

CSTAR-IF를 이용한 다국어 대화체 번역시스템

최운천

한국전자통신연구원 교환전송기술연구소 음성언어팀
대전시 유성구 가정동 161 우: 305-350,
ucchoi@etri.re.kr

Multi-Lingual Spoken Language Translation System using CSTAR-IF

Un-Cheon Choi

Spoken Language Processing Team,
ETRI-Switching & Transmission Technology Laboratory

요약

다국어 대화체 번역 시스템은 미국의 카네기 멜론 대학과 일본의 ATR 및 한국의 전자통신연구원 등이 가입한 CSTAR의 99년 국제간 음성언어번역 시스템 데모를 위한 한국어측 번역 시스템이다. CSTAR-IF는 국제간 데모를 위해 각 국의 시스템끼리 주고 받는 정보의 단위 혹은 형태로서, 중간언어 표현의 한 가지 방법으로 간단하면서도 단순한 표현으로 특정 영역 내에 나타나는 의미를 표현할 수 있도록 정의되었다. 다국어 번역 시스템은 크게 두 가지로 나누어 진다. 하나는 한국어 음성인식 결과를 IF로 변환하는 해석 시스템이고, 다른 하나는 IF로부터 한국어 문장을 생성하여 음성으로 들려주는 생성 시스템이다. 한국어 해석 시스템은 현재 92%의 해석 성공률을, 생성 시스템은 98%의 생성 성공률을 보이고 있다.

1 서론

미국의 카네기 멜론 대학과 일본의 ATR 및 한국의 전자통신 연구원 등이 가입한 CSTAR (Consortium for Speech Translation Advanced Research)¹에서는 CSTAR-IF(interchange format)[1]를 이용하여 99년 국제간 음성언어번역 시스템 데모를 할 예정이다. 그 데모에서 각 회원국은 자국의 언어를 음성 인식한 후 해석하여 IF를

만드는 해석 시스템과, IF로부터 자국의 언어를 생성하여 음성으로 들려 주는 생성 시스템을 책임진다. 그리고 필요하면 자국의 언어에 대한 해석 시스템과 생성 시스템을 다른 회원국에 제공할 수 있다. 이런 방식으로 회원국 상호간의 언어번역이나 자체 데모를 할 예정이다. 그림 1은 CSTAR-IF를 이용하여 한국과 또 다른 언어를 사용하는 한 국가와의 음성언어번역 시스템을 보인 것이다. 입력된 음성은 한국어 음성 인식기를 거쳐 한국어 텍스트로 변환되고, 그 텍스트는 한국어 해석 시스템을 거쳐 CSTAR-IF로 변환되어 인터넷을 거쳐 상대국의 시스템에 전달된다. IF를 받은 상대국의 생성 시스템은 그것을 그 나라 언어로 생성하여 음성으로 들려주게 된다. 전달된 IF가 질문일 경우, 그에 대응되는 답이 상대방 시스템에서 인터넷을 통해 다시 한국어 시스템으로 보내 올 것이다. 이와 같은 방법으로 국제간 대화가 가능하다.

IF는 크게는 인터넷을 통해 다른 나라의 음성언어번역 시스템과 연결하여 데이터를 주고 받는데 사용되고, 작게는 음성언어번역 시스템 내의 해석기와 생성기 사이에 데이터를 주고 받는데 사용된다. CSTAR-IF는 CSTAR 회원국들끼리 정보를 주고 받는 수단으로 정의된 것이다. 이 방법으로 기계번역을 하는 것은 전통적인 기계번역의 분류 중 중간언어 방식의 기계번역에 해당된다.

CSTAR-IF는 6개 언어(일본어, 영어, 독일어, 한국어, 이탈리아어, 프랑스어)를 사용하여 모아진 여행 안내 영역 즉, 호텔 예약, 비행기 예약, 여행지 안내 및 예약 등을 포함하고 있다. IF는 여행 안내 영역에 대한 데이터베이스로부터 6개 언어의 전문가들에 의해, 공통의 의미를 갖는 간단하면서도 명료한 표현방법을 정의한 것이다. 이것은 여행안내 및 호텔예약 영역이라는 특정 영역을 한정하기 때문에 가능했다. 한정된 영역에서 대부분의 발화는 제한된 수의 언어 행위(DA:dialog action or dialog

¹CSTAR에 대한 보다 자세한 내용은 CSTAR 홈 페이지 <http://www.is.cs.cmu.edu/cstar/> 참조.

act)[1], 즉 발화자의 의도로 구분할 수 있다. 예를 들면, 호텔 예약이 가능한지 여부를 묻는 행위, 호텔의 가격에 대한 정보를 주는 행위 등이다. 이런 언어 행위들이 모여서 IF를 만든다.

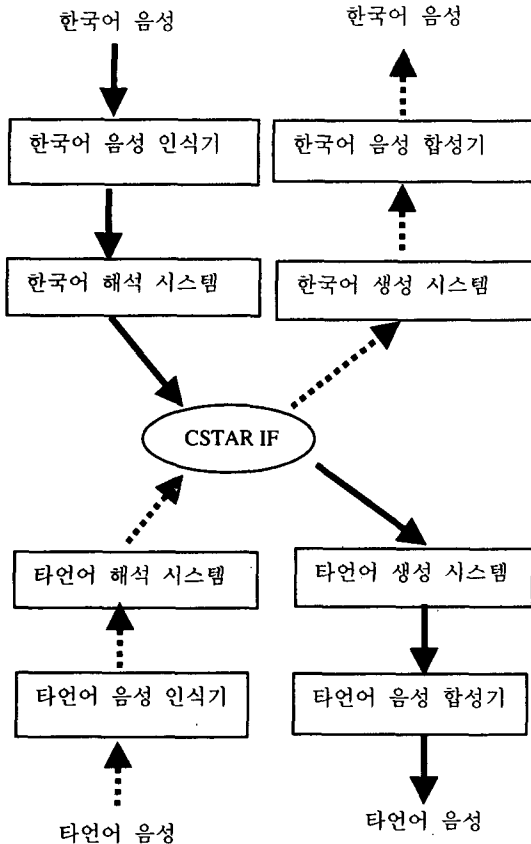


그림1. CSTAR IF를 이용한 다국어 번역시스템

2 한국어 해석 시스템

한국어 해석시스템은 한국어 음성인식기의 인식결과를 받아서 IF로 변환해 준다. 그림2는 한국어 해석시스템을 보인 구성도이다. 한국어 해석전처리기는 기준단어 사전을 이용하여 기준단어별로 입력된 문장을 분리한다[3,5]. 또한 명확히 잘못된 표현은 제거하고, 비슷한 표현은 하나의 표현으로 바꿔준다. 기준단어 사전에는 의미를 파악하는데 핵심이 되는 단어나, 분리할 경우에 해석문법 작성에 유리한 용언들이 들어있다.

SOUP 해석기는 미국의 카네기 멜론 대학에서 개발한 해석기로서 CFG(context-free grammars)를 다루며, 여러 영역의 의미문법(semantic grammars)을 실시간에 처리해 주는 통계 추론(stochastic), 차트기반(chart-based), 탑-다운(top-down) 방식의 파사이더, 대화체 문장 처리를 위해 개발되었다. SOUP 해석기는 해석문법을 이용하여 입력된 문장을 해석하여 파싱트리를 출력해 주는 역할을 한다². 해석문법은 DA(Dialog Act)[1] 별로 규칙을 기술하고, 속성과 값에 해당하는 규칙도 기술하여 준다.

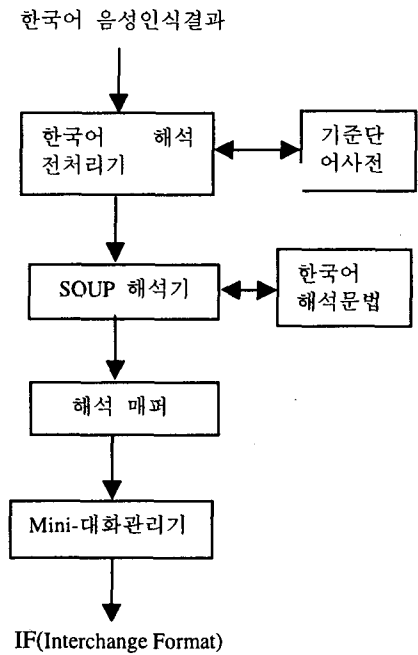


그림2. 한국어 해석 시스템의 구성도

해석매핑은 계층구조인 SOUP 해석기의 파싱트리를 단층 구조인 IF로 변환한다. 그리고 Mini-대화관리기는 만들어진 IF가 모호성(ambiguity)에 의해 잘못된 해석된 경우를 수정한다. 예를 들면, 'introduce-self'란 화행은 인사말로 자신을 소개하는 것으로 대화의 앞부분에만 나와야 하는데, 모호성에 의해 대화의 중간에도 나타날 수 있다. 또한 한국어의 "안녕"이란 표현은 만났을 때 하는 인사말을 의미하는 화행 'greeting'이 될 수도 있고, 반대로 헤어질 때 하는 인사말인 'closing'이 될 수도 있다. 대화의 어느 부분에 나타나는냐에 따라 화행이 달라질 수 있다. 이와 같이 Mini-대화 관리기는 모호성에 의해 잘못된 IF를 올바르게

² SOUP에 대한 자세한 내용은 웹 페이지 http://www.is.cs.cmu.edu/ISL_speech_parsing_soup.html 참조.

고쳐주는 역할을 한다. 현재의 대화 관리기가 단지, 화행이 나타나는 위치와 기존의 IF가 무엇인지에 따라 모호성을 해결하는 아주 단순한 알고리즘만 이용하기 때문에 mini라고 이름을 붙였다. 그러나 앞으로 모호성 처리를 보다 정확하기 위해 대화관리 기능을 보완할 예정이라서 대화 관리기라 명명하였다.

한국어 해석 시스템의 실행 예를 보이면 다음과 같다.

입력(음성인식의 결과): 제가 미국 피츠버그로 여행을 하려고 하는데요 호텔을 예약하고 싶습니다 저는 시월 십오일부터 십팔일까지 머무르고 싶습니다

한국어 해석 전처리기 결과: 제가 미국 피츠버그로 여행을 하려고 하는데요 호텔을 예약하고 싶습니다 저는 시월 십오일부터 십팔일까지 머무르고 싶습니다

SOUP 해석기의 출력(파싱트리):

[give-information+destination+trip] (제가 [destination=] ([pittsburgh] (미국 피츠버그)))로 여행을 하려고 하는데요)

[give-information+reservation+features+hotel] (호텔을 예약 하고 싶습니다)

[give-information+temporal+stay] (저는 [time=] ([start-time=] ([point=] ([month=] ([october] (시월))) [md=] ([md15] (십오일))) 부터) [end-time=] ([point=] ([md=] ([md18] (십팔일)))) 까지) 머무르고 싶습니다)

해석 매퍼의 결과(IF):

c:give-information+destination+trip (destination=pittsburgh)
 c:give-information+reservation+features+hotel
 c:give-information+temporal+stay (time=(start-time=(october, md15), end-time=md18))

위의 예에서 보듯이 하나의 발화는 여러 개의 IF를 만들어 낼 수 있다. CSTAR 회원국끼리 주고 받는 정보는 발화 단위이고, 한 발화에 나타나는 모든 IF를 보내 주어야 한다. 한국어 해석 전처리기 결과 중 밑줄이 쳐진 부분은 기준단어로 분리된 것을 보여 준 것이다.

3. 한국어 생성 시스템

한국어 생성 시스템은 다른 회원들에게서 받은 IF를 이용하여 한국어 텍스트를 만들어 내는 역할을 한다. IF 모호성 처리기는 IF 정의 자체의 모호성으로 인해 나타나는 모호성을 처리하기 위한 것이다. 예를 들면, '누구'를 의미하는 속성 'who='의 값 'i'는 "저는"으로, 글자를 의미하는 속성 letter=의 값인 'i'는

"아이"로 생성되어야 하나, 현재의 생성문법에서는 그 둘을 구분할 수가 없다. 그래서 IF의 속성 letter=의 값인 'i'를 'ii'로 변환하여 주고, 생성문법에 속성 letter=의 값은 'i' 대신 'ii'로 적으면 모호성을 처리할 수 있다.

생성 매퍼는 단층구조의 IF를 생성기가 이해할 수 있는 계층구조의 생성 트리로 변환해 준다. 그리고 생성기는 한국어 생성문법을 이용하여 입력된 IF를 한국어 텍스트로 변환하여 준다. 한국어 생성문법은 해석문법과 비슷한 방법으로 작성한다. 그러나, 해석문법에는 하나의 IF표현이 여러 개의 규칙을 가질 수 있으나, 생성문법에서는 오직 하나의 표현만 가질 수 있다. 한국어 생성 후처리기는 생성기의 결과 중 어색한 부분을 자연스럽게 수정해 준다. 대표적인 경우가 조사어의 이행태 처리이다.

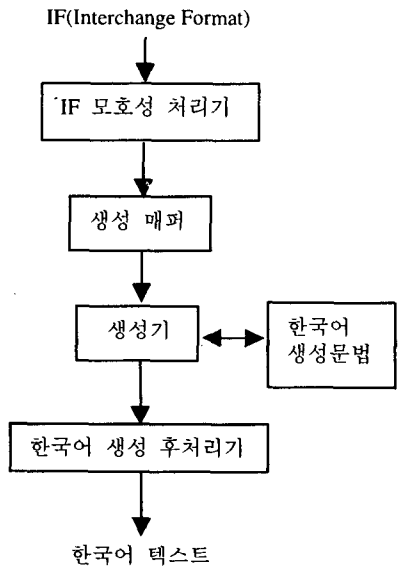


그림3. 한국어 생성 시스템의 구성도

한국어 생성 시스템의 실행 예를 보이면 다음과 같다.

입력 IF:

c:give-information+temporal+stay (time=(start-time=(october, md15), end-time=md18))
 c:give-information+destination+trip (destination=pittsburgh)
 c:give-information+reservation+features+hotel

생성 매퍼의 결과(생성 트리):

```
[c==give-information+temporal+stay] ([time=] ([start-time=]
([point=] ([month=] ([october=] ()) [md=] ([md15] ())))
[end-time=] ([point=] ([md=] ([md18] ())))))
```

```
[c==give-information+destination+trip]
([super-destination=] ([destination=] ([pittsburgh]
( ) ) ) )
```

```
[c==give-information+reservation+features+hotel] ( )
```

생성기의 결과:
 피츠버그로 여행을 하고 싶은데요.
 저는 호텔을 예약하려고 합니다.
 시월 십오일 부터 십팔일 까지 묵을 예정입니다.

한국어 생성 후처리기의 결과(한국어 생성문):
피츠버그로 여행을 하고 싶은데요.
저는 호텔을 예약하려고 합니다.
시월 십오일부터 십팔일까지 묵을 예정입니다

위의 결과 중 밑줄이 있는 부분은 생성 후처리기에 의해 수정된 부분을 의미한다.

4. 한국어 해석 생성 시스템의 성능

4.1 IF 태깅 데이터베이스

CSTAR 회원들끼리 미리 정의된 IF로 정보를 주고받기 때문에 IF 태깅은 CSTAR에서 제시한 원칙에 의해 실시되어야 한다. 한국어 IF 태깅 데이터베이스는 CSTAR에서 제시한 원칙에 근거하여 전문가들이 수동 태깅하여 현재 160 대화가 구축되어 있다. 160 대화는 약 2000개의 발화와 5000개의 IF로 이루어져 있다. 한 대화당 평균 발화수는 9이다. 그리고 하나의 발화는 약 평균 2.5개의 IF로 나뉘어진다. CSTAR에서 제시하지 않은 부분에 대해서는 필요하면 새로 IF를 만들어 CSTAR에 보내 검증을 받아서 사용해야 한다.

4.2 한국어 해석 시스템의 성능

한국어 해석 시스템의 성능을 알아보기 위해 수동 태깅된 것과 해석 시스템의 결과를 비교하였다. 먼저 해석문법 작성에 사용한 125대화를 이용하여 평가한 결과는 92.1%의 해석 성공률이었다. 여기서 발생하는 예러는 주로 모호성과 복잡한 낱자 표현, 여행 안내 외의 다른 영역의 표현 등이다. 다음으로 해석문법 작성에 사용하지 않은 35대화에 대해 평가한 결과는 54.1%의 해석 성공률이었다. 같은 영역에 대한 데이터인데도 해석 성공률이 낮게 나오는 이유는 아직 해석문법이 완전하지 못하기 때문이다. 특히 미등록어 처리나, 간투사가 끼어드는 발화에 대한 처리가 아직 미흡하다. 보다 나은 해석 시스템을 위해 한국어의

특징인 부분 자유어순을 처리할 수 있도록 해야 하며, 미등록 표현에 대한 대처 방안도 고려되어야 한다. 아울러 모호성 처리 부분도 추가되어야 한다.

4.3 한국어 생성 시스템의 성능

생성 시스템의 평가는 수동 태깅된 IF 데이터베이스의 IF로부터 생성 시스템을 실행하여 나온 생성문과 원래의 한국어 문장과의 비교를 통해 이루어졌다. 판단의 기준은 평가자의 주관에 의존한다[4,5]. 평가자가 원문과 생성된 결과가 의미적으로 큰 차이가 없으면 생성 성공으로 간주한다는 기본 원칙만 평가 기준으로 제시하였다. 현재 생성 문법은 50대화에 대해서만 작성되었고, 그 50 대화에 대한 생성 시스템의 평가 결과는 98%이다. 나머지 2%의 예러는 생성 매핑이 복잡한 병렬처리를 생성기가 원하는 형태로 변환해 주지 못 하기 때문에 주로 나오고 있다. 아직은 작은 양에 대해 생성문법이 작성되었기 때문에 98%가 큰 의미는 없지만, 생성 시스템은 해석에 비해 상대적으로 단순하기 때문에 95% 이상의 생성 성공을 기대할 수 있다.

5. 마무리

이 논문은 CSTAR의 99년 국제간 음성언어번역 시스템 데모를 위한 한국어측 번역 시스템에 대해 기술한 것이다. 국제간 음성언어 번역 시스템은 CSTAR-IF라는 정보의 단위를 이용하여 국제간에 정보를 주고 받는다. 다국어 번역 시스템은 한국어 음성인식 결과를 IF로 변환하는 해석 시스템과, IF로부터 한국어 문장을 생성하여 음성으로 들려주는 생성 시스템으로 구분할 수 있다. 한국어 해석 시스템은 해석문법 작성에 사용된 데이터에 대해 현재 92%의 해석 성공률을, 생성 시스템은 98%의 생성 성공률을 보이고 있다. 그러나 해석문법 작성에 사용하지 않은 데이터에 대해서는 60% 미만의 해석 성공률을 보이고 있어 보완이 필요하다. 보완 방안으로 사전에 무의미어 제거, 미등록어 처리, 해석문법 보완, 모호성 처리의 확장 등을 고려하고 있다.

감사의 글

이 연구는 정보통신부의 지원에 의해 이루어진 결과물입니다.

참고문헌

- [1] 최운천, "다국어 대화체 음성언어 번역시스템을 위한 IF와 IF태깅" 제15회 음성통신 및 신호처리워크샵 논문집, pp.409-412, 1998.
- [2] Nam-Yong Han, Un-Cheon Choi and Youngjik Lee, "An Implementation of a Partial Parser in the Spoken Language Translator", The 1998 International Conference on Acoustics,

Speech and Signal Processing(ICASSP'98), Vol. I, pp. 205-208, 1998.

- [3] Nam-Yong Han, Un-Cheon Choi, and Jae-Hoon Kim, "A Machine Translation of Korean Spoken Language Translation System," 17th International Conference on Computer Processing of Oriental Languages (ICCPOL'97), Vol. II, pp. 460-463, 1997.
- [4] Nam-Yong Han and Un-Cheon Choi, "Evaluation Results of Concept-based Translator with Partial Parsing", Proceedings of the Natural Language Processing Pacific Rim Symposium 1997 (NLPRS'97), pp. 513-516, 1997.
- [5] 최운천, 한남용, 김재훈, "대화체 음성언어 번역시스템에서의 개념기반 번역 시스템" 한국정보처리학회 논문지, 제4권 제 8호, pp.2025-2037, 1997.
- [6] Alon Lavie, Lori Levin, Puming Zhan, Maite Taboada, Donna Gates, Mirella Lapata, Cortis Clark, Matthew Broadhead, Alex Waibel, "Expanding the Domain of a Multi-Lingual Speech-to-Speech Translation System," Proceedings of the Workshop on Spoken Language Translation, 1997.
- [7] Alon Lavie, Alex Waibel, Lori Levin, Michael Finke, Donna Gates, Marsal Gavald Torsten Zeppenfeld, Puming Zhan, " JANUS III: SPEECH-TO-SPEECH TRANSLATION IN MULTIPLE LANGUAGES," IEEE International Conference on Acoustics, Speech and Signal Processing, 1997.