

패턴 매칭과 통계 모델링을 이용한 대화 인터페이스 시스템의 구현

김학수[†]

요 약

본 논문에서는 대화 인터페이스 시스템을 구성하는 필수 요소들을 알아보고, 각각에 대한 실용적인 구현 방법을 제안한다. 구현 시스템은 담화 관리자, 의도 분석기, 개체명 인식기, SQL 질의어 생성기, 응답 생성기로 구성된다. 구현 과정에서 비교적 영역 의존성이 낮은 의도 분석 모듈은 통계 기반의 최대 엔트로피 모델을 이용한다. 그리고 높은 영역 이식성을 필요로 하는 개체명 인식기, SQL 질의어 생성기, 응답 생성기는 단순 패턴 매칭 방법을 이용한다. 일정 관리 영역에서의 실험에서 구현 시스템은 88.1%의 의도 분석 정확률과 83.4%의 SQL 질의어 생성 성공률을 보였다.

키워드 : 대화 인터페이스 시스템, 의도 분석기, SQL 질의어 생성기

Implementation of a Dialogue Interface System Using Pattern Matching and Statistical Modeling

Hark-Soo Kim[†]

ABSTRACT

In this paper, we review essential constituents of a dialogue interface system and propose practical methods to implement the each constituent. The implemented system consists of a discourse manager, an intention analyzer, a named entity recognizer, a SQL query generator, and a response generator. In the progress of implementation, the intention analyzer uses a maximum entropy model based on statistics because the domain dependency of the intention analyzer is comparatively low. The others use a simple pattern matching method because they needs high domain portability. In the experiments in a schedule arrangement domain, the implemented system showed the precision of 88.1% in intention analysis and the success rate of 83.4% in SQL query generation.

Keywords : Dialogue interface system, Intention analyzer, SQL query generator

1. 서 론

대화 인터페이스 시스템은 자연어 대화를 통하여 인간과 컴퓨터가 자연스럽게 상호작용할 수

있도록 도와주는 지능형 프로그램을 말한다. 이러한 대화 인터페이스 시스템을 구현하는 방법은 대화 제어 방식에 따라 유한상태 모델 (finite-state model), 프레임기반 모델 (frame-based model)[6], 계획기반 모델 (plan-based model)[12]로 나뉘어 진다. 유한상태 모델은 시스템의 응답 발화를 나타내는 노드 (node)들과 사용자의 입력 발화를 나타내는 아크

[†] 정희원: 강원대학교 컴퓨터정보통신공학전공 교수(교신저자)
논문접수: 2007년 3월 8일, 심사완료: 2007년 5월 7일
* 이 연구(논문)은 산업자원부 지원으로 수행하는 21세기 프론티어 연구개발사업(인간기능 생활지원 지능로봇 기술개발사업)의 일환으로 수행되었습니다. 또한 부분적으로 강원대학교 정보통신연구소의 지원을 받았습니다.

(arc)들을 이용하여 대화의 흐름을 표현하는 모델이다. 그러므로 대화의 흐름이 미리 정해진 순서에 따라서만 이루어진다는 단점이 있다. 프레임 기반 모델은 특정 개수의 슬롯으로 구성된 프레임을 정의하고, 빈 슬롯들을 자유롭게 채워나가는 방법으로 대화의 흐름을 표현하는 모델이다. 프레임 기반 모델은 슬롯을 채우는 순서가 미리 정해져 있지 않기 때문에 대화의 흐름이 유한상태 모델보다 자유롭다는 장점이 있다. 계획기반 모델은 추론 기반의 레시피(recipe; 대화를 구성하는 소단위의 시나리오 대본)의 확장을 통하여 대화의 흐름을 표현하는 모델이다[4][9][12]. 계획기반 모델은 추론 규칙을 이용하기 때문에 복잡한 대화 현상을 다룰 수 있다는 장점이 있다. 그러나 추론 엔진의 구현이 어렵고, 레시피 구축 및 유지보수에 많은 비용이 든다는 단점이 있다. 본 논문에서는 영역 이식에 필요한 비용을 줄이면서 일정 수준의 대화 유연성을 보장하기 위하여 프레임 기반 모델을 기초로 한 실용적인 대화 인터페이스 시스템을 구현한다.

본 논문의 구성은 다음과 같다. 2장에서 대화 인터페이스 시스템을 구성하는 필수 요소들을 알아보고 각각의 구현 방법들을 설명한다. 3장에서는 실험 데이터를 설명하고 실험 결과를 분석한다. 마지막으로 4장에서 결론을 내린다.

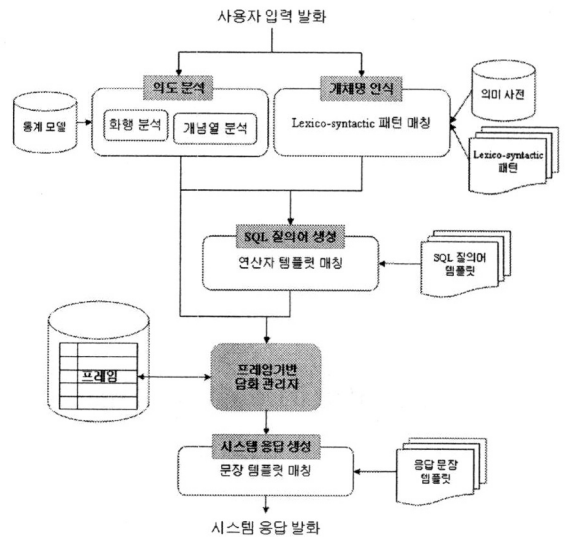
2. 대화 인터페이스 시스템

2.1 전체 시스템 구조

<그림 1>은 본 논문에서 구현한 대화 인터페이스 시스템의 구조도이다. <그림 1>에서 보는 것과 같이 구현 시스템은 다음과 같이 5가지 필수 요소로 구성된다.

- **답화 관리자:** 대화의 흐름을 제어하는 모듈
- **의도 분석기:** 사용자의 발화에 포함된 의도를 파악하는 모듈
- **개체명 인식기:** 프레임의 검색 및 갱신에 필요한 정보를 사용자 발화로부터 추출하는 모듈

- **SQL 질의어 생성기:** 의도 분석과 개체명 인식 결과를 바탕으로 데이터베이스 형식의 프레임에 접근하기 위한 SQL 질의어를 생성하는 모듈
- **응답 생성기:** SQL 질의 결과를 이용하여 시스템의 응답 발화를 생성하는 모듈



<그림 1> 대화 인터페이스 시스템 구조도

2.2 답화 관리

프레임 기반 모델에서 답화 관리자의 역할은 프레임에 정보를 채우기 위해서 사용자에게 질문을 던지고 사용자의 답변으로부터 정보를 얻어 슬롯을 채우면서 대화를 진행하는 것이다. 그리고 사용자의 정보 요구가 있을 경우에 현재 프레임 정보를 이용하여 적절한 검색 결과를 제공하는 것이다. 예를 들어, 일정관리 영역에서 4개의 슬롯을 가진 프레임 [person, date, time, place]이 정의되어 있다고 가정하면, 답화 관리자는 각각의 슬롯을 채우기 위하여 미리 정의된 질문 템플릿에 따라 <그림 2>와 같이 대화를 이끌어 나간다. 대화를 이끌어 나가기 위해 필요한 정보들은 의도 분석, 개체명 인식, SQL 생성 과정을 거치면서 답화 관리자에게 제공된다.

사용자: 내일 약속 있어.
 시스템: 누구와 약속이 있으십니까?
 사용자: 홍길동이랑.
 시스템: 약속 시간은 언제입니까?
 사용자: 12시
 시스템: 약속 장소는 어디입니까?
 사용자: 신촌
 시스템: 약속 내용을 저장하였습니다.
 <중략>
 사용자: 내일 홍길동이랑 어디에서 만나기로 했?
 시스템: 홍길동님과 신촌에서 약속이 있습니다.

<그림 2> 프레임기반 모델의 대화 시나리오 예제

2.3 의도 분석

발화에 포함된 화자의 의도는 <표 1>과 같이 영역 독립적인 문장의 유형을 나타내는 화행 (speech act)과 영역 종속적인 세부 의미를 나타내는 개념열(concept sequence)의 쌍으로 표현될 수 있다[10].

<표 1> 화행과 개념열을 이용한 의도 표현 (U: 사용자, S: 시스템)

발화	화행, 개념열
U: 내일 아침에 깨워줘.	request, alarmtable-select
S: 몇 시에 깨워드릴까요?	ask-ref, alarmtable-insert-time
U: 6시30분	response, alarmtable-insert-time
S: 예, 알겠습니다.	accept, alarmtable-insert-time

본 논문에서는 [3]을 기초로 영역 독립적인 화행을 정의하고, 영역 의존적인 의미들은 개념열을 통하여 보충하는 방법으로 화자의 의도를 일반화한다. <표 2>는 본 논문에서 정의하여 사용하는 화행을 보여준다.

<표 2> 화행 구성

화행	설명	화행	설명
Greeting	대화 서두의 인사말	Request	행위를 요청
Expressive	대화 후미의 인사말	Ask-confirm	이전 발화의 확인
Opening	실제 대화의 시작	Confirm	확인 발화의 응답
Ask-ref	WH-질문	Inform	정보 제공
Ask-if	YN-질문	Accept	호응
Response	응답		

<표 3>은 일정관리 영역에서 데이터베이스 연산을 모델링하여 정의한 개념열의 구성을 보여준다. <표 3>과 같이 본 논문에서는 2가지 테이블, 4가지 연산자, 8가지 필드를 3층 구조의 개념열 부착 방법[11]에 따라 정의하여 사용한다.

<표 3> 일정관리 영역에서의 개념열 구성

테이블명	연산자명	필드명
timetable alarmtable	insert delete select update	agent date day-of-week time person place content field

하나의 대화를 구성하는 i 번째 발화 U_i 의 화행 SA_i 와 개념열 CS_i 가 서로 독립이라고 가정하면 사용자 의도 I_i 를 계산하는 통계 모델 $P(I_i|U_i)$ 는 수식 (1)과 같이 근사화된다.

$$P(I_i|U_i) \approx P(SA_i|U_i)P(CS_i|U_i) \tag{1}$$

수식 (1)에 베이저안 정리(Bayesian rule)를 적용하고, 전체 확률값에 영향을 미치지 않는 $P(U_i)$ 를 제거하면 수식 (2)와 같다.

$$P(I_i|U_i) \approx \frac{P(U_i|SA_i)P(SA_i)P(U_i|CS_i)P(CS_i)}{P(U_i)P(U_i)} \approx P(U_i|SA_i)P(SA_i)P(U_i|CS_i)P(CS_i) \tag{2}$$

수식 (2)에 화행과 개념열이 바로 이전의 문맥에만 의존한다는 1차 마코프(Markov) 가정을 적용하면 수식 (3)과 같다.

$$P(I_i|U_i) \approx P(U_i|SA_i)P(SA_i|SA_{i-1})P(U_i|CS_i)P(CS_i|CS_{i-1}) \quad (3)$$

수식 (3)에서 자연어 발화의 다양성으로 인하여 $P(U_i|SA_i)$ 와 $P(U_i|CS_i)$ 를 직접 계산하는 것은 매우 힘들다. 그러므로 발화 U_i 가 문장 자질들의 집합으로 표현될 수 있다고 가정하면 수식 (4)와 같다.

$$P(I_i|U_i) \approx P(F_i|SA_i)P(SA_i|SA_{i-1})P(F_i|CS_i)P(CS_i|CS_{i-1}) \quad (4)$$

수식 (4)에서 F_i 는 i 번째 발화의 문장 자질 집합으로써 어휘 자질과 품사 자질로 구성된다[1]. 어휘 자질은 품사가 부착된 어휘를 말하며, 품사 자질은 품사 바이그램(bigram)을 말한다. 예를 들어, ‘안녕하세요?’라는 발화의 형태소 분석 결과가 ‘안녕/ncp+하/xsp+세요/ef+?/sf’라고 했을 때, 어휘 자질은 ‘안녕/ncp, 하/xsp, 세요/ef, ?/sf’이고, 품사 자질은 ‘ncp-xsp, xsp-ef, ef-sf’이다.

수식 (4)에서 모든 문장 자질을 이용하여 $P(F_i|SA_i)$ 와 $P(F_i|CS_i)$ 를 계산하는 것은 효율적이지 못하다. 그러므로 본 논문에서는 추출된 자질 후보들과 범주(화행 또는 개념열) 사이의 χ^2 통계량[14]을 계산하여 정보량이 큰 것만을 선택적으로 사용한다.

수식 (4)의 조건부 확률값들은 수식 (5)에 의해서 계산될 수 있으며, $P(a,b)$ 를 계산하기 위하여 본 논문에서는 수식 (6)과 같은 최대 엔트로피 모델(maximum entropy model)을 이용한다.

$$P(a,b) = \frac{P(a,b)}{\sum_a P(a,b)} \quad (5)$$

$$P(a,b) = \pi \prod_{i=1}^k \alpha_i^{f_i(a,b)}, \text{ where } 0 < \alpha_i < \infty, i = \{1, 2, \dots, k\} \quad (6)$$

수식 (6)에서 π 는 정규화 요소이고, α_i 는 자질 함수 $f_i(a,b)$ 에 대한 모델 파라미터(parameter)이다. 자질 함수 $f_i(a,b)$ 는 화행 또는 개념열을 의미하는 범주 a 에서 해당 자질 b 가 나타나면 1, 그렇지 않으면 0의 값을 가진다.

2.4 개체명 인식

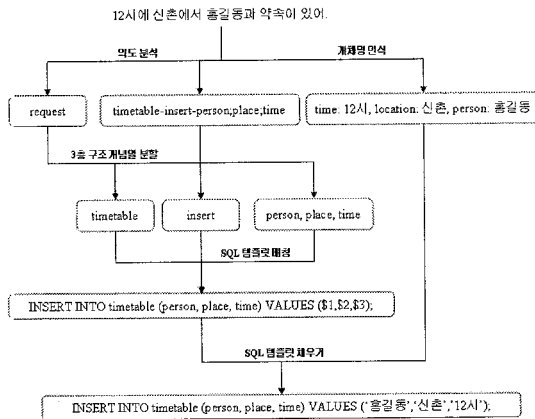
대화 인터페이스 시스템이 사용자의 의도에 따라 프레임 내부의 정보를 검색하거나 갱신하기 위해서는 입력된 발화로부터 근거가 되는 정보를 응용영역에 맞게 추출해야 한다. 예를 들어, “12시에 신촌에서 홍길동과 약속이 있어”라는 발화에 포함된 정보를 프레임 ‘[person, date, time, place]’에 저장하기 위해서는 입력 발화로부터 ‘person: 홍길동’, ‘time: 12시’, ‘place: 신촌’이라는 영역 의존적인 개체명(named entity)을 추출해야 한다. 본 논문에서는 영역 이식성이 높은 개체명 인식기를 구현하기 위하여 lexico-syntactic 패턴 매칭 방법[8]을 이용한다. 개체명 인식을 위한 과정을 자세히 살펴보면 다음과 같다. 먼저, 의미 사전을 이용하여 입력된 발화를 적당한 형태로 변환한다. 의미 사전은 단어들의 의미 표지(semantic marker)를 담고 있으며, 입력 발화는 패턴 매칭에 앞서 의미 표지로 변환된다. 예를 들어, “12시 40분에 약속이 있어”라는 발화는 “num@hour num@minute j 약속 j 있 ef”로 변환된다. 예제에서 ‘@hour’, ‘@minute’가 의미 표지이다. 의미 사전에 없는 단어들은 어휘 형태가 그대로 유지되며, 기능어들은 품사로 대체된다. 입력 발화를 변환한 후, 개체명 인식기는 다음과 같은 형식으로 수동 구축된 lexico-syntactic 패턴과 변환된 질의를 매칭하여 개체명을 추출한다.

- 형식: 패턴 추출위치 -> 개체명
- 예제: (num(@hour|@minute))+.*약속 1 -> date
- 설명: 첫 번째 괄호에 매칭되는 부분을 입력 발화에서 추출하여 date라는 개체명을 부여한다.

만약 여러 개의 lexico-syntactic 패턴이 매칭되면 가장 먼저 매칭된 것을 우선적으로 적용한다.

2.5 SQL 질의어 및 시스템 응답 생성

입력된 발화에 대한 의도 분석과 개체명 인식이 정확하게 이루어졌다고 하더라도 자연어를 SQL 질의어로 변환하는 과정에는 많은 애매성이 존재한다. 본 논문에서는 개체명 인식기와 마찬가지로 영역 이식성을 높이기 위하여 <그림 2>와 같이 템플릿(template) 매칭 방법을 이용한다.



<그림 3> 템플릿 기반 SQL 생성 예제

템플릿 매칭 과정을 자세히 살펴보면 다음과 같다. 먼저, 의도 분석과 개체명 인식이 완료되면 대화 인터페이스 시스템은 사용자 발화가 프레임에 대한 접근을 요구하는 것인지 화행을 기준으로 결정한다. 즉, 사용자 발화의 화행이 ask-ref, request, inform에 해당하면 프레임에 대한 검색이나 갱신을 요구하는 것으로 판단한다. 그리고 프레임에 대한 접근이 필요한 발화이면 3층 구조로 구성되어 있는 개념열을 테이블명, 연산자명, 필드명으로 분리한다. 다음으로 연산자명에 해당하는 템플릿을 선택한 후, 테이블명과 필드명을 템플릿에 추가한다. 마지막으로 개체명 인식 결과를 참조하여 해당 필드명에 대한 속성값을 채우고 SQL 질의어를 완성한다.

SQL 질의어가 완성되면 대화 인터페이스 시스템은 해당 SQL 질의어를 수행하고 그 결과를 미리 정의된 응답 템플릿에 따라 사용자에게 제공한다. 이 과정에서 응답의 다양성을 위하여 동일

한 의미의 응답 템플릿을 여러 개 만들어 두고 무작위로 선택하는 방법을 이용한다.

3. 실험 결과

3.1 실험 데이터 및 환경

구현 시스템의 성능을 실험하기 위하여 일정관리 영역에서 Wizard-Of-Oz 방식으로 시뮬레이션(simulation)한 대화 말뭉치를 수집한 후, 수동으로 화행과 개념열을 부착하였다. 수집된 말뭉치는 일정 추가, 삭제, 변경과 관련된 내용을 포함하고 있다. 말뭉치의 구성은 956개의 대화(21,336개의 발화)로 구성되며, 대화당 평균 발화의 수는 22.32개이다. 의도 분석에 필요한 통계 모델은 MEMT(Maximum Entropy Modeling Toolkit)[13]을 이용하여 구현하였으며, 화행 분석과 개념열 분석에 사용된 문장 자질의 수는 각각 1,000개와 1,200개로 제한하였다. 평가 방법은 2가지로 나누어서 진행하였다. 첫 번째는 의도 분석에 대한 정확률을 살펴보았으며, 두 번째는 SQL 질의어 생성에 대한 성공률을 살펴보았다. 그리고 모든 실험에는 5배 교차검증(5-fold cross validation)을 수행하였다.

3.2 실험 결과

<표 4>는 동일한 학습 데이터와 평가 데이터를 이용하여 화행과 개념열을 분석한 결과이다. 각 발화에 대해 화행과 개념열을 모두 맞춘 정확률은 88.1%였다.

<표 4> 의도 분석 정확률

모델	화행 분석 평균 정확률(%)	개념열 분석 평균 정확률(%)
Kim2003[7]	85.5	-
Choi2005[5]	88.6	-
En2005[2]	90.4	-
구현 시스템	93.0	90.2

구현 시스템과 Kim2003, Choi2005는 서로 다른 문장 자질을 사용하기 때문에 직접적으로 비교하는 것은 어렵다. 그러나 <표 4>에서 보는

것과 같이 구현 시스템은 사용한 입력 자질의 동일성 여부에 상관없이 이전의 모델들보다 좋은 결과를 보였다. 또한 구현 시스템은 화행 분석뿐만 아니라 개념열 분석까지 동일한 확률 모델을 이용하여 동시에 수행할 수 있다는 장점이 있다. 의도 분석에 실패한 유형을 살펴본 결과, 많은 오류들이 문맥 정보를 잘못 사용하여 생긴 것이었다. 구현된 의도 분석기는 문맥 정보로써 바로 이전의 화행이나 개념열을 이용하였다. 그러나 부대화가 삽입된 경우에 대화는 계층적 담화구조를 형성하며, 바로 이전의 화행이나 개념열보다는 계층적으로 인접한 것들이 올바른 문맥이 된다. 이런 문제를 해결하기 위해서는 담화 구조 분석 모델을 구현하고 구현된 의도 분석기와 통합하여 올바른 문맥을 자질로 사용해야 한다.

SQL 질의어 생성에 대한 성공률을 평가하기 위하여 먼저 프레임 정보의 검색이나 갱신과 관련된 발화들을 화행을 기준으로 선별하였다. 그리고 선별된 발화에 대하여 SQL 질의어 생성기가 만들어낸 모든 SQL 질의어들 중에 올바른 질의어의 비율을 계산하였다. 이러한 방법으로 계산된 구현 시스템의 SQL 질의어 생성 성공률은 83.4%였다. SQL 질의어 생성에 실패한 유형을 살펴본 결과, 많은 오류들이 생략이나 대응 현상으로 인하여 올바른 개체명을 인식하지 못한 경우였다. 이러한 문제를 해결하기 위해서는 대응어 처리 모듈을 개발하여 개체명 인식기의 전처리 엔진으로 사용하는 것이 필요하다.

4. 결 론

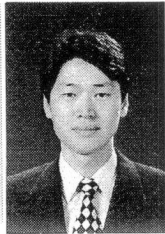
본 논문에서는 실용적인 대화 인터페이스 시스템의 구현 방법에 대해서 살펴보았다. 구현 시스템은 영역 이식성을 높이기 위하여 사용자의 의도를 화행과 개념열로 분리하였으며, 개체명 인식, SQL 질의어 생성, 응답 생성에 패턴 매칭 방법을 적용하였다. 그리고 의도 분석의 성능을 높이기 위하여 통계 기반의 최대 엔트로피 모델을 적용하였다. 일정 관리 영역에서의 실험 결과, 구현 시스템은 88.1%의 의도 분석 정확률과 83.4%의 SQL 질의어 생성 성공률을 보였다.

참 고 문 헌

- [1] 김경선 · 서정연 (2003). 자질 선택 기법을 이용한 한국어 화행 결정. 한국정보과학회 논문지 30(3). pp. 278-284.
- [2] 은종민 · 이성욱 · 서정연 (2005). 지지벡터기계(support vector machines)를 이용한 한국어 화행분석. 한국정보처리학회 논문지 12B(3). pp. 365-368.
- [3] 이현정 (1997). 한국어 대화체 문장의 화행 분석. 석사학위논문. 서강대학교.
- [4] Caberry, S. (1989). A pragmatics-based approach to ellipsis resolution. Computational Linguistics 15(2). pp. 75-96.
- [5] Choi, W. · Kim, H. · Seo, J. (2005). An Integrated dialogue analysis model for determining speech acts and discourse structures. IEICE Information and Systems E88D(1). pp. 150-157.
- [6] Goddeau, D. · Meng, H. · Polifroni, J. · Seneff, S. · Busayapongchai, S. (1996). A form-based dialogue manager for spoken language applications, Proceedings of ICSLP. pp. 701-704.
- [7] Kim, H. · Seo, J. (2003). An efficient trigram model for speech act analysis in small training corpus. Journal of Cognitive Science 4(1). pp. 107-120.
- [8] Kupiec, J. (1993). Murax: A robust linguistic approach for question answering using an on-line encyclopedia, Proceedings of SIGIR.
- [9] Lambert, L. · Caberry, S. (1991). A tripartite plan-based model of dialogue, Proceedings of ACL. pp. 47-54.
- [10] Langley, C. (2002). Analysis for speech translation using grammar-based parsing and automatic classification, Proceedings of ACL Student Research Workshop.
- [11] Lee, H. · Kim, H. · Seo, J. (2006). Efficient domain action classification using neural networks. Lecture Notes in Computer Science 4233. pp. 150-158.

- [12] Litman, D. J. · Allen, J. F. (1987). A plan recognition model for subdialogues in conversations. Cognitive Science 11. pp. 163-200.
- [13] Ristad, E. (1996). Maximum entropy modeling toolkit, Technical Report. Department of Computer Science, Princeton University.
- [14] Yang, Y. · Pedersen, J. O. (1997). A comparative study on feature selection in text categorization, Proceedings of ICML.

김 학 수



1996 건국대학교 전자계산학과
(공학사)
1998 서강대학교 컴퓨터학과
(공학석사)
2003 서강대학교 컴퓨터학과
(공학박사)

2004~2005 CIIR in UMass, Amherst
(박사후연구원)

2005~2006 한국전자통신연구원 (선임연구원)

2006~현재 강원대학교 컴퓨터정보통신공학전공
교수

관심분야: 자연어처리, 대화시스템, 정보검색, 질
의응답시스템

E-Mail: nlpdrkim@kangwon.ac.kr