

# 영역 전환 전략을 사용한 다 영역 대화 프레임워크

최원석<sup>○</sup>, 강상우, 서정연\*

서강대학교 컴퓨터공학과, 서강대학교 컴퓨터공학과/바이오융합기술협동과정\*

panglister@gmail.com, {swkang, seojy}@sogang.ac.kr

## Multi-domain Dialog Framework using Domain Switching Strategy

Wonseok Choi<sup>○</sup>, Sangwoo Kang, Jungyun Seo\*

Department of Computer Science and Engineering, Sogang University

Department of Computer Science and Engineering, and Interdisciplinary Program  
of Integrated Biotechnology, Sogang University\*

### 요약

다 영역 대화 시스템 개발에서는 영역 확장이 쉬워야 하며 처리하는 대화 영역이 늘어나더라도 대화 과정에서 사용자 편의성을 유지해야 하는 점이 중요하다. 본 논문에서는 이런 특성을 가지는 다 영역 대화 시스템을 작성하기 위한 프레임워크를 제안한다. 이 프레임워크는 공통의 인터페이스를 구현하는 영역 전문가(Domain Expert) 기반으로 동작하므로 영역 확장이 용이하다. 또한 진행 중이던 대화를 종료하지 않은 채 타 대화 영역으로 이동하는 영역 전환(Domain Switching) 현상은 다 영역 대화를 복잡하게 만드는 주요한 원인 중 하나로써 이를 효과적으로 관리할 수 있는 영역 전환 전략을 사용하여 사용자 편의성을 확보하였다.

주제어: Multi-Domain Dialog Framework Expert

### 1. 서론

다 영역 대화 시스템은 여러 개의 영역에 대한 대화를 하나의 대화 시스템에서 모두 처리할 수 있도록 한 것이다. 최근 서비스 로봇 등이 이슈화 되면서 다양한 영역에 대해 대화를 진행할 수 있는 시스템에 대한 필요성이 증가하고 있으며 시스템이 다루어야 할 대화 영역도 다양해지고 있다.

처리할 대화 영역의 수가 증가하면서 다 영역 대화 시스템 개발에서 2가지 사항이 중요하게 부각되었다. 1) 영역이 확장되는 경우에도 기존 시스템의 변경을 최소화 하여야 하며 2) 대화 영역이 넓어지더라도 사용자가 대화를 진행하는데 어려움이 없도록 해야 한다는 것이다. 본 연구에서는 대화 영역이 늘어나면서 발생하는 문제들 중, 진행 중이던 영역 대화가 종료되지 않은 상황에서 타 대화 영역으로 대화가 전환되는 경우 발생할 수 있는 문제에 대해 초점을 맞추었다.

본 논문에서는 위에서 언급한 2가지 사항을 만족하는 프레임워크를 제안하고 이를 사용하여 4개의 대화 영역을 처리할 수 있는 대화 시스템을 구성하였다.

이 프레임워크에서는 대화 영역마다 영역 전문적인 지식을 가지고 대화를 처리할 수 있는 영역 전문가(Domain Expert)가 존재하며 각각 대화 영역에 대한 처리를 담당한다. 논문에서 제안하는 영역 전환 전략(Domain Switching Strategy)은 진행 중이던 대화 영역이 종료되지 않은 상태로 타 대화 영역으로 전환되는 경우 사용자가 대화 흐름을 놓치지 않도록 도와준다.

프레임워크의 구조와 영역 대화 전략의 보다 자세한 내용은 3장에서 다룰 것이며 4장에서는 프레임워크를 사용하여 구현한 대화 시스템에 대해 설명할 것이다.

### 2. 관련 연구

다 영역 대화 시스템에 대한 연구들은 대화 시스템을 객체 지향 구조로 설계하거나[1] 분산 아키텍처를 응용하여[2][3][4] 새로운 대화 영역을 쉽게 추가할 수 있는 방법들에 대해서 다루고 있다. 또한 대화 영역이 늘어나면서 사용자 입력 발화의 영역 결정 문제가 발화 인식 에러에 취약할 수 있는 문제에 대해서도 연구되었다[5]. 그러나 이런 연구들은 대화 시스템이 다루어야 할 대화 영역이 증가할 때 정량적으로 처리할 수 있는 능력에만 초점을 둔 것이며 1장에서 언급한 사용자 대화 편의성은 연구 범위에 포함되지 않는다.

### 3. 프레임워크 구조 및 동작

#### 3.1 전체 구조

그림 1은 논문에서 제안하는 프레임워크의 구조를 보여준다. 이 프레임워크에서는 각 대화 영역마다 하나의 영역 전문가가 할당되며 각 전문가는 다른 대화 영역에 대해서는 인지하지 못한다. 이 경우 영역 전문가는 서로 독립적으로 제작될 수 있으므로 확장이 쉬운 장점을 가진다. 모든 영역 전문가는 미리 정의된 인터페이스의 기능들을 구현해야 하며 이에 대한 보다 자세한 내용은 3.3절에서 설명한다.

음성 인식기는 사용자 음성 발화를 인식하여 문장 텍스트 형태로 바꾸어 주며 문장 분석기는 인식된 발화 텍스트에 대해 의미 분석을 수행하여 컨텍스트 매니저(Context Manager)에 결과를 기록한다.

컨텍스트 매니저는 대화 처리 과정에서 생성된 다양한 정보를 기록하고 관리한다. 분석된 사용자 입력 발화 정보도 이곳에 저장되며 새로운 정보가 추가되었을 경우 상위에 존재하는 대화 관리자(Dialog Manager)에게 알려 정보에 대한 처리를 시작하도록 하는 것도 컨텍스트 매니저의 역할이다.

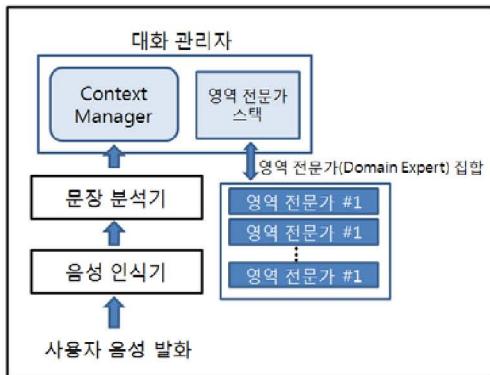


그림 1 프레임워크 전체 구조

대화 관리자는 구조의 최상위에서 전체적인 동작을 관리한다. 들어온 사용자 입력 발화가 어떤 대화 영역에 속하는지 결정하고 이에 따라 적절한 대화 영역 전문가가 입력 발화를 처리할 수 있도록 한다.

영역 전문가 스택은 사용자가 현재 진행 중인 영역의 대화를 종료하지 않고 타 영역으로 대화를 전환할 경우 영역 대화를 처리하던 영역 전문가를 미종료 상태로 보관한다. 저장된 영역 전문가들은 영역 전환 전략에 따라 처리된다.

새로운 사용자 입력 발화가 들어오면 대화 관리자는 이를 컨텍스트 관리자로부터 분석된 정보를 가져와 각 영역 전문가에게 보내준다. 각 영역 전문가는 입력 발화와 담당하는 대화 영역과의 연관 정도를 점수로 돌려주고 대화 관리자는 점수가 가장 높은 영역 전문가를 선택한다. 선택된 영역 전문가는 영역 전문가 스택이 비었을 때에는 최상위에 추가되고 사용자 입력 발화를 처리하면서 영역 대화를 진행한다. 만약 이미 동일한 모듈이 최상위에 추가되어 있다면 스택에 중복 추가는 하지 않는다.

영역 대화가 종료된 후에는 스택에서 해당 영역 전문가는 제거된다.

스택이 비어있지 않은 경우는 이전에 진행 중인 영역 대화가 존재하는 것이므로 영역 전환이 발생하였다고 간주하고 영역 전환 전략에 따라 추가적인 대화 관리 작업을 수행한다.

### 3.2 영역 간에 적용 가능한 대화 전략

다양한 대화 영역이 뒤섞여 진행될 수 있으므로 한 영역에 대한 대화가 종료되기 전에 다른 영역의 대화로 이동하는 영역 전환이 빈번하게 발생한다. 대화가 길어지고 이런 영역 전환이 많을수록 사용자는 이전 대화에 대해서 기억하기 어려우므로 시스템

과 사용자의 미 종료 대화에 대한 기억의 차이가 발생하기 마련이다. 이 때 사용자가 대화 흐름을 파악하고 효과적인 대화를 진행할 수 있도록 도와주는 대화 전략이 중요하다. 이런 대화 전략들은 서로 독립적인 대화 영역 간에 적용할 수 있어야 하므로 영역 내용과 무관하여야 하며 본 논문에서는 이를 영역 전환 전략으로 표기하였다.

표 1은 프레임워크에서 사용되는 영역 전환 전략들을 보여준다.

표 1. 영역 전환 전략 리스트

전략	내용
사용자 의도 확인	<p>적용 조건 :</p> <ol style="list-style-type: none"> <li>새로운 사용자 입력 발화가 들어온 경우.</li> <li>입력 발화에 대한 영역 연관 점수가 비슷한 경우.</li> </ol> <p>수행 :</p> <ol style="list-style-type: none"> <li>사용자에게 입력 발화가 연관된 영역에 대해 명시적인 확인을 요청.</li> <li>사용자 응답에 따라 영역을 결정.</li> </ol>
이전 대화 정리	<p>적용 조건 :</p> <ol style="list-style-type: none"> <li>영역 전환이 발생한 후 현재 영역 대화를 종료하고 이전 영역 대화로 돌아온 경우.</li> <li>이전 영역 대화가 전환된 후 시간이 오래 지난 경우.</li> </ol> <p>수행 :</p> <ol style="list-style-type: none"> <li>돌아온 영역에서 대화 상황을 간략히 기술.</li> <li>돌아온 영역의 마지막 발화 재출력.</li> </ol>
대화 종료 권유	<p>적용 조건 :</p> <ol style="list-style-type: none"> <li>영역전환이 발생한 경우.</li> <li>이전 영역 대화에서 종료 상태까지 얼마나 않은 경우.</li> </ol> <p>수행 :</p> <ol style="list-style-type: none"> <li>사용자에게 현재 영역 대화를 먼저 종료할 것을 요청.</li> <li>사용자 응답에 따라 영역전환 여부를 결정.</li> </ol>

표 1의 영역 전환 전략들은 3.1절에서 언급하였듯이 영역 전문가 스택에 이전에 진행 중이었던 영역 전문자가 존재할 때 적용되며 추가로 각 영역 전문가가 사용자 입력 발화에 대해 계산한 영역 연관 점수가 서로 비슷한 값을 가질 때에도 사용된다. 이 경우 사용자 의도 확인 전략이 적용되며 서로 다른 대화 영역이라 해도 사용하는 어휘가 유사할 수 있으므로 사용자 입력 발화가 어떤 대화 영역과 연관성을 가지는지 판단하기 어려운 경우에 유용하다.

대화 종료 권유 전략은 영역 전환을 제한하는 전략으로 불필요한 영역 전환이 너무 많이 발생하는 것을 막는다.

이전 대화 정리 전략은 영역 전환된 대화가 종료되어 이전 대화 영역으로 돌아온 경우 사용자에게 대화 상태에 대해 정보를 제공하도록 하여 사용자가 대화 흐름을 파악하는데 도움을 준다.

### 3.3 영역 전문가 공통 인터페이스

논문에서 제안하는 프레임워크는 대화 영역을 확장하기 쉽도록 영역 전문가가 구현해야 할 기능들을 공통의 인터페이스 형태로 구현하도록 하였다. 공통 인터페이스는 객체 지향 프로그래밍에서 사용하는 방법이며 대화 프레임워크 연구에서도 자주 사용된다[6]. 미리 정의된 인터페이스를 구현한 영역 전문가는 다른 전문가들과 동일한 방식으로 동작하는 것을 보장하므로 기존 시스템의 변경을 최소화하면서 새로운 대화 영역을 추가 가능 한 장점을 가진다. 표 2는 프레임워크에서 사용한 영역 전문가 공통 인터페이스를 보여준다.

표 2. 영역 전문가 인터페이스 내역

인터페이스 이름	내용
Evaluate	입력 : 사용자 입력 발화 동작 : 사용자 입력 발화와 영역 관련 정도를 나타내는 점수를 계산
Respond	입력 : 사용자 입력 발화 동작 : 영역 전문가에서 사용자 입력 발화에 대한 처리를 수행하고 결과를 리턴
Confirm	입력 : 사용자 발화 출력 : 영역에 맞는 사용자 의도 확인 전략을 수행하고 결과를 리턴
Remind	입력 : 사용자 발화 출력 : 영역에 맞는 이전 대화 정리 전략을 수행하고 결과를 리턴
Insist	입력 : 사용자 발화 출력 : 영역에 맞는 대화 종료 권유 전략을 수행하고 결과를 리턴

#### 4. 대화 시스템 구현

프레임워크를 이용하여 구현된 대화 시스템은 노인을 대상으로 하는 가정용 음성 도우미 시스템이다. 대화 영역은 간단한 건강 정보와 게임, 운세 및 일정 관리이며

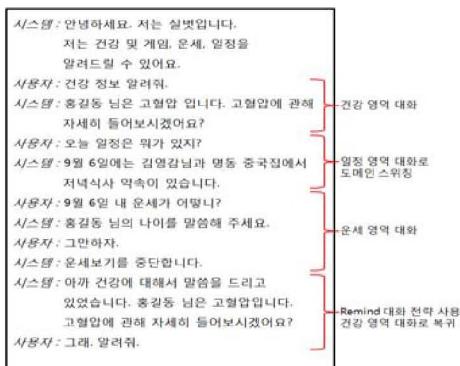


그림 2. 구현 대화 예제

대화 시스템은 4개의 영역 전문가로 구성된다. 사용자 입력 발화는 어휘의미패턴[7]을 사용하여 분석하였으며 대화에서 사용되는 총 패턴의 종류는 약 80개이다.

그림 2는 구현된 시스템의 대화 예제를 보여준다.

#### 5. 결론

본 논문에서는 영역 전문가를 사용하여 확장성과 사용자 편의성이 좋은 다 영역 대화 프레임워크를 제안하였다. 제안된 프레임워크는 전문모듈에 공통의 인터페이스를 적용하여 새로운 대화 영역을 쉽게 작성할 수 있으며 영역을 확장하는 경우에도 기존 시스템의 변경이 적은 장점이 있다. 또한 영역에 독립적인 영역 전환 전략을 사용하여 처리하는 대화 영역의 개수가 증가하더라도 사용자 편의성을 유지할 수 있다.

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

#### 참고문헌

- [1] O' Neill, P. Hanna, X. Liu, and M. McTear, Cross domain dialogue modelling: an object-based approach, In Proc. Interspeech-2004, pp. 205-208, 2004.
- [2] B. Lin, H. Wang, and L. Lee, Consistent dialogue across concurrent topics based on an expert system model, In Proc. Eurospeech-99, pp. 1427-1430. 1999
- [3] M. Hartikainen, M. Turunen, J. Hakulinen, E.-P. Salonen, J. A. Funk, Flexible dialogue management using distributed and dynamic dialogue control, In Proc. Interspeech-2004, pp. 197-200, 2004
- [4] Bor-shen Lin, Hsin-min Wang, and Lin-shan Lee, A distributed agent architecture for intelligent multi-domain spoken dialogue systems, IEICE Trans. on Information and Systems, E84-D(9):1217--1230, Sept. 2001
- [5] K. Komatani, N. Kanda, M. Nakano, K. Nakadai, H. Tsujino, T. Ogata, H. G. Okuno, Multi-domain spoken dialogue system with extensibility and robustness against speech recognition errors, In Proc. 7th SIGdial Workshop, pp. 9-17, 2001
- [6] M. Nakano, K. Funakoshi, Y. Hasegawa, H. Tsujino, A framework for building conversational agents based on a multi-expert model, Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue, pp. 88-91, 2008
- [7] Sangkeun Jung, Cheongjae Lee, Seokhwan Kim, Gary Geunbae Lee, DialogStudio: A Workbench for Data-driven Spoken Dialog System Developement and Management, Speech Communication, 50(8-9):683-697, Aug-Sept 2008.