

# EPG 정보 검색을 위한 예제 기반 자연어 대화 시스템\*

김석환\* 이청재 정상근 이근배  
포항공과대학교 컴퓨터공학과 지능소프트웨어연구실  
{megaup, lcj80, hugman, gblee}@postech.ac.kr

## An Example-Based Natural Language Dialogue System for EPG Information Access

Seokhwan Kim\*, Cheongjae Lee, Sangkeun Jung and Gary Geunbae Lee  
iSoft Lab., Dept. of Computer Science and Engineering,  
Pohang University of Science and Technology

### 요 약

본 논문에서는 EPG 정보 검색을 위한 자연어 대화 시스템에 대해 논한다. 자연어 대화 시스템 구축을 위한, 대화 예제를 이용한 상황 기반 대화 관리 방법론은, 효율적이고 실용적인 대화 시스템 구축을 가능하게 한다. 대화 시스템은 사용자 발화에 대해 적합한 시스템 응답 발화를 출력하는 과정으로 진행되며, 이를 위해, 사용자 발화 의미 분석, 대화 관리, 시스템 응답 발화 생성의 과정을 거친다. 정확하고 신속한 정보의 전달이 중요한 EPG 정보 검색 도메인의 특성상 EPG 데이터베이스의 관리 및 갱신이 중요한 요소로 작용한다. 이를 위해 웹마이닝 기반의 EPG 데이터베이스 관리자를 구현함으로써 데이터베이스 구축에 필요한 비용을 최소화하고, 신속하고 정확한 정보를 제공할 수 있었다.

### 1. 서론

케이블 TV 및 위성 방송의 보급으로 인해 사용자들은 다양한 채널을 통해 다양한 방송을 시청할 수 있게 되었다. 하지만 그에 따라 원하는 채널과 프로그램을 리모컨만으로 찾는 일은 매우 어려워졌다. 디지털 방송 서비스의 시작과 함께 디지털 방송 서비스의 대표적인 어플리케이션으로 자리매김한 EPG(Electronic Program Guide)는 방송 편성표 및 방송 프로그램에 대한 상세한 정보를 디지털 TV 화면을 통해 제공함으로써, 원하는 채널과 프로그램을 좀 더 쉽게 찾을 수 있도록 도와주고 있다. 하지만 EPG 역시 리모컨 인터페이스의 한계로 인해, 제한적인 정보 검색 기능만을 제공하고 있으며, 이로 인한 복잡한 사용법은 사용자의 편의성을 저해하고 있다. 이 논문에서는 사용자 편의성을 고려한 효율적인 EPG 정보 검색을 위한 자연어 대화 시스템을 제안한다. 자연어 대화 시스템(Natural Language Dialogue System)은 자연어를 이용하여 컴퓨터와 의사소통을

하는 차세대 인터페이스이다. 자연어 대화 시스템은 많은 분야에서 이용되고 있으며 21세기의 유비쿼터스 시대를 위한 필수적인 지능형 인터페이스로 활발한 연구가 진행되고 있다[1]. 자연어 대화 시스템은 일반적으로 자연어 이해(Natural Language Understanding), 대화 관리(Dialogue Management), 자연어 응답 생성(Natural Language Generation) 등의 모듈로 구성된다. 본 연구에서는 EPG 영역의 데이터베이스 검색을 위한 사용자 인터페이스로서의 기능을 갖는 대화 시스템을 구현하였다.

EPG 정보 검색 영역이 다른 정보 검색 영역에 비해 갖고 있는 특성은 검색 대상 정보가 주기적으로 갱신되어야 한다는 점이다. 본 연구는 EPG 정보 데이터베이스의 주기적인 갱신을 위한 웹 마이닝(Web Mining) 기술을 포함한다. 인터넷의 발달과 보급으로 인해 웹을 통해 다양한 정보를 얻을 수 있으며, EPG 정보의 주를 이루는 방송 프로그램 관련 정보 역시 방대한 양의 정보가 웹 상에 존재하고 있다. 웹을 통해 제공되는 정보를 자동으로 추출하여 EPG 정보 데이터베이스를 구축하고 갱신함으로써, 데이터베이스 구축 및 갱신에 필요한 비용을 절감하고, 신속하게 정보를 갱신할 수 있다.

\* 본 연구는 산업 자원부 21C 프론티어 연구 “인간생활자원 지능로봇” 과제의 지원을 받아 수행되었음.

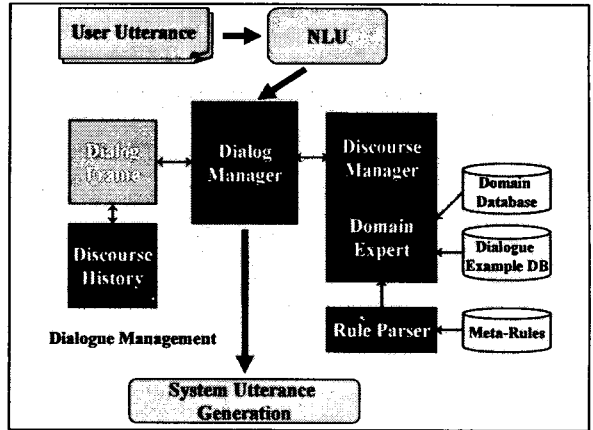
## 2. 대화 예제를 이용한 상황 기반 대화 관리 시스템

본 연구에서는 대화 예제를 이용한 상황 기반 대화 관리 (Situation-Based Dialogue Management with Dialogue Examples)[2] 방법론에 기반한 시스템을 개발하였다. 상황 기반 대화 관리 시스템의 목적은 실용적이며 유연한 대화 처리를 할 수 있고, 다양한 응용 분야에 적용 가능한 높은 도메인 확장성을 갖는 시스템을 만드는 데에 있다.

### 2.1 기존의 대화 모델

대부분의 기존 대화 관리 시스템은 유한 상태 기반 방식 (Finite State-Based Model)에 바탕을 두고 있다 [3]. 현재 대화 상태를 특정 상태로 사상을 하여 그 상태에서 적절한 대화 진행을 하고 다음 상태를 예측하는 방식이다. 이것은 실제로 고정된 상태 전이 모델 위에서 대화를 진행하기 때문에 특정 도메인의 목적에 맞는 상태 전이 네트워크를 디자인하여 빠른 대화 관리 시스템을 구축할 수 있다. 이 모델은 전형적으로 대화 형태가 고정되어 있는 대화 시스템에 많이 사용된다. 그러나 사용자의 현재 발화에 의해 다음 턴에 사용자의 발화가 제한이 되거나 도메인이 바뀌면 모델 자체를 바꿔야 하기 때문에 도메인 간의 확장성이 떨어지는 단점을 가지고 있다. 이를 개선하기 위해 제시된 방식이 프레임 기반 대화 모델 (Frame-Based Dialogue Model)이다[4]. 이 모델은 자연스러운 대화 처리를 위해 고안되었으며 사용자의 발화에서 필요한 정보를 추출하여 의미 프레임을 채우고 그 정보를 바탕으로 대화를 진행하는 방식이다. 그러나 실제 대화 관리는 이러한 의미 프레임을 이용한 상태 전이나 수작업으로 만든 규칙에 의존하는 경우가 많아 유한 상태 모델과 같은 단점을 가지게 된다. 또한 최근에는 복잡한 도메인을 처리하기 위해 계획 기반 모델 (Plan-Based Model)이 연구되고 있다 [5]. 이것은 사용자와 컴퓨터 간의 대화 목적을 정하고 그것을 위해 다양한 대화 계획 스키마를 적용하여 대화를 진행하는 방식이다. 그러나 이 방식은 사용자의 발화 자유도가 높고 어려운 작업을 위한 도메인을 처리하기 때문에, 대화 관리 성능의 저하로 상용성이 떨어진다는 단점이 있다.

최근에는 단순히 특정 도메인을 위한 음성 대화 시스템을 벗어나 도메인 확장성을 향상시키기 위한 객체 기반의 구현 방식이 연구되고 있다[6]. 이것은 대화 시스템의 각 모듈을 객체로 구현하여 도메인에 무관한 공통적인 처리 부분과 도메인에 관련된 처리 부분을 나누어 도메인 확장성을 향상시킨 것이다. 그러나 이 시스템도 대화 관리를 위해 수작업으로 많은 규칙을 인코딩해야 하기 때문에 실용적인 대화 관리 시스템 개발이 힘들다.



<그림 1> 상황 기반 대화 관리 시스템 구조

### 2.2 상황 기반 대화 관리 시스템

상황 기반 대화 관리 시스템 (Situation-Based Dialogue Management System)의 기본적인 아이디어는 프레임 기반의 대화 모델을 취한다. 전반적인 상황 기반 대화 관리 시스템의 구조는 <그림 1>과 같으며 기본적인 구조는 [6]의 구조와 동일하다. 여기서 대화 관리자(Dialogue Manager)는 필요한 도메인 전문가(Domain Expert)를 호출해 자연어 이해(NLU, Natural Language Understanding) 모듈에서 만들어진 의미 프레임과 외부 도메인 데이터베이스, 대화 예제 데이터베이스를 연결하여 현재 사용자의 상황에 맞는 적절한 시스템 발화를 생성한다. 이 논문에서 정의된 “상황 (Situation)”이라는 것은 현재 사용자의 발화와 의도, NLU 결과인 의미 프레임, 담화 이력(Discourse History) 등을 포함하는 전반적인 현재 대화 상태를 말한다. 상황 기반 대화 관리 시스템은 유한 상태 기반의 상태 전이를 지양하고 현재 대화 상황에서 자유로운 상태 전이를 허락한다. 그러므로 현재 턴에 규제된 대화 진행 규칙이 없고 사용자의 발화를 입력 받아 현재 대화 상황을 분석하게 된다.

### 2.3 예제 기반 대화 모델링

기존의 규칙 기반 대화 모델링에서는 개발자가 미리 지식 작업을 하여 일관성 있고 적용성이 큰 규칙을 인코딩 해야 하며, 효과적인 상황 기반 대화 관리를 위해서는 수작업에 의한 규칙 작성이 필요하다. 이러한 규칙이 일관성이 없거나 잘못된 경우에는 적절한 대화 진행을 하지 못하는 경우가 있다. 즉, 대화 관리 시스템에는 수많은 대화 상태가 존재하며, 이러한 대화 상태를 모두 처리할 수 있는 규칙을 수작업을 통해 만드는 것은 불가능하다.

이와 같은 규칙 기반 대화 모델링의 단점인 규칙

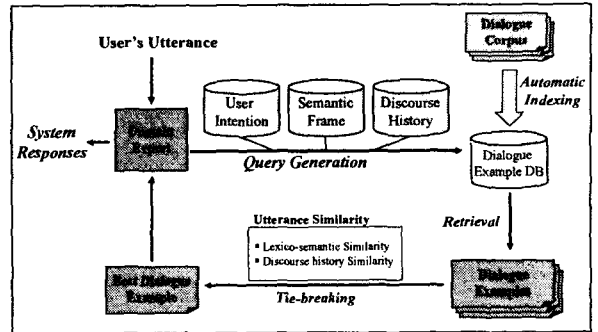
학습의 인력 비용을 간소화하기 위해 대화 말뭉치로부터 자동적으로 규칙을 학습하여 시스템 발화를 생성하는 대화 예제 기반 대화 모델링을 사용한다. 우선, 대화 모델링을 위해 대화 말뭉치로부터 자동적으로 대화 예제 데이터베이스(Dialogue Example Database)를 만든다. 우리는 말뭉치의 태깅을 최소화하기 위해 NLU 모델을 훈련하기 위한 말뭉치를 재사용하였다. 즉, 프레임 기반의 대화 관리 시스템을 구축하기 위해서는 기본적으로 사용자의 발화에 대한 의미 프레임 태깅 작업이 필요하다. 여기서는 대화 모델링을 위한 추가적인 태깅 작업으로 담화 이력 벡터(Discourse History Vector)와 시스템 응답(System Response)만을 추가적으로 태깅하였다. 여기서 담화 이력 벡터는 현재 대화 상황까지 의미 프레임의 슬롯이 채워졌는지 아닌지를 나타내는 이진 벡터로 표시한다.

이러한 태깅 정보들은 대화 예제 데이터베이스의 색인 키로써 이용되고 이를 가지고 대화 예제 데이터베이스에 질의를 던져 적절한 시스템 응답을 찾아낸다. 대화 예제를 찾기 위한 질의에서 현재 대화 상황을 바탕으로 검색한다면 항상 전체 일치(Exact Match)가 가능하지 않을 수 있다. 그러므로 부분 일치(Partial Match)가 가능하도록 제약을 완화하여 사용자의 의도(Dialog Act and Main Action)만을 가지고 대화 예제를 검색한다. 이것은 실제 시스템 응답이 현재 상황의 사용자 의도에 주로 의존하기 때문이다.

대화 예제 데이터베이스에서 현재 상황에 적합한 대화 예제들을 가져올 때, 여러 개의 예제 후보들이 생긴다. 이러한 것들 중에서 가장 적합한 예제를 찾기 위해서 발화 유사성(Utterance Similarity)을 정의하였다. 발화 유사성은 담화 이력 유사성(Discourse History Similarity)과 어휘-의미 유사성(Lexico-Semantic Similarity)의 선형 보간(Linear Interpolation)으로 정의하였다. 담화 이력 유사성은 담화 이력 벡터의 코사인 계수(Cosine Measure)로 나타내며, 어휘-의미 유사성은 추출된 슬롯의 값을 미리 정의된 슬롯 이름으로 대체하여 대화 예제의 사용자 발화와 현재 사용자 발화의 유사성을 측정하였다. <그림 2>는 어휘-의미 유사성을 계산하기 위한 입력 예제를 보여준다. 대부분 대화의 상황은 대화 예제를 통해 시스템 발화를 결정할 수 있지만, 대화 예제가 없는 경우 등을 처리하기 위해 미리 만들어 놓은 메타 규칙도 이용하였다. <그림 3>은 대화 예제 기반의 대화 모델링의 전반적인 전략을 나타낸다.

User Utterance	그림 SBS 드라마는 언제 하지?
Component Slots	[channel = SBS, genre = 드라마]
Lexico-Semantic Input	그림 [channel] [genre]는 언제 하지?

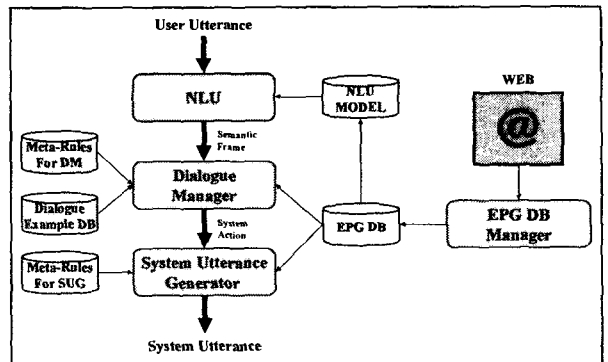
<그림 2> 어휘-의미 유사성 입력 예제



<그림 3> 대화 예제 기반 대화 모델링 전략

### 3. EPG 정보검색 대화 시스템

<그림 4>는 예제기반 대화 모델링 구조를 이용한 EPG 정보 검색을 위한 자연어 대화 시스템의 전반적인 구조를 나타낸다. EPG 정보 검색 대화 시스템은 각각의 사용자 발화 입력에 대한 시스템의 응답 발화를 출력하는 구조를 갖고 있으며, 이 때, 사용자 발화는 EPG DB 검색 요청에 대한 자연어 발화로 간주하며, 시스템 발화는 사용자의 요청에 대한 EPG DB 검색 결과를 주된 내용으로 포함한다. 이를 위해, 시스템은 사용자 발화의 의미를 분석하는 NLU 모듈, 의미 분석 결과와 기존의 담화 기록 등에 기반하여 적합한 시스템 발화 요소를 선정하는 대화 관리자(Dialogue Manager) 모듈, 선정된 시스템 응답 요소에 기반하여 실제 시스템 발화를 생성하는 시스템 발화 생성(System Utterance Generator) 모듈로 구성된다. 또한 EPG 데이터베이스 관리자(EPG Database Manager)는 웹상의 EPG 정보로부터 대화 시스템에 필요한 정보를 추출하여 EPG 데이터베이스를 구축 및 갱신하고, 시스템의 각각의 모듈들은 EPG 데이터베이스에 기반하여 구동된다.



<그림 4> EPG 정보 검색 대화 시스템 구조

### 3.1 NLU 모듈

대화 시스템을 위한 의미 분석 기술은 사용자의 발화로부터 의도나 상황을 분석하여 시스템이 처리 가능한 의미 프레임(Semantic Frame) 구조 형태로 구성하는 기술이다. 본 시스템에서 사용된 의미 프레임은 화행(Dialog Act)과 행위(Main Action), 그리고 구성 성분(Component Slot)으로 이루어진다. 화행은 일반적으로 문형 정보를 나타내며 상황에 따라서 반드시 문형과 일치하지는 않는다. 그리고 행위는 실제 사용자가 그 도메인 내에서 원하는 행위를 나타내는 의미적인 정보이다. 구성성분 요소는 실제 사용자 발화에 나타나는 도메인에 대한 개체 정보이다. 이것은 슬롯과 그것에 해당하는 값을 추출하여 표현 된다. <그림 2>의 사용자 발화의 경우 화행은 “wh-question” 이, 행위는 “search\_start\_time” 이 된다. 그리고 사용자가 말한 구성성분 요소는 채널(Channel)과 장르(Genre)에 대응하는 값으로 “SBS”와 “드라마”가 된다. NLU 모듈은 정보 추출 기술을 이용한 언어 이해 방법론에 기반하여 구현되었다[7].

### 3.2 대화 관리자 모듈

대화 관리자(Dialogue Manager) 모듈은 앞서 언급한 대화 예제를 이용한 상황 기반 대화 관리 방법론에 기반하여 구현되었다. 하지만 기본적인 상황 기반 대화 관리 방법론이 다양한 도메인의 대화를 처리할 수 있는 확장성에 초점을 맞추고 있는 것에 비해, 본 시스템은 EPG 정보 검색의 단일 도메인에 대한 대화 처리에 초점을 맞추고 있다. 그러므로 대화 관리자 내의 세부 모듈들은 EPG 정보 검색 도메인에 해당하는 각각의 단일 모듈들로 구성되어 있으며, 이는 앞서 언급한 상황 기반 대화 관리 구조의 단일 도메인에 대한 특별한 경우라 볼 수 있다. 하지만 기본적인 상황 기반 대화 관리 방법론의 구조에 따르고 있기 때문에, 추후 필요에 의해 별도의 도메인의 확장이 가능하다.

### 3.3 시스템 발화 생성 모듈

시스템 발화 생성 모듈은 대화 관리자의 결과인 시스템 발화 요소를 기반으로 자연어 시스템 응답 발화를 생성한다. 시스템 발화의 생성은 앞선 과정들로부터 얻은 의미 분석 결과와 시스템 응답 요소, EPG 데이터베이스 검색 결과에 기반하여 메타 규칙 기반으로 이루어진다.

### 3.4 EPG 데이터 베이스 관리자

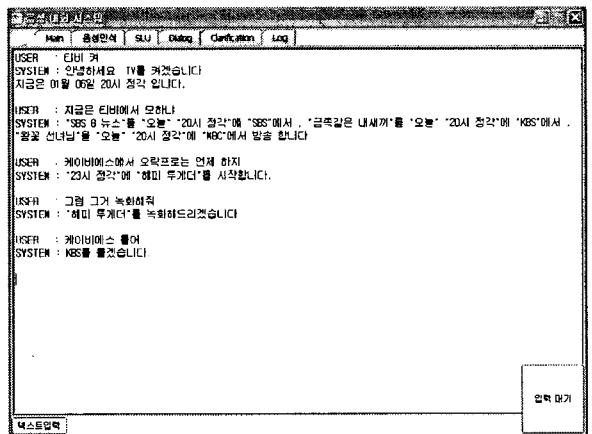
EPG 정보의 가장 큰 특성은 주기적으로 새로운 정보가 제공되며, 정보가 갖고 있는 시간적 특성이 각

시점에서의 정보의 가치를 좌우하는 기준이 된다는 점이다. 그러므로 EPG 정보 검색 시스템에서는 신속하고 정확한 정보의 갱신 및 관리가 중요한 요소이다. EPG 데이터베이스 관리자 모듈은 EPG 정보의 갱신을 위해 웹을 통해 얻을 수 있는 HTML 형식의 EPG 정보 테이블로부터 원하는 정보를 추출하여, 데이터베이스로 구축하는 기능을 수행한다. HTML 테이블로부터 정보를 추출하는 과정은 규칙 기반으로 이루어지며, 이 결과 생성된 EPG 데이터베이스는 NLU 모듈 구축과 대화 관리, 시스템 발화 생성의 대화 시스템 전반에 걸쳐 반영됨으로써, EPG 정보 검색 대화 시스템이 항상 최신의 정보를 제공할 수 있도록 한다.

## 4. 구현 및 실험

### 4.1 구현 결과

EPG 정보 검색을 위한 대화 시스템은 텍스트 입력과 함께 HTK 기반의 음성 인식기를 이용하여 음성 입력도 처리할 수 있도록 구현되었다. 시스템은 텍스트 입력이나 음성 인식 결과를 통해 입력된 사용자 발화에 대한 처리 결과를 자연어 형태의 시스템 발화로 출력한다. 이 시스템 발화를 음성 합성 모듈(TTS)을 이용하여 음성으로 합성함으로써 음성 입력과 음성 출력으로 이루어지는 음성 대화 시스템 (Spoken Dialog System)을 구성할 수 있다. <그림 5>는 EPG 정보 검색 대화 시스템의 실험 예제이다. 각각의 EPG 정보 검색 요청에 관련된 사용자 발화에 대한 시스템의 정보 검색 결과에 기반한 응답 발화가 출력되고 있다.



<그림 5> EPG 정보 검색 대화 시스템 실험 예제

## 4.2 실험

실험을 위해 EPG 정보 검색 영역의 대화 말뭉치를 이용하여 대화 예제 데이터베이스를 구축하였다. 이 대화 말뭉치는 380개의 사용자 발화로 이루어진 88개의 한국어 대화로 구성되어 있다. 각 사용자 발화는 NLU 모델을 위한 태깅이 되어 있고 여기에 대화 예제 기반 대화 모델을 구축하기 위해 담화 이력 벡터와 시스템 응답을 태깅하였다.

실험은 텍스트 입력과 음성 입력으로 나누어 진행하였으며, 음성 인식기는 HTK를 기반으로 EPG 정보 검색 영역의 언어 모델로 훈련이 되어 있다. 이번 실험 환경에서 단어 오인식률(WER)은 15.3%이며, 텍스트 입력과 음성 입력 각각에 대한 NLU 모듈의 성능은 <표 1>과 같다[7].

Slot Type	Textual Input (WER 0.0%)	Spoken Input (WER 15.3%)
Dialogue Act	95.33	85.34
Main Action	93.50	81.78
Component Slot	90.85	80.12

<표 1> NLU 성능

대화 예제 기반의 대화 모델링 능력을 측정하기 위해 대화 예제 일치 비율(Example Matching Rate, EMR)과 그것에 따른 턴 성공 비율(Success Turn Rate, STR)을 측정하였다. EMR은 사용자 발화 입력에 대한 대화 예제 일치의 평균 일치 비율로 전체 일치와 부분 일치로 나누어서 측정하였다. 그리고 그것에 따라 시스템 발화가 적절한지를 확인하기 위해 STR을 측정하였다. 실험을 위해 5명의 사용자에게 EPG 정보 검색 영역에 대한 10개의 임의의 텍스트 입력을 훈련에 사용한 88개의 대화와 독립적으로 넣도록 하였으며 각 턴마다 시스템 발화의 적절성을 평가하도록 하였다. <표 2>는 각 예제 일치 분류에 따라 EMR과 STR을 나타낸다. 실험 결과를 통해 대부분의 대화는 전체 일치나 부분 일치로 진행이 가능하며 훈련에 사용하지 않은 입력에 대해서도 대부분 (94%의 커버리지) 원만한 대화 진행이 이루어지는 것을 확인할 수 있다.

Example Match Type	EMR	STR
Exact Match	0.42	0.9
Partial Match	0.52	0.73
No Example	0.06	0.33

<표 2> EMR과 STR

또한 대화 관리 시스템의 성능 측정을 위해 텍스트 입력과 음성 입력에 대한 PARADISE 성능 측정법 [5]의 평가가 이루어졌다. 사용자 만족도 (User Satisfaction)는 음성 인식률 (Mean Recognition Accuracy, MRA)과 대화 성공률(User Perception of Task Completion Rate, TCR), STR의 값들에 각각 1/3씩의 가중치를 부여한 뒤, 선형 보간법을 이용하여 정의하였다. 평가는 5명의 사용자에게 5개의 다른 EPG 정보 검색 대화 시나리오에 기반한 대화를 수행하도록 하고, 각각의 턴과 대화에 대해서 시스템 발화의 정확성과 성공 유무를 평가하도록 하였다. <표 3>은 대화 성능 평가 결과를 나타낸다.

Evaluation	Textual Input	Spoken Input
TCR	0.92	0.76
STR	0.88	0.65
MRA	1.00	0.85
User Satisfaction	0.93	0.75

<표 3> 대화 성능 평가 결과

## 5. 결론

본 연구에서는 대화 예제를 이용한 상황 기반 대화 관리 방법론에 기반하여 EPG 정보 검색 대화 시스템을 개발하였다. 대화 예제를 이용하여 대화 관리를 위한 규칙을 자동으로 구축함으로써 효율적이고 실용적인 대화 시스템을 구축할 수 있었으며, 웹마이닝을 이용하여 EPG 데이터베이스를 자동으로 구축 및 갱신함으로써 데이터베이스 구축을 위한 비용을 절감할 수 있었다. 향후 계획은 EPG 데이터베이스 구축 모듈을 보강하여 테이블 형태의 자료뿐 아니라, 일반적인 자연어 문서에 대해서도 원하는 정보를 추출함으로써 좀 더 다양한 정보를 제공할 수 있는 EPG 대화 시스템을 개발할 예정이다.

## 참고 문헌

- [1] J. Allen, D. Byron, M. Dzikovska, G. Ferguson, L. Galescu, and A. Stent, "Towards conversational human-computer interaction", *AI Magazine*, Vol. 22, no. 4, pp 27-37, 2001.
- [2] Cheongjae Lee, Sangkeun Jung, Jihyun Eun, Minwoo Jeong, and Gary Geunbae Lee, A Situation-based Dialogue Management using Dialogue Examples, *Proceedings of the 2006 IEEE international conference on acoustics, speech and signal processing (ICASSP-2006)*, May 2006, Toulouse.
- [3] McTear, "Modeling spoken dialogues with state transition diagrams : Experiences with the CSLU

- toolkit” , *Proceedings of the 5th International Conference on Spoken Language Processing*, Vol. 4, pp. 1223-1226, 1998.
- [4] O. Lemon, A. Gruenstein et al., “Multi-tasking and collaborative activities in dialogue systems” , *Proceedings of the 3rd SIGDIAL Workshop on Discourse and Dialogue*, pp. 113-124, 2002.
- [5] J. Chu-Carroll, “MIMIC: An adaptive mixed initiative spoken dialogue system for information queries” , *Proceedings of the 6th Applied Natural Language Processing Conference*, pp. 97-104, 2000.
- [6] I. O’ Neil, P. Hanna et al., “Implementing advanced spoken dialogue management in Java” , *Speech Communication*, Vol. 54, no. 1, pp 99-124, January 2005.
- [7] J. Eun, C. Lee, and G. Lee. "An Information extraction approach for spoken language understanding", *Proceedings of the 8th International Conference on Spoken Language Processing*, pp. 2145-2148, 2004.