

# 유사도 계산과 대화 전이 네트워크를 이용한 말뭉치 기반 대화 시스템

강상우\* 박홍민\* 고영중\*\* 서정연\*

\*서강대학교 컴퓨터학과, 121-742, 서울특별시, 마포구 신수동 1번지

\*\*동아대학교 컴퓨터공학과, 604-714, 부산광역시, 사하구 하단2동 840

swkang@sogang.ac.kr bluely@sogang.ac.kr yjko@dau.ac.kr seojy@sogang.ac.kr

## The Corpus-based Dialogue System Using a Dialogue Transition Network and a Similarity Measure Method

Sangwoo Kang\* Hongmin Park\* Youngjoong Ko\*\* Jungyun Seo\*

\*Dep. of Computer Science, Sogang University, Korea

\*\*Dep. of Computer Engineering, Dong-A University, Korea

### 요 약

본 연구는 말뭉치로부터 추출된 정보를 사용하여 대화 시스템에 필요한 과정들을 통합 처리하는 시스템을 제안한다. 기존 연구는 영역 확장 시 대화 시스템의 각 과정들을 위해 많은 노력이 필요하였지만, 제안하는 방법은 말뭉치를 사용하여 각 과정들을 통합적으로 업데이트함으로써 이 문제를 해결하고자 한다. 사용자 입력문장과 말뭉치의 각 문장들 간의 유사도 계산을 통하여 의미적으로 가장 유사한 말뭉치 문장의 정보를 이용하고, 시스템 응답에 필요한 정보를 선택한다. 또한, 문맥에 관련된 정보를 자동으로 추출하여 대화 관리를 위한 대화 전이 네트워크(network)를 생성한다. 따라서, 제안 시스템은 말뭉치의 추가 및 수정만으로 새로운 영역 확장과 관리에 용이한 구조를 갖는다. 실험으로 관찰한 제안된 시스템의 성능은 유사도 계산 만족도 약 77%, 시스템 응답의 적절성 약 84%로 충분히 작업 수행이 가능한 점수를 보여주었다.

### 1. 서 론

본 논문은 말뭉치를 이용하여 대화 전이 네트워크(DTN: Dialog Transition Network)를 생성하고 사용자 입력문장 분석과 시스템 응답 생성과정을 통합하여 처리하는 대화 시스템에 관한 연구이다.

대화 시스템은 인간과 기계가 상호작용의 도구로서 자연어(natural language)를 사용하여 특정 작업을 성취하도록 하는 시스템을 말한다. 대화 시스템은 자연어가 가지고 있는 유연성, 명료성, 표현력으로 인해 기존의 인터페이스의 한계를 극복할 수 있을 것으로 기대하고 있다. 최근 대화 시스템은 ASRS(Automatic Speech Recognition System)의 성능향상과 더불어 기능성 로봇, 텔레메틱스(telematics), 전화교환시스템 등과 유비쿼터스(ubiquitous) 환경에 대응할 수 있는 차세대 인터페이스로 기존의 시스템을 대체하기 위한 연구가 활발히 진행되고 있다.

대화 시스템은 인간의 대화현상을 이해하기 위한 대화

모델의 연구로부터 시작하였고, Kautz와 Allen[1]은 그것을 계획 이론(plan theory)으로 설명하였다. 계획인식은 협상대화(negotiation dialog)와 같은 복잡한 대화구조를 설명할 수 있는 모델로 발전하였다[2]. 하지만 점차 인간의 명령이나 질문에 반응할 수 있는 시스템이 요구되면서 인간-기계 상호작용을 위한 시스템들이 시도되고 있으며 실제 환경에서 응용할 수 있도록 제한된 영역에서 강건하게 동작할 수 있는 모델에 관련된 연구가 진행되고 있다. 이 연구들은 특정 목적을 순차적으로 처리할 수 있도록 유한 상태 기계(finite state automata)형태를 사용하여 정형화된 대화 영역에서 강건한 동작을 보장한다. 그러나 아직 대부분의 연구들은 극히 제한된 응용영역을 가지고 있으며 영역의 전환과 대화 범위에 확장에는 전문가의 많은 시간과 노력이 필요하다. 따라서 많은 연구자들은 새로운 대화모델의 연구뿐만 아니라 영역 확장과 유지 관리에 용이한 구조를 갖는 대화 시스템 연구에 초점을 맞추고 있다.

본 연구는 말뭉치로부터 추출된 정보를 사용하여 대화

시스템에 필요한 과정들을 통합 처리하는 시스템을 제안한다. 사용자 입력문장과 말뭉치의 각 문장들 간의 유사도 계산을 통하여 의미적으로 가장 유사한 말뭉치 문장의 정보를 이용하고 대화 관리 필요한 정보를 자동으로 추출하여 사용하였다. 따라서, 제안된 시스템은 말뭉치의 추가 및 수정만으로 새로운 영역 확장과 관리에 용이한 구조를 갖는다. 기존의 연구는 영역 확장 시 대화 시스템의 각 과정들을 위해 많은 각각의 노력이 필요하였지만 제안하는 방법은 말뭉치를 사용하여 각 과정들을 통합적으로 업데이트함으로써 이 문제를 해결하였다. 또한 제안 시스템은 로봇 플랫폼과 멀티모달 환경을 위해 디자인 되었다. 따라서 음성 대화뿐만 아니라 로봇 플랫폼의 다양한 리소스를 활용하여 기존의 시스템보다 다양한 상호작용이 가능하다. 예로 제스처어 인식정보와 실제 환경에 대한 정보를 포함하는 온톨로지를 사용하여 멀티모달 대응어를 처리할 수 있으며, 사용자 모델 정보를 사용하여 개인화 된 서비스 등의 제공이 가능하다.

본 논문에서는 기존의 시스템의 범위인 대화 전이 네트워크의 생성과 관리 그리고 유사도 계산을 통한 사용자 입력문장의 분석에 초점을 맞추어 설명한다.

## 2. 관련 연구

대화 시스템에 관한 기존 연구는 크게 두 가지 분야로 나눌 수 있다. 그 중 첫 번째는 실제 환경에 적용할 수 있도록 강인한 성능을 보장할 수 있는 시스템을 위한 연구이다. 이 연구의 범위는 비교적 간단하고 제한적인 영역에 한정되며 대화 형태의 변이가 없거나 예측할 수 있는 형태의 대화에 적용된다. 대표적으로 스크립트기반 모델(script-based model)에 관한 연구가 있다. 다른 연구 분야로는 인간의 복잡하고 다양한 대화 현상을 설명하기 위한 대화 모델에 초점을 집중한 분야이다. 계획기반 모델(plan-based model)은 위 분야에 대표적인 연구이며 현재까지 지속적인 연구가 진행되고 있다. 최근에는 확률기반 모델을 도입하여 각 모듈간의 에러 전파와 유지 관리 측면에서의 문제점들을 해결하려는 방법들도 시도되고 있다.

스크립트기반 모델은 특정 목적을 수행하기 위한 작업 절차를 순차적으로 처리할 수 있도록 대화 전이 네트워크 모델을 사용하고 일반적으로 양식기반 대화 모델(form-based dialogue model)에 사용된다[2]. 대화 전이 네트워크 모델은 대화의 상태와 전이 과정을 네트워크로 구현하여 예측할 수 있는 사용자 발화의 담화 구조를 표현한다. 이 모델은 비교적 설계가 쉽고 강건하게 동작한

다. 따라서 이 모델은 대화가 간결하고, 항상 일정한 형태를 유지하며 모호성을 배제할 수 있는 강건함이 요구되는 영역에 적합하다. 하지만 순차적 구조를 가지기 때문에 예외사항에 대처할 수 있는 유연성이 떨어지며 영역 이식성(portability) 또한 매우 낮다.

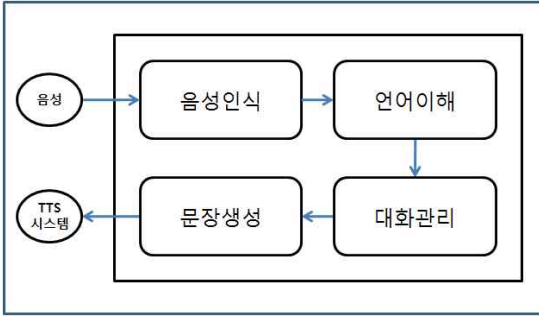
계획기반 모델은 관측된 단위 행위(action)로부터 사용자가 원하는 목적을 추론(plan recognition)하는 방법을 사용한다. 사용자 믿음(belief)의 상태에 따라 다양한 추론이 가능하고 레서피<sup>1)</sup>(recipe)의 활용으로 복잡한 담화 구조를 설명할 수 있다. Kautz와 Allen[1]은 계획 이론(plan theory)을 처음으로 대화에 적용하였고 계획 인식에 필요한 대화 모델을 담화 수준(discourse level)과 영역 수준(domain level)으로 구분하였다. Lambert[3][4]는 보다 복잡한 담화 구조를 처리할 수 있도록 담화 수준과 영역 수준 사이에 문제-해결 수준(problem-solving level)을 추가하였다. Kautz와 Allen, 그리고 Lambert의 연구는 인간-기계의 대화를 모델링 것이 아니라 인간-인간의 대화를 설명할 수 있는 모델링의 관점에 초점을 맞추었다.

기존의 대화 시스템들이 음성 인식 시스템과의 결합이 시도 되면서 각 시스템들의 오류 전파 문제들이 제기되고, 시스템 구축과 유지관리의 효율성을 극복하기 위하여 확률기반 모델을 통한 접근 방법들이 시도되고 있다. Roy[5]는 MDP(Markov Decision Process)을 이용하여 대화에서 일어날 수 있는 각각의 모든 상태를 모델링 하여 현재 상태가 다른 상태로 전이될 확률을 예측함으로써 대화를 진행하는 방법을 제안하였고, Singh[6]은 강화학습(Reinforcement Learning)을 사용하여 사용자의 보상(reward)을 통해 시스템의 대화 전략을 학습해 나가는 방법을 제안하였다. 이 방법들은 극히 제한된 상태들을 정의할 수 있고, 수차례의 반복 학습하는 과정이 필요하기 때문에 아직까지 실제 환경에서의 효율성은 매우 낮다고 할 수 있다.

## 3. 제안 시스템 구성

음성 대화 시스템은 일반적으로 음성 인식, 언어 이해, 대화 관리, 그리고 문장 생성 과정으로 구성 된다[그림 1]. 음성 인식은 발화음성을 텍스트로 변화시키며 언어 이해는 발화 텍스트를 화행, 개체명 등의 의미표지를 부착하여 시스템이 처리 가능한 형태로 분석한다. 대화 관리의 대화의 문맥을 관리하며 사용자 입력에 대한 응답

1) 계획 인식을 위한 담화, 영역 지식 표현 형식



[그림 184] 대화 시스템의 일반적인 처리절차

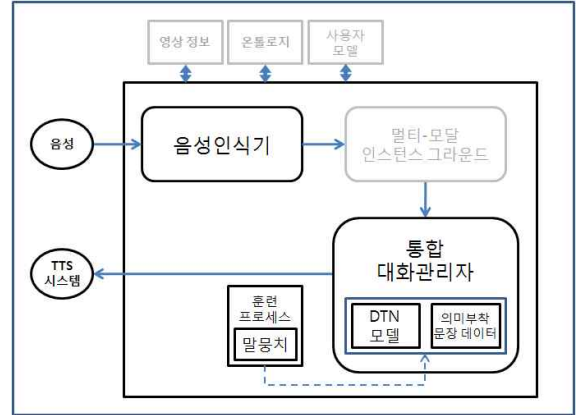
을 결정한다. 문장 생성은 결정된 시스템 응답을 시스템이 발화할 문장으로 생성한다. 하지만 대부분 시스템의 각 모듈들은 서로 독립적으로 동작하며 각 모듈에서 발생하는 오류나 환경에서 발생하는 오류들은 다음 모듈로 전파되어 전체 시스템의 성능을 떨어뜨리는 요인이 된다. 또한, 대부분의 시스템들은 대화의 범위를 확장하기 위해서 각 모듈들은 데이터의 추가와 모듈의 수정이 필요하다. 따라서, 음성 대화 시스템의 활용을 위해서는 영역의 전환과 확장에 필요한 새로운 시도가 필요하다.

제안하는 시스템은 언어 이해, 대화 관리 그리고 문장 생성모듈이 통합된 형태를 갖는다[그림 2]. 제안하는 통합 대화 관리자는 언어 이해를 위해 사용자 입력문장과 말뭉치의 사용자 발화 문장들의 유사도를 계산하여 가장 높은 점수를 획득한 문장의 의미표지들을 사용자 입력문장의 의미표지로 사용한다. 통합 대화 관리자는 대화의 유지와 관리를 위해 대화 전이 네트워크를 포함한다. 대화 전이 네트워크는 별도의 훈련 절차를 통해 말뭉치로부터 자동 생성되며 생성된 대화 전이 네트워크 일반성을 높이기 위한 네트워크 생성 규칙들을 제안하였다. 그리고 문장생성을 위해서는 선택된 사용자 발화문장에 대응하는 응답문장의 정보와 문맥 정보를 사용하여 시스템 응답문장을 생성한다.

#### 4. 통합 대화 관리자

##### 4.1 사용자 입력문장 분석

사용자 입력문장을 분석하기 위한 언어 이해 과정은 화행 분석, 행위 분석, 개체명 인식 등의 처리 과정을 수행해야 한다. 제안하는 방법은 사용자 입력문장과 말뭉치내의 사용자 발화문장들과의 유사도 계산을 통하여 선택된 사용자 발화문장의 의미표지 정보를 사용함으로써 이 과정을 간소화하였다. 유사도 계산은 다음 식(1)과 같이 코사인 유사도 (cosine similarity)를 사용하여 측정



[그림 185] 제안 시스템 구성도

하였고 측정에 사용된 자질들은 [표 1]에 나타내었다.

$$sim(IU, CU) = \frac{\sum_{i=1}^n iu\_w_i \times cu\_w_i}{\sqrt{\sum_{i=1}^n iu\_w_i^2 \times \sum_{i=1}^n cu\_w_i^2}} \quad (1)$$

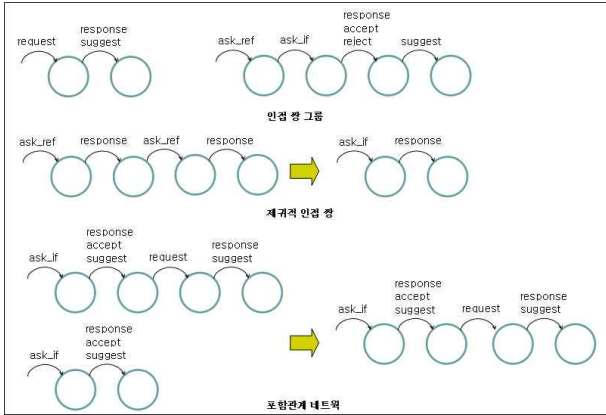
식 (1)에서 IU는 입력 발화문장을, CU는 코퍼스에 있는 발화문장을 각각 표시하고,  $iu\_w_i$ 는 입력 발화문장에서 출현하는 자질을 표시한다. 기본적으로는 이진가중치 (binary weighting) 기법을 사용했으며, [표 1]에서의 자질 중에서 개체명에 해당하는 문장 자질에 좀더 높은 가중치를 할당하였다

[표 1] 코사인 유사도 계산에 사용된 자질

문장 자질	문맥 자질
형태소	토픽
형태소 표지	-
개체명	-
개체명 표지	-

##### 4.2 대화 전이 네트워크

대화 전이 네트워크는 대화의 상태가 화행에 따라 전이되도록 구성된 유한상태 기계형태를 갖는다. 따라서, 대화 전이 네트워크는 사용자 발화와 시스템 응답의 상태를 예측하기 위한 모델로써 대화 관리에 사용된다. 기존의 연구는 대부분 고정된 스크립트의 형태로써 정형화된 대화를 처리하기 위해 사용되었다. 본 연구에서는 시스템의 확장성을 높이기 위하여 대화 전이 네트워크의 자동 생성 방법을 사용하였다. 말뭉치로부터 각 대화별로 화행의 시퀀스를 추출한 다음 [그림 3]에서 나타난 규칙들을 순차적으로 사용하여 추출된 시퀀스(sequence)들을 결합한다. 사용된 화행의 개수는 12개이며 인접 쌍



[그림 3] 화행 결합 규칙 적용 예

그룹으로 구성하기 위해, [표 2]에서와 같이 'request type' 과 'response type' 그리고 'single type' 으로 분류하였다. 인접 쌍 그룹은 동일한 'request type' 후에 나타나는 여러 타입의 'response types' 패턴을 말뭉치에서 추출하여 정의하였다. 따라서, 인접 쌍 그룹에 의해 시스템 응답의 범위가 결정된다. 재귀적 인접 쌍은 반복적으로 나타난 인접 쌍 그룹의 반복 횟수와 관계없이 동일하게 표시하였다. 포함관계 네트워크는 한 네트워크가 다른 네트워크에 포함되는 경우이며, 포함되는 네트워크는 대화 전이 네트워크 구성에서 제외한다. 그리고, 'single type'인 화행들은 대화 전이 네트워크에 포함되지 않으며 독립적으로 사용된다.

[표 2] 타입별 화행 분류

request type	response type	single type
ask_ref	suggest	inform
ask_if	response	acknowledge
ask_confirm	reject	closing
suggest	accept	opening

### 4.3 대화 전이 네트워크 운용

대화 시스템은 대화 전이 네트워크를 사용하여 현재 상태에서 선택될 수 있는 행위(action)들을 예측할 수 있다. 제안된 시스템에서 행위는 화행에 해당하므로 대화 전이 네트워크는 다음 상태로 전이 될 수 있는 화행의 후보를 결정한다. 그 다음 유사도 계산을 통하여 화행 후보들 중 한 개의 화행을 결정하고, 네트워크는 결정된 화행에 의하여 다음 상태로 전이된다. 또한, 대화 전이 네트워크는 유사도 계산 시 사용자 입력문장과 비교할 말뭉치 문장들의 범위를 줄이는 역할을 담당한다. 대화 전이 네트워크에서 결정된 화행 후보의 표지를 갖는 문

장들만이 유사도 계산의 비교 대상이 되기 때문에 유사도 계산의 정확도를 높일 수 있다.

화행을 결정하기 위하여 유사도 계산값의 임계점은 실험을 통하여 0.7로 결정하였다. 만약 유사도 계산 결과 가장 높은 점수를 갖는 문장의 점수가 임계점 이하이면 이 결정은 거부되며 새로운 화행 후보를 선택하여 유사도 계산한다. 이 경우는 대화의 주제가 전환된 것으로 판단하며 새로운 화행 후보들은 모든 대화 전이 네트워크 초기 상태에서 전이 될 수 있는 화행들이 선택된다. 이 과정을 선택된 모든 대화 전이 네트워크가 종료될 때까지 반복적으로 수행한다.

### 4.4 시스템 응답 생성

시스템 응답의 결정은 먼저 유사도 계산을 통해 말뭉치 문장이 결정되면 그 문장의 다음 차례 나타난 시스템 응답문장의 정보를 이용한다. 시스템 응답문장의 정보 중에 문장 유형에 해당하는 <TEMPLATE> 을 문장 틀로 사용하고 대화에서 유지하고 있는 문맥정보를 사용하여 시스템 응답문장을 생성한다. 시스템 응답은 유사도 계산의 결과에 따라 문장 유형이 결정되기 때문에 규칙 기반 생성모델 보다 사용자 입력문장에 따라 다양한 유형의 시스템 응답이 생성될 수 있다.

```

<ID>1-4
<SP>system
<DA>response
<MA>load_game
<KS>A게임 준비하겠습니다.
<NE>
<KW>
<TEMPLATE>game_name게임 준비하겠습니다.
<POS>game_name/SL+게임/NNG 준비/NNG+했/SSV+습니다/EP+습니다/EF+/SF
<DB>
<FORM>[0,1,0,0,0,0]
<EXECUTION>
    
```

[그림 5] 시스템 응답문장 정보의 예

## 5. 실험

실험을 위한 대화 말뭉치는 107개 대화로 구성되어 있으며 사용자발화와 시스템발화가 혼합되어 있는 1205 발화로 이루어져 있다. 말뭉치를 사용하여 생성된 대화 전이 네트워크는 [표 3]에서와 같이 결합 규칙을 모두 사용한 결과 45개의 네트워크가 생성되었다. 대화 전이 네트워크는 결합 규칙을 사용한 결과 서로 연관성이 있는 전이 관계를 추출하여 결합함으로써 크기가 작은 말뭉치에서 일반성이 높은 네트워크를 얻어 낼 수 있었다.

[표 3] 결합 규칙 적용에 따른 네트워크 개수

결합 규칙	개수
인접쌍그룹	75
인접쌍그룹+제귀적인접쌍	65
인접쌍그룹+제귀적인접쌍+포함관계	45

사용자 만족도를 위한 실험은 평가자 4명에게 3개의 작업 주제에 대하여 설명하고 각 주제 당 10개의 발화를 시도하게 하였다. 유사도 계산에 대한 만족도와 시스템 응답문에 대한 만족도 점수는 0에서 10까지의 점수로 표현하게 실험하였다. 유사도 계산에 대한 만족도와 시스템 응답문장에 대한 만족도는 모든 항목에서 70% 이상의 결과를 보여주었으며 대화의 주제가 변경된 경우 역시 높은 만족도를 보여 주었다.[표 4].

[표 4] 유사도 계산을 사용한 문장 선택 만족도

	유사도만족도	시스템 응답문 만족도
A	7.4	8
B	7.7	7.9
C	7.1	7.7
D	8.7	10
평균	<b>7.7</b>	<b>8.4</b>

0-2: 작업 수행 불가능, 3-6: 모호함, 7-10: 작업 수행 가능  
A, B, C, D는 평가자

## 6. 결론 및 향후과제

본 연구는 말뭉치로부터 추출된 정보를 사용하여 대화 시스템에 필요한 과정들을 통합 처리하는 시스템을 제안하였다. 사용자 입력문장과 말뭉치의 각 문장들 간의 유사도 계산을 통하여 의미적으로 가장 유사한 말뭉치 문장의 정보를 이용하고, 시스템 응답에 필요한 정보를 선택함으로써 언어 이해와 문장 생성 과정을 통합 처리하고 결합규칙을 통해 작은 크기의 말뭉치에서 높은 일반성을 갖는 대화 전이 네트워크를 자동으로 생성하였다. 따라서, 본 연구는 말뭉치를 사용하여 각 과정들을 통합적으로 업데이트함으로써 시스템의 유지 관리의 효율성을 높이고 대화 관리의 일반성을 증가시키는 기법을 제시하였다. 향후 연구로 새로운 말뭉치를 추가함에 따라 생기는 대화 전이 네트워크의 변화 여부와 말뭉치 문장들의 모호성 문제, 그리고 양식 기반 대화 등 여러 형태의 대화에 관련한 세분화된 실험이 앞으로 해결할 과제로 남아 있다.

## 감사의 글

이 연구(논문)는 지식경제부 지원으로 수행하는 21세기 프론티어 연구개발사업(인간기능 생활지원 지능로봇 기술개발사업)의 일환으로 수행되었습니다.

## 참고 문헌

- [5] H.A. Kautz and J.F. Allen. "Generalized Plan Recognition," In *Proceedings of the 14th National Conference on Artificial Intelligence*, pp. 32-37, 1986.
- [6] D. Goddeau, H. Meng, J. Polifroni, S. Seneff, and S. Busayapongchai, "A Form-Based Dialog Manager for Spoken Language Applications", In *Proceedings of the ICSLP 96*, pp. 701-705, 1996
- [7] L. Lambert and S. Carberry, "A Tripartite Plan-based Model of Dialogue," In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pp. 47-54, 1991.
- [8] L. Lambert, *Recognizing Complex Discourse Acts : A Tripartite Plan-Based Model of Dialogue*, Ph.D. thesis, University of Delaware, 1993.
- [9] N. Roy, J. Pineau and S. Thrun, "Spoken Dialog Management Probabilistic Reasoning," In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*, pp. 93-100, 2000.
- [10] S. Singh, D. Litman, M. Kearns and M. Walker, "Optimizing Dialogue Management with reinforcement Learning," *Journal of artificial Intelligence Research*, vol. 16, pp. 105-133, 2002