

소 떨어지며 의미정보의 표현 능력이 다소 떨어진다는 것이다. 수식 표현 인식 시스템은 추가 정보, 환율정보, 일기예보 정보 등을 추출하는 정보추출 시스템이나 질의응답 시스템에 널리 사용될 수 있다. 주어진 문장에 수식 표현을 먼저 인식함으로써 자연언어처리 시스템의 처리 부담을 덜어줄 뿐 아니라 속도도 크게 개선할 수 있을 것이다.

본 논문에서는 수식 표현 인식을 위한 정규표현을 렉스(lex)라는 어휘 분석기를 사용하였다. 렉스는 입력 문자열에서 정규표현으로 기술된 토큰들을 찾아내는 프로그램을 작성하는 데 유용한 도구이다. 수식 표현 인식 시스템은 단계형 유한상태 오토마타를 사용하였다. 이것은 유한상태 오토마타가 여러 단계를 이루고 있는 경우를 말하며, 각 단계에서는 정규표현에 의해서 특별한 구를 인식한다. 이때 어떤 정규표현에도 일치되지 않은 입력은 무시되어 그대로 출력된다.

한국어에서 수사는 숫자와 고유어 수사와 한자어 수사가 있으며, 이들 수사는 기호와 단위성 의존명사와 함께 쓰여 문서에서 의미있는 정보를 나타낸다. 이 때 의미있는 정보의 단위는 크게 시간 표현과 수량 표현이 있다. 시간 표현은 다시 날짜, 시간, 기간으로 나눌 수 있으며, 수량 표현은 금전, 비율, 측도, 계수로 나눌 수 있다.

정규표현으로 각 단계의 규칙을 기술하였고, 이미 인식한 내용에 대해서는 아무런 처리를 하지 않고 넘기는 우회 방법과, 가능한 가장 큰 의미 정보를 나타내기 위한 최장 일치 방법을 사용하였으며, 또 어절 가운데서 인식하는 오류를 줄이기 위해 어절의 시작 단위에서 수식 표현을 인식하게 하였다. 그리고 수사와 단위성 의존명사와의 연결가능성도 고려하여 잘못된 인식을 줄이도록 하였다.

본 논문에서는 시스템의 성능을 평가하기 위하여 조선일보 94년 기사 중 일부를 수동으로 태깅하여 평가 집합으로 사용하였다. 객관적인 평가를 위해서 정보검색에서 사용하는 평가 기준인 정확률과 재현율을 사용하였으며, 평가 집합과 시스템 결과물을 비교하여 얻은 전체 성능은 정확률과 재현율이 각각 86.9%와 90.8%를 보였다. 오류 분석을 통해 단순한 규칙의 첨가나 조사와 같은 주변 품사 정보를 고려한다면 성능향상을 기대할 수 있을 것이다.

81. 이동 단말기를 위한 다자간 회의 시스템 구조 설계

컴퓨터공학과 이 호
지도교수 손 주 영

20세기 후반 50년은 컴퓨터 시대였으며 21세기에도 컴퓨터는 우리 생활에 지대한 영향을 미치고 있다. 특히 1980년대 초반에 본격적으로 보급이 확대되기 시작한 개인용 컴퓨터(PC: Personal Computer)는 기능이 날로 다양해지고 성능이 향상되고 있으며 새로운 디자인의 값싼 컴퓨터가 지속적으로 개발되고 있다.

20세기는 또한 통신 혁명의 시대였다고 말할 수 있다. 원시적인 모르스 부호 통신에서 시작된 통신 혁명은 전세계를 잇는 자동 전화 시스템으로 발전하고 이제는 휴대폰으로 확대 발전하여 전세계 어디로든지 원하기만 하면 통화할 수 있는 상태에까지 이르게 되었다. 20세기의 마지막 10년은 컴퓨터 시대이면서 동시에 인터넷 시대였다. 미국에서 ARPANET으로 시작된 인터넷은 전세계 컴퓨터를 네트워크로 연결함으로써 인류 전체에 심대한 영향을 미치고 있다. 그러나

아무도 인터넷이 인류에게 앞으로 어떠한 영향을 미칠 것인지 정확히 예측하지 못하고 있다.

휴대 정보 기기는 1980년대 말경에 PDA(Personal Digital Assistance)가 등장한 이래 아직까지는 보급이 크게 확대되지는 못하였으나, 최근 인터넷의 확산과 더불어 향후 5년 이내에 폭발적인 보급 확대가 이루어질 것으로 예상되고 있다. 이러한 이동 단말기들의 보급과 더불어 이동 통신기술의 발전으로 인터넷 정보검색, 파일공유, 메일서비스, 화상회의 서비스와 같은 기존의 인터넷 서비스를 그대로 제공받고자 하는 사용자들의 수요가 최근 급증하고 있는 추세다. 특히 이러한 추세와 더불어 이동 단말기의 사용자들에게 다자간 회의 서비스와 같은 복잡하고 제한된 환경에서의 서비스를 제공하기 위해 활발한 연구가 진행되고 있다.

기존의 회의 시스템에서는 PC와 같은 많은 데이터의 처리능력, 그리고 다양한 입력장치와 출력장치를 보유한 단말기를 통하여 음성, 화상, 데이터를 모두 포함하는 멀티미디어 서비스를 제공하고 있다. 이에 반해 이동 단말기는 소규모 데이터 처리능력, 제한된 입력장치와 출력장치로 인하여 기존 회의 서비스를 제공하기에는 많은 문제점을 가지고 있다. 특히 음성 서비스 위주의 이동 단말기들은 음성입력 이외에 DTMF(Dial Tone Message Frequency) 키 입력과 같이 단순한 입력만 가능하다. 또한 스피커를 통한 음성출력, 그리고 아주 제한적인 디스플레이를 통해서만 데이터의 출력이 가능하다. 뿐만 아니라 이동 단말기들은 기존 유선 망에 접속된 단말기들과는 근본적으로 네트워크를 통하여 데이터를 전송하는 능력에서 현저한 차이를 나타낸다.

이동 단말기들의 이러한 제한적인 기능들을 극복하고자 하는 노력들이 WML, VoiceXML, 음성인식이나 합성과 같은 기술들을 통하여 나타나고 있다. 그러나 이러한 기술들은 이동 단말기를 이용하여 아주 단순하고 기본적인 인터넷 서비스(문자전송, 메일확인, 게시판 등)를 제공하는 정도에 불과하다. 따라서 음성 통화중 파일전송 및 공유, 인터넷 자료검색과 같은 데이터 처리 기능을 가진 다자간 회의 서비스는 이동 단말기들이 제공하기에 어려움이 있다.

본 논문에서는 이러한 다양한 서비스 중에서 제약적인 기능을 가진 환경에서 이동 단말기를 통한 다자간 회의 서비스 제공이 가능한 구조 모델을 제시하고자 한다. 위에서 언급하였듯이 음성서비스 위주의 이동 단말기를 이용하여 회의 서비스 도중 데이터 처리를 위한 서비스를 제공하는 것은 매우 까다로운 부분이다. 이에 본 논문에서는 이러한 문제점들을 기존 회의 시스템에 WML과 VoiceXML이라는 새로운 기반기술의 접목과 발생하는 문제점의 극복을 통하여 보다 확장된 회의 서비스를 제공하는 새로운 형태의 회의 시스템 구조 모델을 제시하고 구조모델에 대한 구체적인 동작과 실험을 통하여 그 가능성을 타진한다.

82. 차분 파워 스펙트럼을 이용한 음성구간 추출 및 잡음제거

컴퓨터공학과 정 성 일
지도교수 신 욱 근

음성인식 기술은 지난 수십 년간의 연구결과 최근에는 괄목할만한 진보가 이루어졌으며, 음성인식 기술과 뿌리를 같이하는 화자인식 기술도 많은 진전이 있었다. 이러한 음성인식과 화자인식의 기술적인 진보는 그 기술이 실생활에 이용될 수 있으리라는 기대를 가지도록 하기에 충분했다.