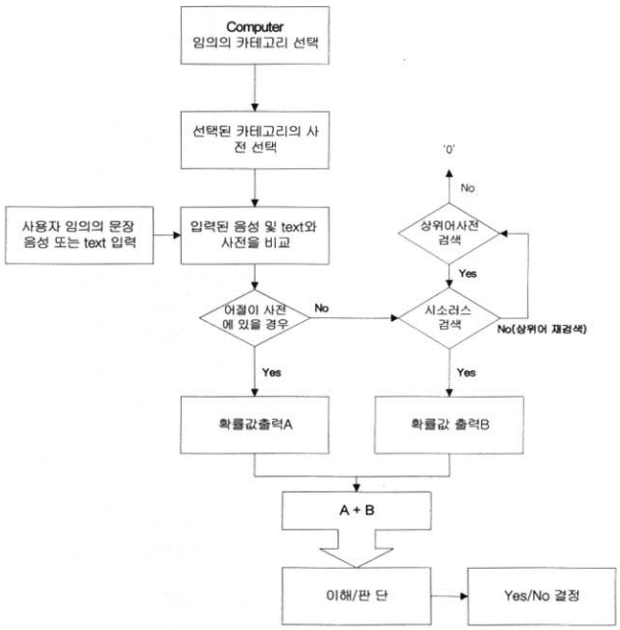
[6, 7]들이 계속되고 있으며 이는 언어자원 구축의 방향을 제시 하고 있다. 객체 지향 시소러스에서 참조 질의 조건 완화 기법의 방식은 객체지향 시소러스의 구조적인 특징을 이용하여 질의 조건을 일반화 시키는 질의 처리 기법이다 [8]. 과거에는 사전만을 사용하거나 시소러스 만을 사용하여 확률값을 얻거나 항상 같은 값의 확률값을 부여하였다.

본 논문에서는 시소러스와 사전을 사용하여 문장을 이해하여 확률값을 산출하는 이해 모델을 제안하였다. 이해 모델로부터 구해진 확률값은 실험을 통해 얻어진 임계값과 비교하고 관련이 있는지의 여부를 판단한다. 본 논문에서 제안한 확률모델의 성능을 확인하기 위하여 문장음성이해 시스템을 구현하였고 스무고개 게임에 적용시켜 보았다. 실제의 스무고개 게임은 영역의 제한이 없지만 본 논문에서는 동물영역으로 제한하여 구현하였다. 동물 스무고개 게임에서 컴퓨터는 선택한 동물의 사전을 검색하여 확률값을 얻어 낸다. 확률값이 주어진 임계값 이상일 경우 관련이 있다고 생각하며 이하일 경우 동물과 관계가 없다고 판단한다. 사전과 시소러스를 관련 분야에 맞게 추가하면 다른 분야도 이해시스템에 적용 가능하도록 설계되었다.

2. 문장 음성 이해 시스템

문장음성 이해 시스템은 문장 텍스트가 입력되면 입력된 문장을 이해하는 것이다. 이해 시스템은 시소러스와 영역별 사전, 상위어 사전을 사용하여 구현되며, 입력되는 문장에 대해서 사전을 검색하고 사전에 없는 경우에 시소러스를 이용하여 문장내의 단어에 대한 상위어를 추출하고, 상위어의 사전을 사용하여 검색한다. 문장 음성 이해 시스템은 컴퓨터가 임의의 카테고리를 선택하여 선택된 카테고리에서 정보를 가져온다. 사용자는 임의의 문장 또는 텍스트를 사용하여 질문을 한다. 질문과 카테고리가 관계가 있으면 예, 관계가 없으면 아니오라고 판단한다. 사전과 상위어 정보를 가진 시소러스를 가지면 어느 분야이든 이해과정에 있어서 적용이 가능하다. 문장 음성 이해 시스템에서는 사용자가 음성으로 카테고리 내용을 질문해서 컴퓨터가 생각한 카테고리를 맞추도록 구현하였다.

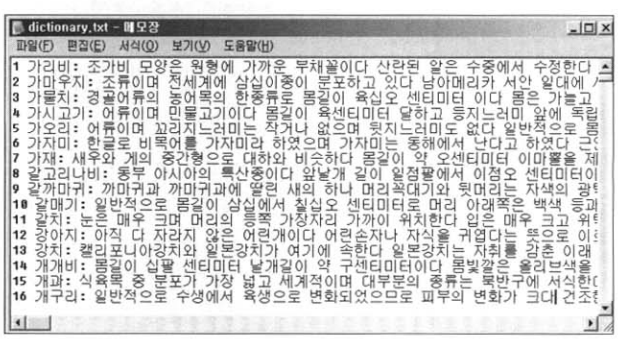
문장음성 이해시스템의 개념도를 (그림 1)에 나타내었다. (그림 1)에서 컴퓨터가 임의의 카테고리를 선택하고 선택한 사전의 정보를 가져온다. 사용자는 스무고개에 관련된 임의의 문장을 음성 또는 텍스트로 질의 한다. 사용자의 질의와 컴퓨터의 사전을 비교하여 확률값 A를 출력하고 사전에 질의의 어절이 존재 하지 않을 경우 시소러스를 검색하여 상위어 정보를 가져오고, 상위어 사전을 비교하여 확률값 B를 출력한다. 출력된 두 개의 확률값을 더한 후 이해 및 판단을 한 다음 예, 아니오를 결정한다.



(그림 1) 문장 음성 이해 시스템의 개념도

2.1 사 전

문장음성 이해 시스템에서 사전은 대용량 사전이나 영역별 사전, 상위어 하위어 사전등 여러가지 사전을 이용하며 컴퓨터가 문장을 이해하는데 기본적으로 필요한 도구이다 [6, 7], 사전의 사용 예로 스무고개 게임에서의 사전 구성을 설명한다. 본 논문에서는 일반 동물 사전과 상위어 사전으로 구성하였다. 일반 동물 사전은 Yahoo 동물 백과사전을 참조하여 구성하였으며 510개의 동물을 선정하였다. 동물 사전에는 510개의 동물들에 대한 설명이 나와있으며 (그림 2)와 같다.



(그림 2) 동물 사전

동물 사전은 맨 처음 동물 번호를 부여하였으며 가나다 순으로 정리하였다. 510종류의 동물 선정은 동물 전공자가 아닌 일반인 5명이 한번이라도 들어본 적 있는 동물만 선택하여 구성하였다. 사정은 약 5만 어절정도로 구성되어 있으며 한 단어를 설명하는 평균어절이 100어절 정도이다.

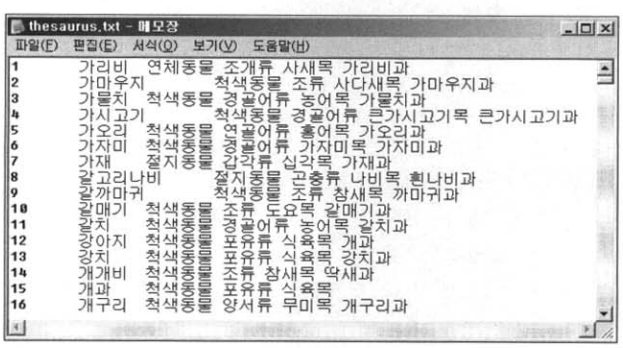
상위어 사전은 동물 사전에 등록되어 있는 510개 동물 상위어에 대한 내용을 뜻풀이 하였다. 시소러스에 동물의 상위어 정보들을 상위어 사전으로 따로 구성하여 뜻풀이를 하였다.

그 예로 시소러스에 호랑이, 척수동물, 포유류, 식육목, 고양이과에서 척수동물, 포유류, 식육목, 고양이과에 대한 사전은 상위어 사전에 나타내었으며 사전과 같은 뜻풀이문 형식으로 되어 있다. 상위어 사전은 일반사전에서 그에 관련된 정보가 없을 때 참조하게 된다. 상위어 사전은 (그림 3)과 같다.

(그림 3) 상위어 사전

2.2 시소러스

시소러스는 사전에 없는 상하위관계를 나타내 주는 것으로 정보 검색에서 일반적으로 많이 사용한다[9, 10]. 시소러스는 전체 시소러스와 영역별 시소러스로 나누어지며 전체 시소러스의 경우 모든 객체를 나타내주기 어렵기 때문에 영역별, 또는 부분 시소러스를 사용한다[11]. 문장 음성 이해 시스템에서는 사전과 시소러스를 사용하여 확률모델을 구성하고 확률값을 얻는다.



(그림 4) 시소러스

시소러스의 한 예로 동물 시소러스는 (그림 4)와 같이 구성한다. 동물 시소러스는 일반 동물 사전의 상위어 표시이며 사전의 상위어를 텍스트로 나타내었다. 시소러스에서 첫 단어는 사전과 동일한 동물이 기재되어 있으며 선정한 동물 510개에 대한 상위어 정보를 기재하였다. 상위어는 척색동물, 포유류, 소목, 소과와 같은 형태 및 순서로 되어 있으며 최상위어부터 하위어 순으로 구성하였다. 한 동물의 상위어들의 최대 개수는 4개이다.

동물 시소러스는 트리형태로 구성되며 트리형태를 바탕으로 상위어 정보가 기재된다.

2.3 확률모델

확률모델은 이해하는 과정의 일부분으로 판정 하기 이전의 단계에서 필요한 도구이다. 이 모델은 이해 판정 하기 위해 필요한 과정이며 확률 임계값을 기준으로 문장의 이해도를 측정하게 된다[15, 16].

$0 < \alpha < 1$, $0 < \beta < 1$ (α : 상위어사전 검색시 확률값, β : threshold)

① $S = (W_1, W_2, \cdots, W_n)$ (S : 입력되는 문장, W: n개의 어절로 이루어 짐)

② $Wi \supset Po$이면 $w_i = W_i - Po$ 여기서, $Po = (Po_1, Po_2, \cdots, Po_k)$, ($P_o$: 조사가 있는 경우, w : 조사가 있는 어절에서 조사를 삭제한 단어)

$D \supset w_i$이면 w_{D_0}, 이것의 확률값 $\Pr(w_{D_0})$ 여기서 $w_i = = (w_1, w_2, \cdots, w_n)$, ($\Pr(w_{D_0})$: 사전에 포함되는 단어의 확률값 ; 포함된 단어 개수$*\frac{1}{n}$(전체 어절의 개수), D : 사전)

③ $TD_1 \supset w_i$이고 $D \not\supset w_i$일때 w_{D_1}, 이것의 확률값 $\alpha \Pr(w_{D_1})$ (TD_1 : 첫 번째 상위어 사전, $\Pr(w_{D_1})$: 첫 번째 상위어 사전에 포함되는 단어의 확률값)

④ $TD_2 \supset w_i$이고 $D \not\supset w_i$이고 $TD_1 \not\supset w_i$일때 w_{D_2}, 이것의 확률값 $\alpha^2 \Pr(w_{D_2})$(TD_2 : 두번째 상위어 사전, $\Pr(w_{D_2})$: 두 번째 상위어 사전에 포함되는 단어의 확률값)

시소러스가 최상위어에 도달할때까지 계속,

⑤ $TD_m \supset w_i$이고 $D \not\supset w_i$이고 $TD_{m-1} \not\supset w_i$ 일 때 w_{D_m}, 이것의 확률값 $\alpha^m \Pr(w_{D_m})$(m : 상위어의 개수, $\Pr(w_{D_m})$: m번째 상위어 사전에 포함되는 단어의 확률값)

⑥ $\Pr(S) = \sum_{j=0}^{m} \alpha^j \Pr(W_{D_j})$)($\Pr(S)$: 한 문장에 대한 확률값, j는 시소러스에서의 상위어 단계)

⑦ 문장에 대한 확률값으로 판단

$\Pr(S) \geq \beta$이면 예(관련 있음), 아니면 아니오(관련 없음)

2.4 판 단

판단 및 결과 과정에서는 확률모델에 의해 나온 확률값 ($\Pr(S)$)를 사용하여 사용자 문장을 컴퓨터가 이해 판단한다. 판단의 임계값 β를 사용하여 예, 아니오를 판단한다.

본 알고리즘을 스무고개 게임에 적용시켜 β가 임계값 이상인 경우 예, 임계값 이하인 경우 아니오라고 판단한다. β를 결정하는 것은 실험 및 결과에서 초기치 확률을 정하는데 기반으로 하였다. α의 값을 얼마로 정하느냐에 따라 β의 값도 바뀌므로 어느 분야를 이해 하느냐에 따라 최적의 확률값을 구하기 위해 α와 β를 변화시켜서 가장 높은 이해도를 얻을수 있게 하는 α와 β를 실험을 통해 구한다.

3. 스무고개 게임을 위한 음성인식

스무고개의 질문은 50명에게 실제로 스무고개를 하도록 하여 받아 적은 후 정리해서 1,839개의 문장을 얻었으며 동일한 질문이나 상식 밖의 질문을 추려내 956개의 문장을 선별하였다. 스무고개 시스템은 41개의 동물, 곤충의 이름이 등록되어져 있으며 956개의 문장을 이용하여 사용자가 선택한 동물의 이름을 맞추도록 구성하였다[12].

놀이방법은 컴퓨터가 임의의 동물, 곤충을 선택한 후 사용자로부터 956개의 문장 가운데 원하는 한가지 질문 형태를 받아 들여 문장을 인식하고, 선택된 동물과 관련이 있으면 /예/, /아니오/라고 합성음을 들려준다. 사용자는 스무번안에 컴퓨터가 선택한 답을 맞추어야 하며, 사용자가 정답을 알면 /답은 ***입니다/라는 문장으로 발성하여 정답을 맞춘다[12].

3.1 어절 단위 문장인식

본 논문에서는 문장음성 인식을 위하여 다음과 같은 어절 후보 규칙을 적용하여 모델을 구성하였다.

① 체언과 조사가 함께 연결된 어절 단위를 기본 후보 어절 단위로 한다. 예를 들면 /목적+으로/, /육식+을/, /모양+의/과 같은 경우이다.
② 체언과 체언, 체언과 기본 후보어절 단위로 연결된 각각의 어절들을 그룹으로 묶어 후보 어절 단위로 한다. 예를 들면 /이+동물은/, /육식+동물/, /서울+시내에서/의 경우 같은 경우이다.
③ 두 음절 이하의 용언 + 체언의 경우에도 그룹핑하여 후보 어절 단위로 한다. 예를 들면 /큰+동물/, /작은+동물/, /센+동물/의 경우이다.
④ 마지막 어절에서 '명사+입니까'인 경우에는 /명사/+/입니까/로 분리한다.

이상과 같은 규칙을 적용하지 않을 경우 어절의 총 수는 1,151개였으며, 규칙을 적용한 경우의 총 어절 수는 1,095개로 나타나 후보의 수를 줄일 수 있었다.

3.2 음절수를 이용한 인식 후보 감소

음절 개수 추출은 기준모델 작성시 인식후보의 개수를 줄일 수 있다. 먼저 입력 데이터로부터 에너지와 영 교차율을 이용하여 유성음과 무성음영역을 검출한다. 또한 좀더 정확한 유성음 영역을 검출하기 위해 제 1포먼트, 제 2포먼트가 존재하는 215Hz와 2,756Hz사이의 에너지 정보를 이용하여 안정된 유성음 영역을 추출하여 음절수를 정하였다[12].

음성인식 시스템의 인식단위는 부단어 단위인 CV(consonant Vowel), VCCV, VC를 사용하였다. 특징 벡터는 16차 멜캡스트럼 특징 파라미터와 그것의 일차 미분계수를 사용한다.

인식을 위한 음향 모델은 DHMM(discrate hidden Markov model)을 사용하였다.

스무고개 시스템에 사용된 총 어절 수는 1,095개이며, 첫 어절에 나올 수 있는 목록의 수는 561개이다. 따라서 첫 어절과 561개의 어절과 비교하게 되면 인식률의 저하 및 인식 시간이 상당히 소요된다. 따라서 음절개수 추출 방법에 따라 첫 어절의 음절 분포 사전을 구축을 한 다음, 입력 된 음성 데이터의 첫 어절의 음절 개수를 추출하여 선택된 음절 개수에 해당하는 단어만을 후보로 두고 인식하도록 시스템을 구현하였다.

〈표 1〉 첫 어절의 음절 개수에 따른 인식후보 감축율

음절개수의 종류	1음절	2음절	3음절	4음절	5음절	6음절	7음절
감축률	9.98%	47.8%	66.1%	39.0%	16.8%	6.06%	1.6%

<표 1>은 첫 어절의 후보 561개와 비교하여 음절개수에 따라 후보를 설정하는 방법에 따른 인식 후보 감소율을 나타내고 있다. 3음절을 제외하고 다른 모든 음절에서 50%의 미만의 감축율을 가져왔으므로 실제로 인식 시간 향상이 기대된다.

3.3 언어 모델

현재의 어절과 다음 어절과의 구문규칙을 적용하여 구성하였으며 그 과정은 다음과 같다.

① 956개 문장의 모든 어절 후보 1,095개에 대해 인덱스를 부여하고 reference 목록에 저장한다. 이 방식은 스무고개 시스템을 구현하기 위해 어절 단위의 기준모델을 메모리에 로딩하기 위해서 번호를 부여한다.
② 각각의 어절에 따라 현재의 어절과 다음에 오는 어절과의 문맥관계에 따라 그룹을 작성한다. 그룹의 규칙은 먼저 현재의 어절이 속한 reference 목록의 순번 인덱스, 다음 어절에 나올 수 있는 후보의 수, 다음 어절 후보의 그룹의 순번으로 하며 그룹의 수는 문장을 구성하고 있는 최대 어절이 5개이기 때문에 5개의 그룹을 구성한다.
③ 첫 번째 어절의 그룹1은 후보가 reference 목록과 같은 번호를 가지며, 두 번째 어절과의 문맥관계를 저장한다. 예를 들어 문장이 한 어절로 이루어진 문장인 /무섭습

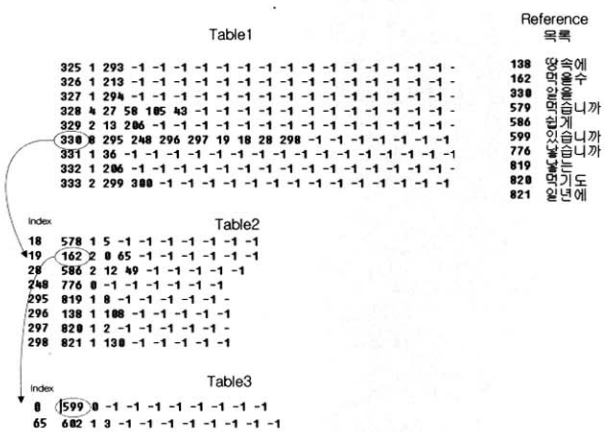
니까/는 reference 번호가 171이라고 하면 171 0 -1 -1 …의 순으로 저장된다. 첫 번째 번호는 reference 목록의 번호이고 다음의 후보가 없기 때문에 두 번째 값이 0으로 나타낸다. 다음 번호들은 -1의 초기화값을 가지게 된다. 두 번째로 /답은/이라는 어절인 경우 다음의 후보가 41개의 동물과 곤충의 이름이 후보로 올 수 있으므로 <표 2>와 같이 저장된다. 442는 reference 목록의 번호이고 41은 다음 어절의 후보의 수가 41개가 있다는 것을 의미하며, 그 다음의 번호는 그룹2에 동물과 곤충의 이름이 저장된 번호를 나타내고 있다.

〈표 2〉 Reference 목록의 예

442 41 363 364 365 366 367 368 369 370 371 372 373 374 375 312 346 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 -1 -1

④ 인식된 첫 번째 후보로부터 그룹2의 순번을 가져오면 맨 처음 번호의 값들을 불려 들여 두 번째 어절 단위와 비교하여 인식을 수행한다.

⑤ 인식된 어절들을 어절 순서에 따라 텍스트창에 표시한다. (그림 5)에 인식 과정을 도시하였다.



(그림 5) "알을 먹을 수 있습니까"의 인식 과정

(그림 5)는 /알을 먹을 수 있습니까/라고 발성한 문장을 테이블에 따라 인식하는 과정을 나타내고 있다. /알을/의 reference 번호는 330, /먹을 수/는 162, /있습니까/는 599번이다.

그룹1에서 /알을/ 인식하게 되면, 그룹2에서 8개의 후보와 비교하여 인덱스 19번 /먹을 수/가 인식되었고, 마지막 테이블3에서 인덱스 0번과 65와 비교하여 reference 번호 599인 /있습니까/가 인식되는 과정을 설명하고 있다.

3.4 복잡도(perplexity)

복잡도는 언어모델의 복잡한 정도를 나타내는 것으로 일반적으로 사용하는 언어 모델에서의 객관적인 평가척도를 말한다[6]. 본 논문에서 실험에 사용한 문장에서는 11.027이고, 음절 개수 추출 알고리즘을 사용하면 복잡도가 8.47 로써 23% 감소하였다.

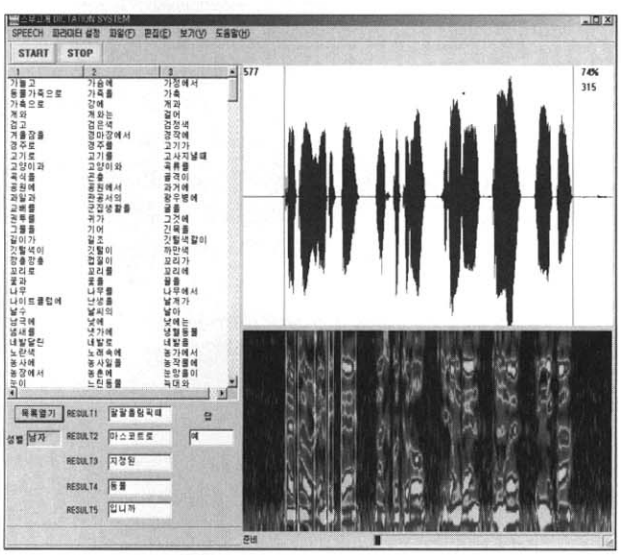
4. 실험 및 결과

본 논문에서의 실험은 3가지 형태로 이루어진다. 첫 번째 실험은 음성인식용 스무고개 시스템을 구현하여 단어 인식률 및 문장 인식률을 측정하였다. 두 번째는 문장음성 이해 확률 모델에서 사용할 최적의 α값과 β값을 구하는 실험이다. 세 번째 실험은 두 번째 실험에서 구해진 최적의 확률값을 사용하여 스무고개 문장의 판단 정확도를 측정하였다.

4.1 음성인식 스무고개 시스템

음성인식용 스무고개 시스템에서는 총 956 종류의 문장을 인식 가능하도록 하였다. 여기서 사용한 음성 인식기는 CV(Consonant + Vowel), VCCV, VC단위 HMM(Hidden Markov Model) 기반의 화자독립 가변어휘 음성인식기를 사용하여 구현하였다[12].

(그림 6)에 음성인식용 스무고개 시스템의 실행화면을 나타내었다. 먼저 목록열기를 선택하여 인식어절 모델을 메모리에 로딩을 하면, 리스트 박스에 후보 어절 목록이 나타난다. 스타트 버튼을 누르면, 컴퓨터는 임의의 문제단어를 선택한다. 사용자가 문장을 발성하면, 먼저 어절 분할을 한 후, 첫 어절의 음절 개수를 추출한다.



(그림 6) 음성인식용 스무고개 시스템의 실행화면

추출된 음절개수로부터 음절 개수 표에 등록되어 있는 어절후보와 비교하여 인식한다.

그림에서는 /팔팔올림픽때 마스코트로 지정된 동물 입니

까/라고 발성한 것에 대한 결과이다. 인식된 어절의 결과가 에디터 박스에 어절에 따라 /팔팔올림픽때/, /마스코트로/, /지정된/, /동물/, /입니까/라고 쓰여지며, 문장음성 이해, 판단의 결과를 답에 표현하고 있다.

어휘독립 음성인식 시스템에 언어모델과 음절 개수 추출 알고리즘을 추가한 스무고개 시스템의 문장인식 성능 평가 실험을 하였다. 실험 방법은 화자 7명이 임의의 100개의 문장을 발성하였으며, 첫 어절 인식률, 총어절 인식률, 문장단위 인식률을 <표 3>에 나타내었다.

오 인식된 문장을 살펴보면, /털로 옷을 만들 수 있습니까/가 /털로 옷을 만들기도 합니까/로 유사음절이 포함되어져 있는 문장에서 오인식되었다.

〈표 3〉 스무고개 시스템의 문장단위 인식 성능

화 자	첫어절 인식률(%)	총어절 인식률(%)	문장단위 인식률(%)
화자 1	87	91.64	87
화자 2	86	89.10	85
화자 3	88	90.48	88
화자 4	87	89.33	86
화자 5	92	96.30	92
화자 6	84	85.90	83
화자 7	88	91.02	88
평 균	87.42	90.15	87.00

4.2 확률모델 평가

확률모델을 평가하는 실험에서는 1,000문장의 DB가 사용되었으며 판단의 기준이 되는 α와 β를 변화시키면서 최적의 확률값을 구하는 실험을 하였다. β를 0부터 0.5까지 0.02의 간격으로 하고 α을 0.5부터 1까지 0.05간격으로 변화 시켜 최적의 확률값을 구하여 α와 β 값을 찾는다.

실험 과정은 문장 10개를 추출하여 동물 510 종류에 대해 참값이 분포도와 거짓의 분포도를 구하였다. 총 5,100개의 질문에 대해 참인 경우 1,087개 거짓인 경우 4,013개를 이용하였다. 상위어 값은 0.5일 때부터 1까지 0.05간격으로 11개, 판단의 기준은 0.02부터 0.52까지 0.02간격 나누어 실험하였다.

실험에 사용한 스무고개 문장이 일반적으로 /예/보다는 /아니오/를 답하는 문장이 많아서 /아니오/를 답하는 문장이 3.7배 많았다.

<표 4>는 모든 상위어에 대한 참값의 분포도를 나타낸 것이며 5,100문장 중 참인 1,087개의 문장에서 상위어 각각에 대한 확률값 분포를 나타낸 것이다.

만일 상위어의 확률이 0.75일 경우 판단 초기치가 0.17일 때 참인 문장이 참으로 판단되는 경우는 424개이고 참인 문장이 거짓으로 판단되는 경우 663개가 된다.

<표 5>는 모든 상위어에 대한 거짓값의 분포도를 나타낸 것이며 4,013문장에서 상위어 각각에 대한 확률값 분포를 나타낸 것이다.

〈표 4〉 상위어 각각에 대한 참 값의 분포

β \ α	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
0	146	146	146	146	146	146	146	146	146	146	146
0.02	49	40	38	0	0	0	0	0	0	0	0
0.04	0	9	11	38	0	0	0	0	0	0	0
0.06	16	0	0	0	38	0	0	0	0	0	0
0.08	452	16	16	11	11	0	37	0	0	0	0
0.1	0	452	0	16	0	38	0	0	0	0	0
0.12	328	0	452	0	16	11	9	46	0	0	0
0.14	0	328	0	452	0	16	2	2	0	0	0
0.16	0	0	0	0	452	452	17	17	38	0	0
0.18	0	0	328	0	0	0	0	0	9	0	0
0.2	0	0	0	328	0	0	452	452	18	47	0
0.22	0	0	0	0	0	0	0	0	452	18	0
0.24	61	61	61	61	389	389	61	61	61	61	124
0.26	0	0	0	0	0	0	0	0	0	0	0
0.28	0	0	0	0	0	0	0	0	0	0	0
0.3	0	0	0	0	0	0	328	328	0	452	0
0.32	12	12	12	12	12	12	12	12	12	12	466
0.34	0	0	0	0	0	0	0	0	0	0	0
0.36	0	0	0	0	0	0	0	0	0	0	0
0.38	0	0	0	0	0	0	0	0	0	0	0
0.4	0	0	0	0	0	0	0	0	328	0	0
0.42	0	0	0	0	0	0	0	0	0	328	0
0.44	0	0	0	0	0	0	0	0	0	0	0
0.46	0	0	0	0	0	0	0	0	0	0	0
0.48	0	0	0	0	0	0	0	0	0	0	0
0.5	23	23	23	23	23	23	23	23	23	23	351

〈표 5〉 상위어 각각에 대한 거짓 값의 분포

β \ α	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
0	2797	2797	2797	2797	2797	2797	2797	2797	2797	2797	2797
0.02	158	62	0	0	0	0	0	0	0	0	0
0.04	56	140	62	32	0	0	0	0	0	0	0
0.06	200	12	96	164	32	32	0	0	0	0	0
0.08	312	200	56	18	104	0	0	0	0	0	0
0.1	0	312	178	178	56	104	32	32	0	0	0
0.12	351	0	334	0	200	256	74	0	0	0	0
0.14	0	351	0	334	0	0	30	104	32	0	0
0.16	0	0	102	102	312	436	234	234	0	0	0
0.18	0	0	249	0	124	0	0	0	74	0	0
0.2	0	0	0	249	0	0	436	436	178	106	0
0.22	0	0	0	0	0	0	0	0	132	178	0
0.24	84	84	84	84	333	333	84	106	140	186	470
0.26	0	0	0	0	0	0	22	0	312	30	0
0.28	0	0	0	0	0	0	0	0	0	0	0
0.3	0	0	0	0	0	0	0	0	0	368	0
0.32	30	30	30	30	30	30	279	279	52	30	428
0.34	0	0	0	0	0	0	0	0	0	0	0
0.36	0	0	0	0	0	0	0	0	22	0	0
0.38	0	0	0	0	0	0	0	0	0	0	0
0.4	0	0	0	0	0	0	0	0	0	22	0
0.42	0	0	0	0	0	0	0	0	249	22	0
0.44	0	0	0	0	0	0	0	0	0	249	0
0.46	0	0	0	0	0	0	0	0	0	0	0
0.48	0	0	0	0	0	0	0	0	0	0	0
0.5	25	25	25	25	25	25	25	25	25	25	318

〈표 6〉 상위어 0.9에 대한 초기 확률 설정

0.90 일때	참 값->참	참->거짓	거짓->거짓	거짓->참	판단이 맞은것
0	1087	0	0	4013	1087
0.02	941	146	2797	1216	3738
0.04	941	146	2797	1216	3738
0.06	941	146	2797	1216	3738
0.08	941	146	2797	1216	3738
0.1	941	146	2797	1216	3738
0.12	941	146	2797	1216	3738
0.14	941	146	2797	1216	3738
0.16	941	146	2829	1184	3770
0.18	903	184	2829	1184	3732
0.2	894	193	2903	1110	3797
0.22	876	211	3081	932	3957
0.24	424	663	3213	800	3637
0.26	363	724	3353	660	3716
0.28	363	724	3665	348	4028
0.3	363	724	3665	348	4028
0.32	363	724	3665	348	4028
0.34	351	736	3717	296	4068
0.36	351	736	3717	296	4068
0.38	351	736	3739	274	4090
0.4	351	736	3739	274	4090
0.42	23	1064	3739	274	3762
0.44	23	1064	3988	25	4011
0.46	23	1064	3988	25	4011
0.48	23	1064	3988	25	4011
0.5	23	1064	3988	25	4011
0.52	0	1087	4013	0	4013

〈표 7〉 α와 β에서 올바른 판단을 한 개수

β \ α	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
0	1087	1087	1087	1087	1087	1087	1087	1087	1087	1087	1087
0.02	3738	3738	3738	3738	3738	3738	3738	3738	3738	3738	3738
0.04	3847	3760	3700	3738	3738	3738	3738	3738	3738	3738	3738
0.06	3903	3891	3751	3732	3738	3738	3738	3738	3738	3738	3738
0.08	**4087**	3903	3847	3896	3732	3770	3738	3738	3738	3738	3738
0.1	3947	**4087**	3887	3903	3825	3770	3701	3738	3738	3738	3738
0.12	3947	3947	**4065**	**4065**	3881	3836	3733	3770	3738	3738	3738
0.14	3970	3947	3947	**4065**	**4065**	**4081**	3798	3724	3738	3738	3738
0.16	3970	3970	3947	3947	**4065**	4065	3826	3826	3770	3738	3738
0.18	3970	3970	4049	4049	3925	4049	4043	4043	3732	3738	3738
0.2	3970	3970	3970	4049	4049	4049	4043	4043	3797	3738	3738
0.22	3970	3970	3970	3970	4049	4049	4027	4027	3637	3797	3738
0.24	3970	3970	3970	3970	4049	4049	4027	4027	3716	3957	3738
0.26	3993	3993	3993	3993	3993	3993	4050	**4072**	4028	**4082**	**4084**
0.28	3993	3993	3993	3993	3993	3993	**4072**	**4072**	4028	4112	**4084**
0.3	3993	3993	3993	3993	3993	3993	**4072**	**4072**	4028	4112	**4084**
0.32	3993	3993	3993	3993	3993	3993	3744	3744	4068	4028	**4084**
0.34	4011	4011	4011	4011	4011	4011	4011	4011	4068	4046	4046
0.36	4011	4011	4011	4011	4011	4011	4011	4011	4068	4046	4046
0.38	4011	4011	4011	4011	4011	4011	4011	4011	**4090**	4046	4046
0.4	4011	4011	4011	4011	4011	4011	4011	4011	**4090**	4046	4046
0.42	4011	4011	4011	4011	4011	4011	4011	4011	3762	4068	4046
0.44	4011	4011	4011	4011	4011	4011	4011	4011	4011	3762	4046
0.46	4011	4011	4011	4011	4011	4011	4011	4011	4011	4011	4046
0.48	4011	4011	4011	4011	4011	4011	4011	4011	4011	4011	4046
0.5	4011	4011	4011	4011	4011	4011	4011	4011	4011	4011	4046
0.52	4013	4013	4013	4013	4013	4013	4013	4013	4013	4013	4013

<표 6>는 α를 0.9로 했을 때의 5,100개 문장 판단 결과를 나타내었다.

<표 7>는 각각의 상위어 마다 판단이 맞는경우를 표시한 것으로 참값 -> 참값, 거짓 -> 거짓의 경우가 올바른 판단으로 0.38과 0.40에서 가장 높은 확률 4090/5100을 가졌다. 여기서 앞뒤값과 오차가 적은 0.38값을 β값으로, α를 0.9로 정하였다.

<표 8>에서 나오는 값들이 <표 6>의 방식에 의해 각각 계산 되었다.

4.3 문장음성이해 스무고개 시스템

앞절에서 결정한 확률값을 사용하여 문장음성 이해 실험을 하였다. 많이 사용하는 200 종류 동물 단어에 대해 문장음성 이해 판단률을 구하였다.

입력되는 문장이 참인 문장일 때 예로 판단을 하는 경우와 거짓인 문장일 때 아니오라고 판단하는 경우가 올바른 판단이며, 반대의 경우는 거짓을 나타낸다. 즉 참 -> 참, 거짓 -> 거짓의 경우 올바른 판단, 참 -> 거짓, 거짓 -> 참이 잘못된 판단이다. 문장음성이해 판단 결과를 <표 8>에 나타내었다. 문장음성이해 판단 결과는 총 102,000 종류 판단 결과 값을 나타낸 것이다. 정확도는 81425/102000로 79.8%의 판단 정확도를 나타내었다.

〈표 8〉 문장이해 판단결과

	참 -> 참	참 -> 거짓	거짓 -> 거짓	거짓 -> 참
문장수	17,359	7,091	64,066	13,484

5. 결 론

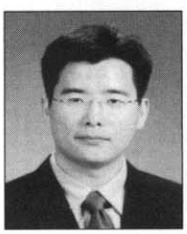
본 논문에서는 문장음성 이해 분야에 대해 연구하였으며 이해 및 판단 모델을 확률적인 모델로 제안하여 구현하였다. 또한 제안한 확률적 모델의 성능평가를 위해 스무고개 게임을 구현하였다. 스무고개 게임은 문장음성 이해모델을 사용한 것으로 컴퓨터가 생각하고 있는 동물을 사람이 맞추어 나가면서 운영 하도록 하였다.

입력된 문장을 이해하기 위하여 동물사전, 상위어 사전, 시소러스를 이용하였고 판단을 하기 위해 확률모델을 사용하였다. 최적의 확률모델을 얻기 위해서 α값과 β값을 변화시켜 α값은 0.9, β가 0.38의 값을 최적의 값으로 얻었으며 판단률을 79.8%였다. 향후 판단의 정확도를 위해 사전의 확장이라든지 이해모델을 개선하여 더 높은 정확도를 얻을 수 있도록 개선하여야 한다.

참 고 문 헌

[1] 이정민, "자연어 처리와 인지", 인지과학, Vol.3, No.2, pp.161-

176, 1992.
[2] 한광록, "한국어 문장이해를 위한 가변패턴네트의 구성과 응용", 정보처리논문지, 제2권 제2호, pp.229-236, 1995.
[3] 김영택, "자연언어처리 기술", 전자공학회지, pp.205-212, 1987.
[4] 이정민 "자연언어처리와 인지", 인지과학, Korean Journal of Cognituve Science, Vol.3, No.2, pp.161-196, 1992.
[5] Biing-Hwang Juang, Sadaoki Furui, "Automatic Recognition and Understanding of Spoken Language-A Frist Step Toward Natural Human-Machine Communication," Proceedings of the IEEE, Vol.88, No.8, pp.1142-1165, August, 2000.
[6] 남지순 "한국어 전자사전", 전자공학회지, 제24권 제9호, pp. 1103-1125, 1997.
[7] 최병진, 이윤재, 이재성, 최기선, "기계가독형 사전 구축을 위한 사전항목의 논리 구조", 인지과학, Vol.7, No.2, pp.75-92, 1996.
[8] 김정애, 박종민, 김원중, 양재동, "객체지향 시소러스에서의 참조 질의 조건 완화 기법", 정보과학회 추계학술대회, pp. 208-211, 2002.
[9] 남영준 "이용자 중심의 시소러스 관리프로그램 설계", Journal of the Institute for Engineering and Technology Jeonju Univ., Vol.4, No.2, pp.225-242, 1998.
[10] 이종인, 한광록, "한국어 단어 시소러스 구축 시스템의 설계", 대한전자공학회, 제21권 제1호, pp.313-316, 1998.
[11] 박계숙, "객체지향 기법을 이용한 시소러스 관리 시스템의 개발에 관한 연구", 정보처리학회지, 제13권 제2호 pp.5-18, 1996.
[12] 노용완, 윤재선, 홍광석 "스무고개 게임을 위한 음성인식", 2002년도 전자공학회 하계 종합 학술대회논문집, 제25권 제1호, pp.203-206, 2002.
[13] 윤재선 "한국어 음성인식 Diction System의 구현", 2001년도 박사학위 청구 논문, pp.69-80, 2001.
[14] 윤재선, 정광우, 홍광석, "모음열과 VCCV 단위 HMM을 이용한 연속 숫자 음성인식," 한국음향학회 추계 학술대회논문집, pp.25-28, 2001.
[15] Quilici, A, Qiang Yang, Woods, S., "Knowldege-based Software Engineering Conference," proceedings of the 11th, pp.25-28, 1996.
[16] Li Junjie, Wang Kaizhu, "Natural language understanding based on background knowledge," TENCON 93. Proceeding. Computer, Communication, Control and Power Engineering. 1993 IEEE Region 10 Conference on Issue, pp. 460-462, Oct., 1993.
[17] Schwartz, R., Miller, S., Stallard, D., Markhoul, J., "Hidden understanding models for statistical sentence understanding," Acoustics, Speech and Signal Processing, ICASSP-97 IEEE International, Vol.2m pp.1479-1482, 1997.



노 용 완

e-mail : elec1004@hotmail.com
2001년 남서울대학교 정보통신공학과(학사)
2003년 성균관대학교 정보통신공학부(공학석사)
2003년~현재 성균관대학교 정보통신공학부 박사과정
관심분야 : 음성인식, 음성이해, 신호처리

홍 광 석

e-mail : kshong@skku.ac.kr
1985년 성균관대학교 전자공학과(학사)
1988년 성균관대학교 전자공학과(공학석사)
1992년 성균관대학교 전자공학과(공학박사)
1990년~1993년 서울보건전문대학 전산정보처리과 전임강사
1993년~1995년 제주대학교 정보공학과 전임강사
1995년~현재 성균관대학교 정보통신공학부 부교수
관심분야 : 음성인식 및 합성, HCI