

이전 문장 자질과 다음 발화의 후보 화행을 이용한 한국어 화행 분석

(Korean Speech Act Tagging using Previous Sentence Features and Following Candidate Speech Acts)

김 세 종 [†] 이 용 훈 [†] 이 종 혁 ^{††}
(Se-Jong Kim) (Yong-Hun Lee) (Jong-Hyeok Lee)

요 약 화행 분석이란 자연언어로 된 발화를 통해서 나타나는 화자의 의도를 인식하는 것으로 대화를 처리하는 여러 응용 분야에서 중요하게 요구되는 과정이다. 기존의 연구에서는 이전 발화의 화행과 현재 발화의 문장 자질을 사용하여 규칙 기반 및 통계 기반의 연구가 진행되었다. 하지만 본 논문에서는 현재 발화 이후의 후보 화행을 추정하여 이를 현재 발화의 화행을 결정하는데 사용함으로써 기존의 연구와 차별화를 두었으며, 실제로 기존 방법보다 3.65%의 성능 향상을 보인 95.27%의 정확도(accuracy)를 보였다. 또한 이전 발화의 화행 정보뿐만 아니라 이전 화행을 결정하는데 사용했던 문장 자질을 현재 발화의 화행 결정에 추가적으로 사용함으로써 이전 발화의 화행 결정을 통해 소실된 정보들의 활용을 최대화하였다. 마지막으로 화행의 종류에 따라 그에 알맞은 확률 모델을 단계별로 적용하여 최종적으로 97.97%의 정확도를 얻었다.

키워드 : 화행 분석, 후보 화행, 문장 자질, 카이제곱 통계량, SVM

Abstract Speech act tagging is an important step in various dialogue applications, which recognizes speaker's intentions expressed in natural language utterances. Previous approaches such as rule-based and statistics-based methods utilize the speech acts of previous utterances and sentence features of the current utterance. This paper proposes a method that determines speech acts of the current utterance using the speech acts of the following utterances as well as previous ones. Using the features of following utterances yields the accuracy 95.27%, improving previous methods by 3.65%. Moreover, sentence features of the previous utterances are employed to maximally utilize the information available to the current utterance. By applying the proper probability model for each speech act, final accuracy of 97.97% is achieved.

Key words : Speech Act Tagging, Candidate Speech Act, Sentence Feature, CHI Statistics, SVM

- 본 논문은 지식경제부 및 정보통신연구진흥원의 정보통신선도기반기술개발사업과 2008년도 두뇌한국21사업의 지원을 받아 수행되었습니다.
- 이 논문은 제19회 한글 및 한국어 정보처리학술대회에서 '이전 문장 자질과 다음 발화의 후보 화행을 이용한 한국어 화행 분석'의 제목으로 발표된 논문을 확장한 것임

[†] 학생회원 : 포항공과대학교 컴퓨터공학과
sejong@postech.ac.kr
yhlee95@postech.ac.kr

^{††} 종신회원 : 포항공과대학교 컴퓨터공학과 교수
jhlee@postech.ac.kr

논문접수 : 2008년 1월 3일
심사완료 : 2008년 4월 11일

Copyright©2008 한국정보과학회 : 개인 목적이거나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지 : 소프트웨어 및 응용 제35권 제6호(2008.6)

1. 서론

화행이란 화자의 발화를 통해서 나타나는 화자의 의도를 가리키는 것으로 이를 올바르게 인식하는 과정을 화행 분석이라고 한다. 화행 분석은 대화를 처리하는 여러 응용 분야에서 매우 중요하게 요구되는 과정으로 특히 목적 지향적 에이전트를 구현하는데 있어서 없어서는 안 되는 요소이다. 현재 목적 지향적 에이전트로 활용되고 있는 분야는 기차 예약 에이전트, 항공 예약 에이전트, 인터넷 쇼핑 도우미 등이 있으며 이러한 대화 에이전트는 사용자가 입력한 발화로부터 해당 도메인에 필요한 정보를 추출하고 화행 분석을 통한 화자의 의도를 파악한 뒤, 시스템 상에서 정의된 대화 전략에 따라 응답을 생성한다[1-3].

그러나 화자는 자신의 의도가 담긴 발화를 다양한 표현으로 상대방에게 전달할 수 있으며 이러한 특성은 시스템이 화자의 의도를 완벽하게 파악한다는 것이 무척 어렵다는 것을 보여준다[4]. 실제 연구 분야에서도 특정 도메인에 한정된 화행 분석은 비교적 높은 성능을 보인 반면, 도메인에 독립적인 화행 분석은 상당히 낮은 성능을 보였다[5-7]. 이는 화행 분석의 한계성을 나타냄과 동시에 이에 대한 더 많은 연구가 요구됨을 뜻한다. 화행의 종류를 결정하는 것도 화행 분석의 성능에 영향을 미치지만 도메인에 독립적인 화행을 결정하는 것 또한 대화 처리 분야의 큰 이슈로서 연구되고 있으므로 본 논문에서는 이에 대한 언급을 따로 하지 않겠다[8].

화행 분석은 '대화를 처음 처리하는 단계'라는 입장에서 볼 때 매우 높은 성능을 보여야 한다. 자연언어처리의 중요 분야인 형태소 분석이 그러하듯이 초기 단계의 성능이 높을수록 이후 단계의 성능이 향상될 수 있기 때문이다. 기존의 연구에서는 이전 발화의 화행과 현재 발화의 문장 자질을 사용하여 규칙 기반 및 통계 기반의 연구가 진행되었다. 하지만 본 논문에서는 현재 발화 이후의 발화에 대한 후보 화행을 추정하여 이를 현재 발화의 화행을 결정하는데 사용함으로써 기존 연구 결과에 대한 성능 향상을 꾀했다. 또한 이전 발화의 화행 정보뿐만 아니라 이전 화행을 결정하는데 사용했던 문장 자질을 현재 발화의 화행 결정에 추가적으로 사용함으로써 이전 발화의 화행 결정을 통해 소실된 정보들의 활용을 최대화하였다. 그리고 화행의 종류에 따라 그에 알맞은 확률 모델을 단계별로 적용하여 최적의 성능을 발휘하도록 하였다.

본 논문의 구성은 다음과 같다. 2장에서는 화행 분석에 관련된 기존의 연구들을 살펴보고 3장에서는 제안하고자 하는 세 가지 방법에 대해서 자세하게 설명한다. 4장에서는 실제 사용자와 항공 예약 전문가 간의 대화 내용을 기록한 SRI ATIS Collection[9]의 일부를 번역 및 정제된 코퍼스를 사용하여 실험 결과를 제시하고 5장에서는 결론을 맺는다.

2. 기존 연구

2.1 규칙 기반 모델

규칙 기반 모델이란 시스템 설계자가 화행을 결정하기 위한 규칙들을 직접 작성하는 것으로, 해당 도메인의 지식을 포함하는 언어 정보 규칙과 문맥 규칙을 사용하여 화행을 결정하는 방법이다[10,11]. 대표적인 방법으로는 화행 분석 단계를 표층적 화행 분석과 심층적 화행 분석으로 나누어 해당 발화의 후보 화행을 점차 줄여가는 것이다[10].

표층적 화행 분석이란 문맥을 고려하지 않고 표층적

으로 나타난 정보만을 가지고 모든 후보 화행을 결정하는 것을 말한다. 즉, 입력된 발화에 대해 형태소 분석, 구문 분석, 의미 분석 등을 거쳐 형성된 정보들을 사용하여 입력 발화의 형태만으로 결정할 수 있는 모든 가능한 후보 화행들을 제시하는 단계이다. [10]에서는 표층적 화행 분석을 위해 2단계의 화행 분석을 수행하는데, 1차 단계는 미리 제시된 화행 결정 규칙을 통해 해당 발화에 대한 하나의 후보 화행을 결정하고 2차 단계는 1차 단계의 결과를 토대로 가능한 표층적 화행들을 추가한다.

심층적 화행 분석이란 대화의 흐름을 고려한 문맥 정보를 사용하여 표층적 화행 분석에서 결정된 화행들 중 하나의 화행만을 선택하는 과정이다. 심층적 화행 분석에서 사용된 문맥 정보는 담화 지식만으로 구현된 재귀적 대화 전이망(RDTN, Recursive Dialogue Transition Network)과 심층적 화행 결정 규칙으로 이루어진다. 재귀적 대화 전이망은 일종의 오토마타(automata)로서, 이를 이루고 있는 노드(node)는 현재 대화의 상태(state)를 나타내고 엣지(edge)는 입력 발화의 화행 흐름을 나타낸다. 심층적 화행 결정 규칙은 표층적 화행과 재귀적 대화 전이망의 상태, 그리고 이전 발화와 현재 발화간의 문장 자질을 비교하여 최종 화행을 결정하는 규칙이다. 이러한 규칙 기반 모델은 화행 결정에 있어서 매우 높은 성능(93.75%, 12개 화행 사용)을 가져왔으나 도메인 지식을 획득하는데 많은 시간이 소모되고, 다른 도메인으로 확장할 경우 해당 도메인에 대한 규칙을 다시 구축해야 하는 단점을 가지고 있다.

2.2 통계 기반 모델

통계 기반 모델은 규칙 기반 모델의 단점을 보완한 것으로, 화행 분석을 위한 규칙들을 직접 작성하는 대신 대량의 대화 코퍼스를 사용하여 기계 학습을 하고 학습된 결과를 바탕으로 각 발화의 화행을 결정하는 방법이다. 본 모델은 기본적으로 다음과 같은 수식을 토대로 설계되었으며 S_i 는 i 번째 화행을, U_i 는 i 번째 발화를 의미한다[5,12].

$$\begin{aligned} S(U_{1...n}) &= \operatorname{argmax}_{S_{1...n}} P(S_{1...n} | U_{1...n}) \\ &\approx \operatorname{argmax}_{S_{1...n}} \prod_{i=1}^n P(S_i | S_{1...i-1}) P(U_i | S_i) \end{aligned}$$

수식 1 통계 기반 모델

통계 기반 모델은 수식 1을 이루는 각각의 확률값을 다른 형태로 변형하거나 어떠한 방법으로 학습할 것인가에 초점을 맞추고 연구되었다. N-gram, 최대 엔트로

피 모델(Maximum Entropy Model), 은닉 마코프 모델(Hidden Markov Model), 결정 트리 모델(Decision Tree Model), 신경망 모델(Neural Network Model) 등이 기존 연구에서 활용된 기계 학습 방법이라고 할 수 있겠다[5,7,12,13]. 기존 연구들 중 [12]는 수식 1의 첫 번째 확률값을 결정하기 위해 선행된 모든 발화의 화행을 참고하는 대신 현재 발화와 가장 근접하고 직접적으로 영향을 미치는 이전 화행만을 선택하여 참고한다. 수식 2는 현재 발화의 화행을 결정하기 위해 [12]에서 사용하는 확률 모델을 표현한 것이다.

$$S(U_n) \approx \underset{S_n}{\operatorname{argmax}} P(S_n | S_j, S_k) P(F_n | S_n)$$

수식 2 현재 발화 화행 결정 모델

여기서 S_n 은 현재 화행을, S_j 와 S_k 는 담화 구조(Discourse Structure)를 통해 추정된 현재 화행에 직접적으로 영향을 미치는 가장 근접한 2개의 화행을, F_n 은 현재 발화의 문장 자질을 의미한다. 또한 담화 구조를 화행 분석 없이 미리 파악한다는 것은 어려운 문제이므로 j 와 k 를 $n-2$ 와 $n-1$ 로 치환하여 사용할 수 있다.

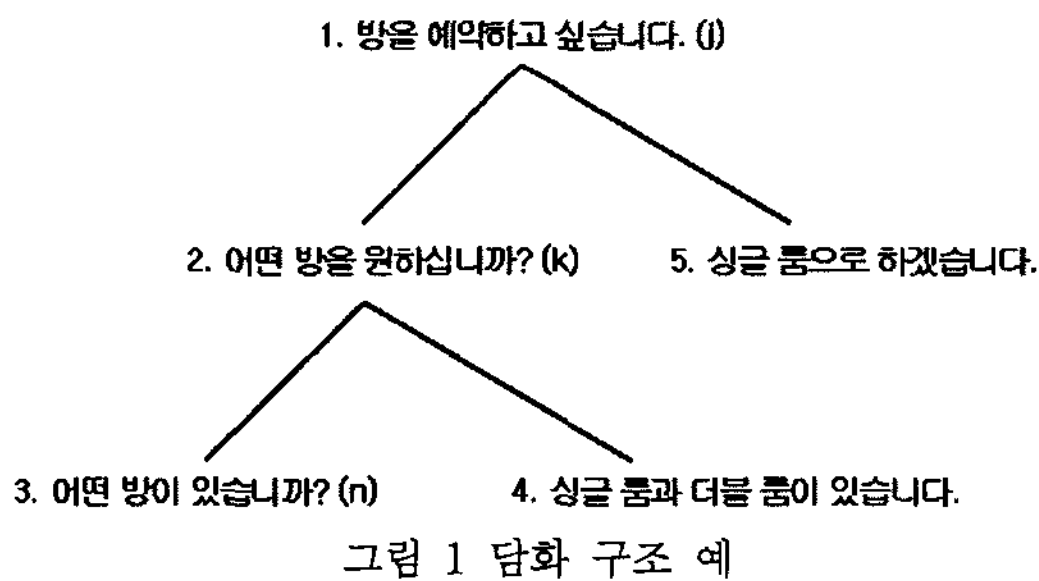


그림 1 담화 구조 예

[7]은 현재 발화의 화행을 결정하는데 유용한 정보를 문장 자질과 문맥 자질로 나누어서 이를 활용하였다. 문장 자질은 현재 문장에서 나타나는 정보를 통틀어서 말하는 것으로, [7]에서는 단어/품사, 품사 바이그램(bi-gram)을 문장 자질로 사용하였다. 문맥 자질은 대화의 흐름을 나타내는 정보로서, 이전 화행과 현재 화행의 쌍(pair)을 그 정보로 사용하기도 하였다.

통계 기반 모델을 구현하는데 있어서 기계 학습 방법을 연구하는 것 이외에도 학습에 사용되는 자질을 선택하는 방법에 대한 연구가 진행되었다. 대표적인 방법으로는 수식 3과 같은 카이제곱 통계량(Chi Statistics)을 사용하여 자질과 화행간의 정보량을 측정 한 후, 화행 결정에 큰 영향을 미치는 효과적인 자질을 선택하는 방법이 있다[7].

$$\chi^2(F,S) = \frac{(a+b+c+d) \times (ad-cb)^2}{(a+c) \times (b+d) \times (a+b) \times (c+d)}$$

- a : 자질 F를 가지고 화행 S로 부여된 발화 수
 - b : 자질 F를 가지고 화행 S로 부여되지 않은 발화 수
 - c : 자질 F를 가지지 않고 화행 S로 부여된 발화 수
 - d : 자질 F를 가지지 않고 화행 S로 부여되지 않은 발화 수
- 수식 3 카이제곱 통계량

2.3 자질의 종류와 한국어 화행

기존 연구에서 사용된 문장 자질 또는 구문 정보는 문장 형태, 본용언, 보조용언, 시제, 부정형 여부, 단서 단어 등이 있으며 문맥 자질로서는 이전 화행, 화자 정보 등이 있다[12,13].

한국어 화행의 종류는 설계자의 판단에 따라 다양하게 정의될 수 있지만 기존의 연구에서 대표적으로 사용된 화행은 표 2와 같다[7,12,13].

표 1 문장 자질

종류	예
문장 형태	평서문, 명령문, Wh의문문, YN의문문
본용언	'알'(동사), '돕'(동사), '어렵'(형용사) 등
보조용언	'드리', '주', '되', '하', '싶', '않' 등
시제	과거, 현재, 미래
부정형 여부	유, 무
단서 단어	'네', '아니요', '안녕히', '감사', '매진' 등

표 2 한국어 화행

화행	설명 (예)
Introducing-oneself	자신을 소개 (홍길동이라고 하는데요.)
Opening	대화를 시작 (무엇을 도와드릴까요?)
Closing	대화를 마칩 (감사합니다. / 수고하세요.)
Accept	승인 (네. / 알겠습니다.)
Acknowledge	상대방이 계속 말하도록 호응 (네. / 788에. / 그리고요?)
Ask-confirm	앞에서 발화된 내용을 확인 (홍길동이라고요?)
Ask-if	Ask-confirm을 제외한 YN의문문 (유나이티드 항공입니까?)
Ask-ref	Wh의문문 (언제 출발하시길 원하십니까?)
Request	특정 행동을 요구 (예약 좀 해주세요.)
Expressive	감정을 표현 (좋네요. / 기쁩니다.)
Correct	상대방의 발화에 대한 수정 (아니요, 그 항공편은 무착륙 항공편입니다.)
Reject	거절 (어렵겠습니다.)
Promise	특정 행동에 대한 약속 (본 내용으로 예약하겠습니다.)
Inform	독립적으로 정보 제공 (샌프란시스코를 경유합니다.)
Offer	부가 (확인해보겠습니다.)
Suggest	제안 (왕복편은 어떠십니까?)
Response	응답 (5월 2일이요. / 예.)

본 논문에서는 위와 같은 문장 자질과 문맥 자질, 그리고 17개의 화행을 사용하며, 특히 단서 단어에 대해서는 실험에 사용된 대화 코퍼스의 형태소 분석 결과로부터 추출된 단어/품사 정보를 수식 3의 카이제곱 통계량을 사용하여 순위화한 후, 해당 발화에 존재하는 최상위 단어/품사 정보를 선택하도록 하였다.

3. 제안하는 방법

3.1 이후 발화의 후보 화행 추정

현재 발화의 화행을 결정하기 위해서는 수식 1을 통해 알 수 있듯이 이전 화행과 현재 발화의 문장 자질을 알아야 한다. 하지만 대화가 화행의 흐름을 벗어나거나 약간의 변화를 일으키기만 해도 이전 화행으로부터 추정할 수 있는 현재 발화의 화행 정보는 불명확하게 된다. 그림 다음과 같은 상황을 생각해보자.

U: 사용자 / A: 에이전트

■ 예 1.
 U: "5월 2일에 출발하고 싶습니다." (Request)
 A: "(평서문 / 단서 단어 : '알'(동사))"

■ 예 2.
 U: "5월 2일에 출발하고 싶습니다." (Request)
 A: "(평서문 / 단서 단어 : '알'(동사))"
 U: "네." (Accept)

그림 2 상황 예시

예 1을 살펴보면 사용자가 5월 2일에 여행을 떠나고 싶으니 해당 정보를 찾아달라는 '요구'를 하고 있음을 알 수 있다. 그럼 에이전트는 사용자의 요구에 따라 해당 정보를 찾아보겠다는 '승인' 및 '부가'적인 발화를 할 것이다. 즉, 에이전트의 대답이 평서형이고 단서 단어가 동사인 '알'이라는 것을 알고 있으므로 해당 발화가 '승인'일 경우에는 "알겠습니다."와 같은 발화를, '부가'적인 표현일 경우에는 "해당 항공편을 알아보겠습니다."와 같은 발화를 했을 것으로 예상할 수 있다. 예 2는 예 1과 동일한 발화가 진행된 후에 사용자가 '승인'의 의미를 가진 발화를 하는 상황이다. 에이전트의 화행이 'Accept'일 때 이후 사용자의 화행이 'Accept'일 수는 없으므로 이러한 사실을 통해 에이전트의 화행이 'Accept'가 아닌 'Offer'라는 것을 알 수 있다. 이렇게 현재 발화 이후의 발화에 대한 화행 정보를 알고 있다면 화행을 결정하는데 있어서 본 정보가 유용하게 사용될 수 있다. 하지만 실제 대화에서 현재 발화 이후의 화

행을 정확하게 알 수 있는 방법은 없다. 다만 다음 화행으로서 추정되는 상위 후보 화행들을 활용하여 현재 화행을 결정하는데 도움을 줄 수 있을 것이다. 즉, 후보 화행들과 연결이 가장 자연스러운 화행이 현재 화행으로서 가능성이 높다는 것이다.

그림 3은 이전 화행, 현재 화행, 그리고 이후 화행간의 관계를 나타내는 예로서, 이전 화행이 A일 경우 현재 화행은 B1 또는 B2가 될 수 있고 현재 화행이 B1일 경우에는 이후 화행이 C1 또는 C2, 현재 화행이 B2일 경우에는 C1만이 이후 화행이 될 수 있음을 뜻한다. 여기서 C1, C2를 현재 발화 이후의 발화에 대한 추정된 화행이라고 할 때 현재 화행으로서 가장 적절할 것은 B1이 될 것이다. 또한 본 그림을 통해서 이후 화행을 정확하게 추정하는 것 이외에도 2개 이상의 화행을 추정하는 것이 현재 화행을 결정하는데 보다 도움이 됨을 알 수 있다.

그림 4는 후보 화행을 추정하는 방법과 추정된 후보 화행이 현재 화행에 영향을 주는 모습을 시각적으로 표

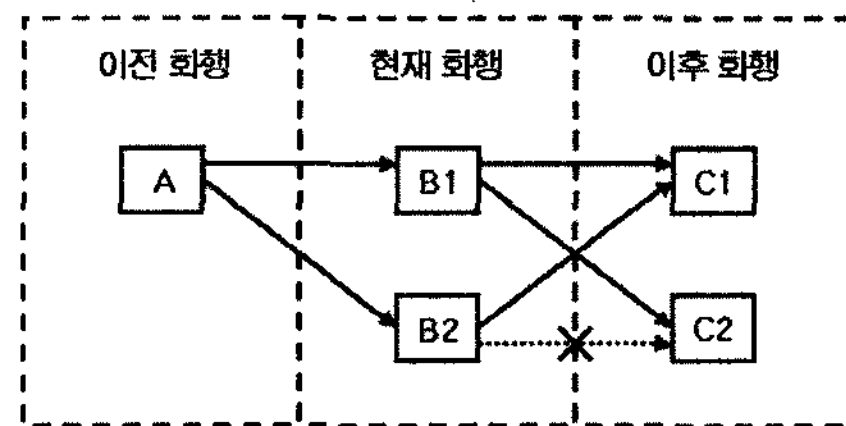


그림 3 화행간의 관계 예

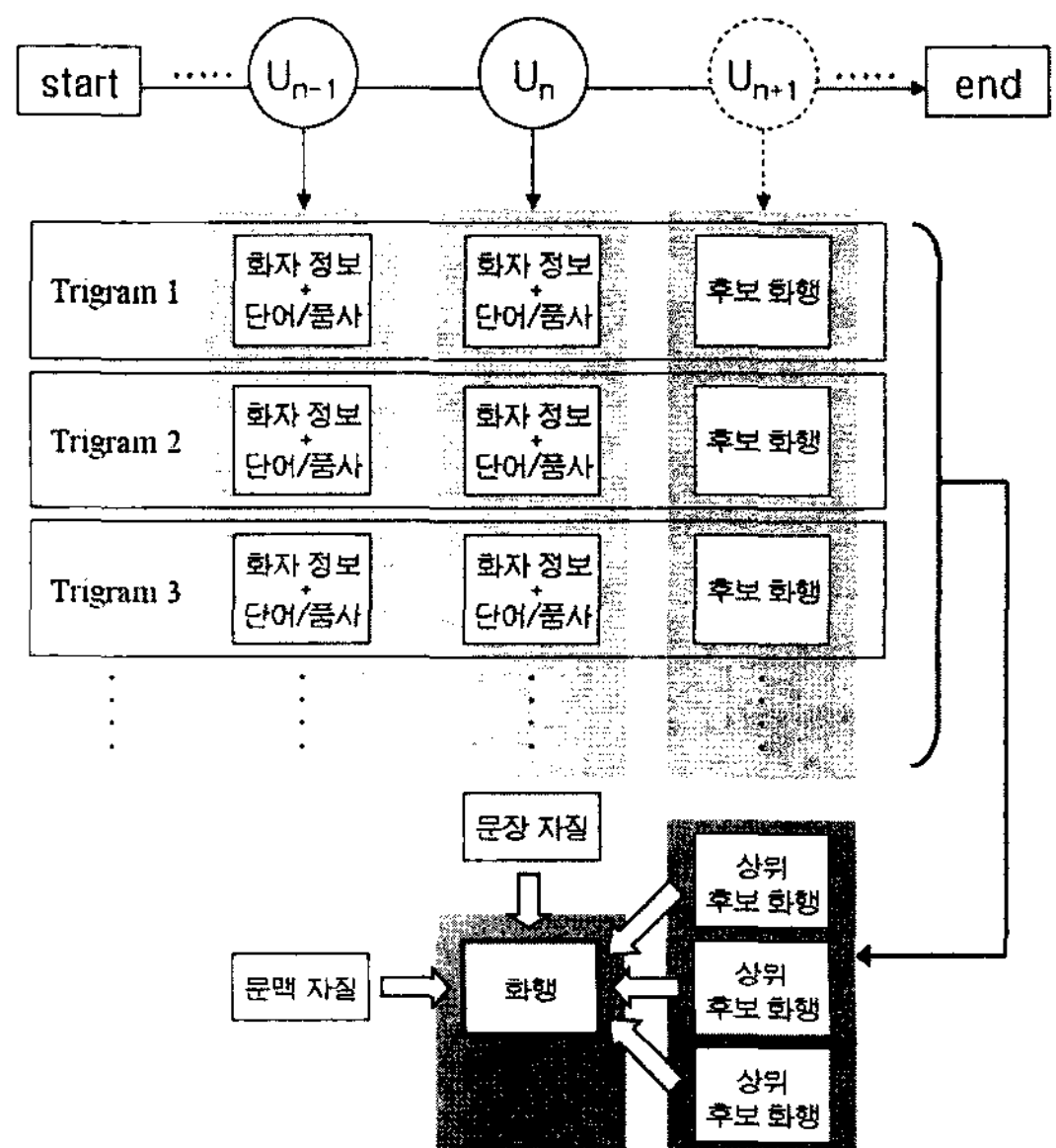


그림 4 후보 화행 추정

현한 것이다. 후보 화행을 구하기 위해서는 현재 발화 이전의 화행 정보를 사용하는 대신 이전 발화의 화자 정보와 단어/품사 정보를 사용한다. 이는 이전 발화의 화행 결정에 독립적인 후보 화행 결정 모델을 구현하기 위함이다. 후보 화행을 추정하는 과정을 자세히 설명하면, 먼저 대화 코퍼스의 형태소 분석 결과와 화자 정보를 추출하여 '화자 변경 정보'를 포함한 '이전 발화의 단어/품사' + '현재 발화의 단어/품사' + '다음 화행'으로 이루어진 트라이그램(trigram) 데이터를 만든다. 그리고 각각의 트라이그램 데이터에 대한 카이제곱 통계량을 구하여 순위화한 후, 구하고자 하는 화행의 이전 발화와 현재 발화의 단어/품사 정보와 일치하는 트라이그램 데이터를 통계량이 큰 것부터 선택하여 해당 데이터의 마지막 항목인 다음 화행 정보를 후보 화행으로서 결정한다. 표 3은 대화 코퍼스로부터 추출된 트라이그램 데이터와 카이제곱 통계량의 일부를 기록한 것이다.

표 3 트라이그램 데이터와 카이제곱 통계량

순위	트라이그램 데이터	화자 변경 유무	카이제곱 통계량
1	blank(blank)/익스프레스(CMCN)/Opening	유	4119.68
9	수고(CMCPA)/.(g)/end	유	3292.06
11	감사(CMCPA)/.(g)/Closing	유	3222.80
12	감사(CMCPA)/하(fpd)/Closing	유	2864.63
77	익스프레스(CMCN)/무엇(CTP3)/Introducing-oneself	무	1348.56
164	으로(fjcao)/알(YBDO)/Promise	유	389.63
262	이(fpd)/비싸(YBHO)/Suggest	유	322.77
708	.(g)/버니까(fmofq)/Response	유	212.63
962	예약(CMCPA)/알(YBDO)/Promise	유	159.50
1390	안녕(CMCPS)/러고(fmocs)/Ask-ref	무	122.49

위 표에서 등장하는 'blank(blank)'는 이전 발화가 존재하지 않을 경우 임시로 사용하는 단어/품사 정보이며 'end'는 다음 발화가 없을 경우에 사용하는 임시 화행 정보이다. 후보 화행은 이전 발화와 현재 발화의 단어/품사 정보에 따라 중복되어 선택될 수 있으며 현재 문장의 문장 자질이 충분한 정보를 제공해주지 못할 경우 높은 효과를 발휘한다. 물론 현재 발화의 문장 자질이 화행을 결정하기 위한 충분한 정보를 제공할 경우에는 오히려 화행 결정에 방해가 될 수 있다.

3.2 이전 발화의 문장 자질 활용

선행된 발화의 화행 정보는 현재 발화의 화행을 결정하는데 큰 영향을 미친다. 기존 연구에서 사용한 수식 1과 이를 변경한 수식 2의 내용을 살펴보아도 현재 발화의 화행을 결정하기 위한 이전 화행의 유용성을 쉽게 알

수 있다. 하지만 실제 대화에서는 매우 다양한 대화의 흐름과 표현 방법들에 의해서 발화가 생성되기 때문에 각 발화의 화행을 결정한 이후에는 해당 발화의 이러한 다양성을 나타낼 수 있는 정보를 소실하게 된다. 결국 이전 발화의 정보로서 화행만을 사용할 경우, 일반적인 화행의 흐름을 벗어나거나 화행 이외의 정보들로 이루어진 대화 패턴에 영향을 받는 대화들에 대해서는 보다 효과적인 결과를 가져오지 못한다. 다시 말하면, 이전 발화의 화행뿐만 아니라 그 발화의 또 다른 특성들에 의해서 현재 발화가 영향을 받을 수 있다는 것이다.

이전 발화에 대한 정보를 화행으로만 한정할 경우 나타날 수 있는 문제점들을 보완하기 위해, 본 논문에서는 이전 발화의 문장 자질을 현재 발화의 화행 결정에 함께 사용함으로써 발화의 다양성을 반영한 화행 결정 모델을 구현하였다. 이전 화행을 결정하는데 사용했던 문장 자질은 잠재적으로 해당 발화의 후보 화행 정보를 포함하고 있으며 발화 자체에 대한 특성을 가지고 있기 때문에 실제 화행 정보와 함께 유용하게 사용할 수 있다. 수식 4는 이와 같은 사실을 바탕으로 수식 2를 변경한 것이다.

$$S(U_n) \approx \underset{S_n}{\operatorname{argmax}} P(S_n | S_{n-2}, F_{n-2}, S_{n-1}, F_{n-1}) P(F_n | S_n)$$

수식 4 이전 발화의 문장 자질을 반영한 화행 결정 모델

앞에서도 언급했듯이 문장 자질은 기존 연구에서 사용된 문장 형태, 본용언, 보조용언, 시제, 부정형 여부, 단서 단어를 사용하였고, 특히 단서 단어에 대해서는 대화 코퍼스의 형태소 분석 결과에서 추출된 단어/품사 정보와 이에 대응하는 화행 정보로 이루어진 바이그램을 활용하였다. 즉, 표 4와 같이 바이그램 결과들을 카이제곱 통계량을 통해 순위화한 후, 분석하고자 하는 발화에 존재하는 최상위 단어/품사 정보를 찾아 이에 해당하는 화행 정보를 화행 결정에 참고하도록 하였다. 이밖

표 4 단서 단어와 카이제곱 통계량

순위	단서 단어와 화행	카이제곱 통계량
1	무엇(CTP3)/Opening	1162.06
2	줬(fmbtf)/Accept	1118.73
3	안녕히(SBO)/Closing	1094.88
5	수고(CMCPA)/Closing	875.67
15	잠시(CMCN)/Offer	388.60
17	매진(CMCPA)/Reject	382.53
43	어디(CTP3)/Ask-ref	187.43
76	부탁(CMCPA)/Request	106.40
147	습니까(fmofq)/Ask-if	50.76
370	그리고(SBJ)/Inform	13.41

에도 발화간의 화자 변경 정보를 문장 자질과 함께 사용하여 현재 화행을 결정하는데 도움을 주었다.

그림 5는 본 방법론을 시각적으로 표현한 것으로, 화행 결정에 'Request' 화행만을 사용하는 것보다 '알'이라는 본용언을 가진 평서문 'Request' 화행을 사용하는 것이 더 효과적임을 보이고 있다.

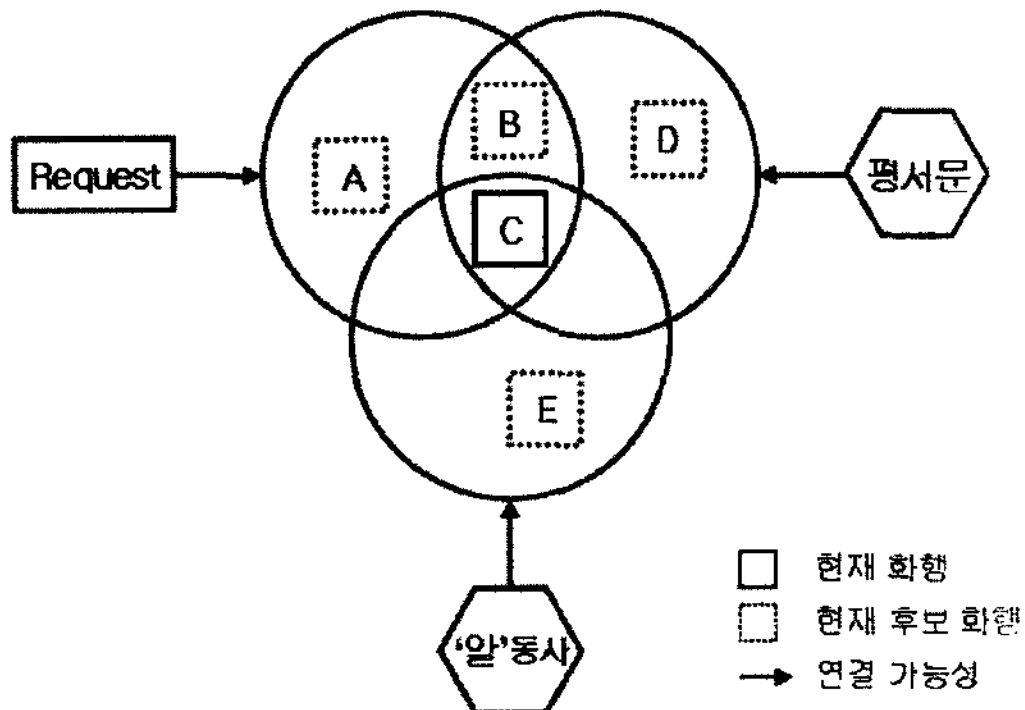


그림 5 문장 자질 활용

본 방법론은 이전 발화의 화행 정보로부터 얻을 수 없는 희박한 화행 흐름과 대화 패턴 정보를 제공해주거나 이전 화행 정보를 강화해주는 역할을 수행한다. 그러나 이전 화행이 현재 화행을 결정하는데 충분한 정보를 제공할 경우, 또는 유용한 문장 자질 자체가 존재하지 않을 경우에는 기존의 방법과 동일한 성능을 보이거나 오히려 성능 감소를 초래할 수 있다.

3.3 화행별 순차 처리(3 Stages Processing)

앞에서 소개한 두 가지 방법론은 화행 결정에 있어서 효과적인 정보들을 제공해줄 수 있지만 몇 가지 한계점을 가지고 있다. 각 절의 마지막 부분에서 언급했듯이 이후 발화에 대한 후보 화행을 추정하여 활용하는 것은 현재 발화의 문장 자질이 화행을 결정하기 위한 충분한 정보를 제공할 경우 오히려 화행 결정에 방해가 될 수 있으며, 이전 발화의 문장 자질을 활용하는 것은 현재 화행이 이전 화행에 전적인 영향을 받거나 해당 문장 자질이 존재하지 않을 경우 유용하지 못하다. 17개의 한국어 화행 중, 위와 같은 문제점을 가지고 있는 화행은 직관적으로 볼 때 'Offer', 'Suggest', 'Introducing-oneself', 'Opening'이 있다. 'Offer'는 "확인해보겠습니다.", "잠시만 기다려주세요."와 같은 거의 고정적인 표현을 가지고 있으며 대화의 어느 위치에서든 등장할 수 있기 때문에 이웃 화행과 같은 문맥 자질보다는 현재 문장 자질이 이를 판단하는데 더 유용하다. 'Suggest'는 일반적으로 'Ask-if' 이후에 나타나는 발화와 거의 동일한 이후 발화를 가지기 때문에 앞에서 제시한 화행 추정

방법을 적용할 경우 올바른 화행 분석을 수행할 수가 없다. 실제 대화에서도 'Suggest'의 빈도수보다 'Ask-if'의 빈도수가 매우 많으므로 본 화행을 결정하는데 있어서 이후 화행의 추정은 방해가 된다. 'Introducing-oneself'과 'Opening'은 대화가 시작되는 부분에서 등장하기 때문에 이전 발화의 정보 자체가 없거나 희박하다. 그러므로 이와 같은 화행을 결정하는 단계에서는 이전 문장 자질을 적용하지 않는 것이 바람직하다.

본 논문에서는 이러한 직관을 바탕으로 화행의 종류에 따라 그에 알맞은 확률 모델을 단계별로 적용하는 화행별 순차 처리 방법을 제시한다. 먼저 기존 연구의 방법론을 모델 A, 모델 A에 이후 발화의 후보 화행을 적용한 것을 모델 B, 그리고 모델 B에 이전 발화의 문장 자질을 적용한 것을 모델 C라고 하자. 각 모델에 대해서 화행 결정을 위한 학습을 수행한 후, 입력된 발화에 대해서 모델 A를 적용한다. 모델 A를 통해서 'Offer'나 'Suggest'로 판단된 발화는 이후 처리를 생략하고 그 외의 화행으로 판단된 발화에 대해서는 모델 B를 적용한다. 모델 B를 통해서 'Introducing-oneself'나 'Opening'으로 판단된 발화는 역시 이후 처리를 생략하고 나머지 화행으로 판단된 발화는 최종적으로 모델 C를 적용한다. 그림 6은 본 순차 처리 단계를 순서도(flow-chart)로 표현한 것이다.

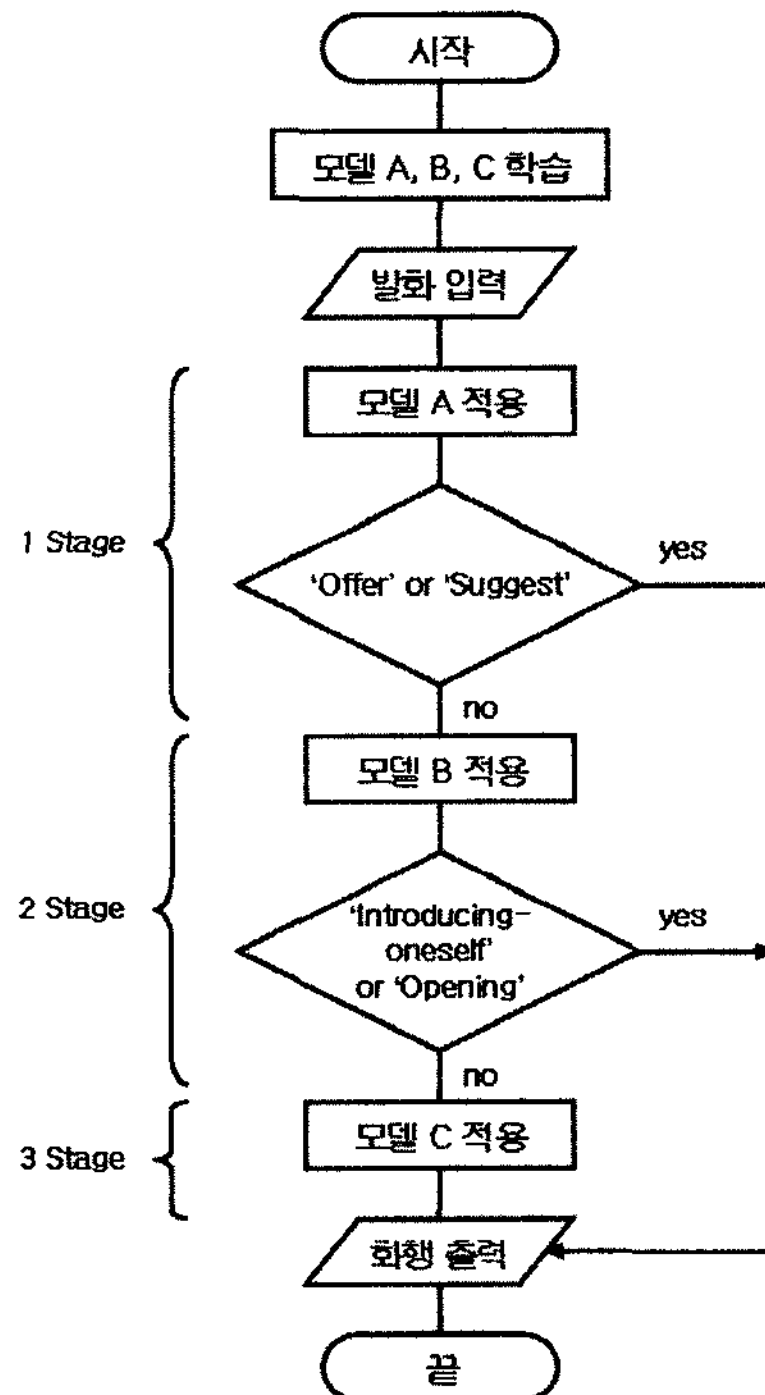


그림 6 화행별 순차 처리

4. 실험 결과 및 분석

본 논문에서는 실제 사용자와 항공 예약 전문가 간의 대화 내용을 기록한 SRI ATIS Collection[9]의 일부를 번역 및 정제한 코퍼스를 사용한다. SRI ATIS Collection은 '컴퓨터와 사람간의 대화'와 '사람과 사람간의 대화'를 비교 분석하기 위해 활용되기도 하였던 대화 스크립트(script)이다[4]. 이렇게 정제된 코퍼스는 총 1479개의 발화와 18개의 대화로 구성되어 있으며 각 발화에 대한 문장 자질과 형태소 분석 결과, 그리고 정답 화행이 부여되어있다. 형태소 분석 결과는 포항공과대학교 지식 및 언어공학 연구실에서 개발한 형태소 분석기(KoMA)를 사용하였다. 실험 데이터의 세부 정보는 표 5와 같다.

실험은 앞에서 소개한 각각의 자질 정보가 화행 결정에 어떠한 영향을 미치는지 확인하고 기존의 방법과 제안한 방법에 대한 비교 성능 평가를 수행하기 위해 그림 7과 같은 순서로 진행하였다. 그림 7에서 S_i 는 i 번째 발화의 화행을, F_i 는 i 번째 발화의 문장 자질을, n 은 현재 발화 위치를 의미한다. 본 실험은 문장 자질, 문맥 자질, 후보 화행, 정답 화행 정보를 대화 코퍼스로부터 추출하고 이를 데이터화하여 입력 자질이 많은 경우에도 안정적인 성능을 보이는 SVM(Support Vector Machine)

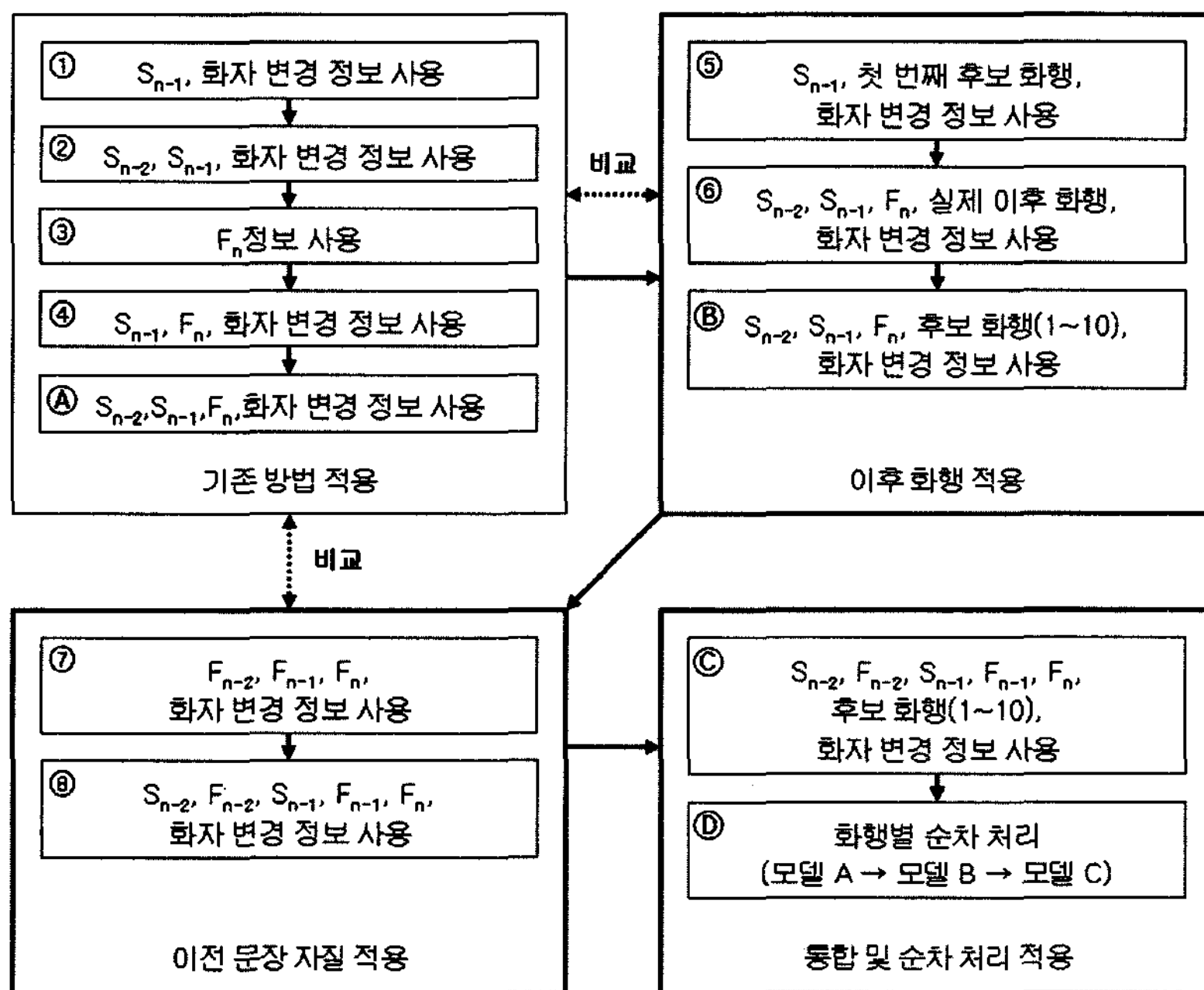
표 5 실험 데이터

화행	개수	화행	개수	화행	개수
Introducing.	33	Ask-if	119	Promise	18
Opening	27	Ask-ref	132	Inform	232
Closing	36	Request	87	Offer	31
Accept	181	Expressive	55	Suggest	21
Acknowledge	102	Correct	3	Response	338
Ask-confirm	49	Reject	15		

기법을 적용하였다. 실제 실험은 자바(Java)기반 기계학습 알고리즘 툴인 웨카(Weka)를 활용하여 학습 및 성능 평가를 수행하였다[14]. 그리고 실험 데이터들의 수가 비교적 적은 편이므로 학습 데이터와 평가 데이터를 정적으로 분리하지 않고 총 실험 데이터의 10%를 평가 데이터로 하고 남은 90%를 학습 데이터로 하여 성능 평가를 수행한 후, 같은 방식으로 평가 데이터와 학습 데이터를 순차적으로 갱신하여 전체 성능 평가를 수행하는 10-fold cross validation 방법을 사용하였다.

4.1 기존 방법 적용 결과

화행 분석에 대한 기존 연구들의 결과들은 표 6과 같다. 비교란의 내용을 통해서 알 수 있듯이, 각 연구별로 자신만의 화행 셋(set)을 사용하고 있으며 그 개수 또한 다양하다. 학습에 사용된 코퍼스도 각기 다르며 언어의



* ○ : 모델 번호 * □ : 제안한 방법 적용

그림 7 실험 순서

표 6 기존 연구 결과

기존 연구	성능	비고
Reithinger, et al., 1997 [15]	74.7%	VERBMOBIL, 18개 화행 사용
Samuel, et al., 1998 [11]	75.12%	VERBMOBIL, 18개 화행 사용
Fukada, et al., 1998 [16]	81.2%	C-Star, 26개 화행 사용
Stolcke, et al., 2000 [5]	65.0%	DAMSL, 42개 화행 사용, 음성(Speech)
Lee, et al., 2002 [13]	81.5%	17개 화행 사용
Kim, et al., 2003 [7]	88.3%	17개 화행 사용
Webb, et al., 2005 [17]	71.98%	21개 화행 사용
Ivanovic, 2005 [6]	81.6%	12개 화행 사용
Choi, et al., 2005 [12]	83.57%	17개 화행 사용

종류 및 발화의 인식 방법(문자/음성)에도 차이가 있다. 이러한 사실들은 본 논문의 실험 결과와 기존 연구의 실험 결과를 직접 비교하는데 있어서 한계점을 가져온다.

하지만 기존에 연구된 모든 방법론들은 화행을 결정하기 위해 이전 화행과 현재 발화에서 추출할 수 있는 정보만을 사용하였다. 그러므로 본 논문에서는 이전 화행과 현재 발화의 문장 자질을 사용하여 평가한 실험 결과를 기존 연구에 대한 성능으로써 가정하였다.

앞에서 소개한 실험 순서에 따라 '이전 화행 1개(S_{n-1}), 화자 변경 정보'를 사용한 것을 모델 1, '이전 화행 2개(S_{n-1}, S_{n-2}), 화자 변경 정보'를 사용한 것을 모델 2, '현재 문장 자질(F_n)'만을 사용한 것을 모델 3, '이전 화행 1개(S_{n-1}), 현재 문장 자질(F_n), 화자 변경 정보'를 사용한 것을 모델 4, '이전 화행 2개(S_{n-1}, S_{n-2}), 현재 문장 자질(F_n), 화자 변경 정보'를 사용한 것을 모델 A라고 하자. 여기서 모델 A는 본 논문에서 제안하는 방법론과 비교하기 위한 기존 연구 결과이다.

표 7의 실험 결과를 살펴보면, 모델 1은 1개의 이전 화행만 사용하여 현재 화행을 결정하였기 때문에 'Introducing-onself'와 'Response'와 같은 대화의 흐름에 의존적인 화행에 대해서만 좋은 성능을 보였다. 모델 2는 2개의 이전 화행을 사용하여 현재 발화의 화행을 결정하였기 때문에 모델1보다 전반적으로 높은 성능을 보였다. 하지만 대화의 흐름에 의존적인 'Opening'과 'Closing'만이 크게 올랐을 뿐 'Correct', 'Reject', 'Promise', 'Offer'와 같은 문장 자질에도 크게 영향을 받는 화행에 대해서는 효과를 보지 못했다. 모델 3은 현재 문장 자질만을 가지고 성능을 평가한 것으로 이전 화행에 의존적인 'Acknowledge', 'Response'와 같은 화행을 제외한 나머지 화행에 대해서 비교적 좋은 성능을 보였다. 모델 4는 1개의 이전 화행과 현재 문장 자질을 가지고 성능을 평가하였는데 모델 3보다 매우 높은 성능을 보였으며, 2개의 이전 화행을 현재 문장 자질과 함께 사용한

표 7 기존 방법 적용 성능 평가 결과

화행	F-Measure (%)				
	모델1	모델2	모델3	모델4	모델A
Introducing.	83.1	82.7	91.7	95.7	95.7
Opening	68.9	87.5	87.5	94.1	94.1
Closing	55.6	98.6	95.7	100.0	100.0
Accept	45.0	47.3	83.0	93.4	94.3
Acknowledge	38.8	37.7	14.3	87.0	90.7
Ask-confirm	11.5	21.1	28.1	45.2	61.9
Ask-if	0.0	17.3	82.8	78.7	86.7
Ask-ref	1.4	29.9	95.0	94.2	96.4
Request	0.0	50.3	79.3	92.2	98.9
Expressive	19.0	50.5	74.8	81.9	84.5
Correct	0.0	0.0	75.0	85.7	100.0
Reject	0.0	0.0	74.1	88.9	92.9
Promise	0.0	0.0	10.0	27.3	25.0
Inform	56.2	58.8	53.1	85.8	89.6
Offer	0.0	0.0	90.3	91.5	95.1
Suggest	0.0	25.0	83.3	87.8	92.3
Response	91.2	91.6	65.9	94.2	95.7
정확도(accuracy)	50.57	55.78	71.06	88.24	91.62

모델 A 또한 보다 높은 성능 향상을 보였다. 이러한 성능 향상은 통계 기반 모델을 구현하기 위해 갖추어야 할 기본적인 정보들을 모두 사용하였기 때문임과 동시에, 실험 데이터 자체의 특성 때문이라고 할 수 있겠다. 즉, 본 실험에서 사용된 코퍼스 자체가 항공 예약이라는 특정 도메인에 한정되어있고 비교적 적은 양의 발화로 인해 대화의 흐름 및 각 발화에서 사용된 문장 자질의 양이 한정되어 있어서 이와 같은 높은 성능이 나타난 것으로 판단된다.

4.2 이후 화행 적용 결과

'이전 화행 1개(S_{n-1}), 첫 번째 후보 화행, 화자 변경 정보'를 사용한 것을 모델 5, '이전 화행 2개(S_{n-1}, S_{n-2}), 현재 문장 자질(F_n), 실제 이후 화행, 화자 변경 정보'를 사용한 것을 모델 6, '이전 화행 2개(S_{n-1}, S_{n-2}), 현재 문장 자질(F_n), 1개 이상의 후보 화행, 화자 변경 정보'를 사용한 것을 모델 B라고 하자.

모델 5는 본 논문의 3장에서 제안했던 다음 발화에 대한 후보 화행들을 추정하여 그 중 최상위에 위치한 후보 화행을 선택하여 실험에 반영한 것이다. 본 모델은 모델 2보다 향상된 성능을 보였고 특히 기존 모델에서 효과를 보지 못했던 화행인 'Correct', 'Promise', 'Offer'의 성능이 향상되었다. 하지만 'Opening'과 'Closing' 같은 대화의 흐름에 보다 의존적인 화행에 대해서는 오히려 성능이 떨어지는 결과를 보였다. 모델 6은 기존 연구 결과로서의 모델 A에 실제 이후 화행을 적용한 것으로, 이후 화행을 정확하게 추정할 수 있다면 현재 화행을

표 8 모델 2·5·A·6 성능 평가 결과

화행	F-Measure (%)			
	모델2	모델5	모델A	모델6
Introducing.	82.7	88.0	95.7	100.0
Opening	87.5	76.2	94.1	100.0
Closing	98.6	57.1	100.0	100.0
Accept	47.3	60.9	94.3	96.2
Acknowledge	37.7	44.7	90.7	96.7
Ask-confirm	21.1	11.5	61.9	60.6
Ask-if	17.3	46.6	86.7	85.7
Ask-ref	29.9	56.8	96.4	98.9
Request	50.3	11.2	98.9	98.3
Expressive	50.5	22.0	84.5	95.3
Correct	0.0	100.0	100.0	100.0
Reject	0.0	0.0	92.9	92.9
Promise	0.0	42.6	25.0	85.7
Inform	58.8	72.3	89.6	94.3
Offer	0.0	16.7	95.1	95.1
Suggest	25.0	0.0	92.3	92.3
Response	91.6	91.0	95.7	95.6
정확도(accuracy)	55.78	61.87	91.62	94.12

표 9 후보 화행 수별 성능 평가 결과

추가한 상위 후보 화행 수	정확도 (%)	추가한 상위 후보 화행 수	정확도 (%)
1	92.49	6	96.08
2	93.17	7	95.94
3	95.33	8	96.08
4	95.27	9	95.74
5	95.61	10	95.88

결정하는데 도움이 된다는 사실을 해당 실험 결과를 통해 알 수 있었다. 그럼 실제 이후 화행이 아닌 3장에서 제안한 방법을 통하여 추정된 화행들을 상위 순위부터 하나씩 추가해가면서 성능을 평가해보자.

표 9의 실험 결과를 살펴보면 현재 발화의 화행을 결정하기 위해 후보 화행을 추가할 경우 서로 중복될 수 있는 6개 또는 8개의 후보 화행을 선택하였을 때 가장 높은 성능을 보였다. 즉, 1개나 2개의 후보 화행을 선택하기보다는 좀 더 많은 후보 화행을 선택하는 것이 성능 향상에 도움이 되며, 특정 양을 초과하여 후보 화행을 선택할 경우 성능에 부정적인 영향을 미친다는 것을 알 수 있었다. 또한 1개의 정확한 이후 화행만을 현재 화행 결정에 사용하기보다는 추정된 여러 후보 화행을 적용하는 것이 효과적임을 알 수 있었다. 하지만 위 실험은 held-out 데이터를 사용하여 평가한 것이 아니기 때문에 이후 실험에서 사용할 후보 화행 수를 결정하는 데이터로서는 활용할 수 없다. 그러므로 본 논문에서는 전체 실험 데이터의 70%를 학습 데이터로, 20%를 held-

표 10 Held-out 데이터를 사용한 성능 평가 결과

추가한 상위 후보 화행 수	정확도 (%)	추가한 상위 후보 화행 수	정확도 (%)
1	94.26	6	94.93
2	94.26	7	95.27
3	95.61	8	95.27
4	95.95	9	95.27
5	94.59	10	95.27

표 11 모델 A·B간의 성능 평가 결과 비교

화행	F-Measure (%)		성능 향상
	모델A	모델B	
Introducing.	95.7	95.7	0.0
Opening	94.1	94.1	0.0
Closing	100.0	100.0	0.0
Accept	94.3	97.5	3.2
Acknowledge	90.7	94.8	4.1
Ask-confirm	61.9	77.4	15.5
Ask-if	86.7	89.9	3.2
Ask-ref	96.4	100.0	3.6
Request	98.9	96.6	-2.3
Expressive	84.5	100.0	15.5
Correct	100.0	100.0	0.0
Reject	92.9	100.0	7.1
Promise	25.0	100.0	75.0
Inform	89.6	94.4	4.8
Offer	95.1	93.5	-1.6
Suggest	92.3	89.5	-2.8
Response	95.7	95.8	0.1
정확도	91.62	95.27	3.65

out 데이터로, 10%를 평가 데이터로 설정하여 재실험하였다.

Held-out 데이터를 사용한 결과, 상위 후보 화행 4개를 사용했을 경우에 최적의 성능을 보였다. 실제로 평가 데이터에 대하여 후보 화행을 사용하지 않았을 경우 93.92%의 성능을 보인 반면, 상위 후보 화행을 4개 사용했을 경우에는 97.97%의 성능을 보였다. 그러나 본 평가 데이터는 그 양이 매우 적은 편이므로 앞에서 소개한 10-fold cross validation 방법을 그대로 적용하되 적용할 상위 후보 화행 수는 4개로 설정하고 이후 실험을 수행하도록 하였다.

표 11은 기존의 화행 결정 방법(모델A)과 제안한 방법(모델 B)의 성능을 비교한 것으로 전체 정확도는 3.65% 향상되었고 몇 가지 화행을 제외한 나머지 화행에 대해서도 전반적인 성능 향상을 보였다. 특히 기존의 방법은 'Promise'를 'Inform'으로, 'Ask-confirm'을 'Ask-if'로 잘못 판단하는 경우가 대부분이었으나 제안한 방법을 통해서 이를 구분하기 위한 단서를 보다 많이 제공

함으로써 각 화행의 성능이 서로 향상되는 결과를 가져올 수 있었다. 한편, 'Request', 'Offer', 'Suggest'는 오히려 성능 감소를 보였는데 'Request'를 'Offer'와 'Response'로, 'Suggest'를 'Inform'로 잘못 판단하여 발생한 결과로서, 추정된 이후 화행보다 현재 문장 자질이 해당 화행들을 결정하는데 보다 효과적임을 보여주는 예라고 할 수 있겠다.

4.3 이전 문장 자질 적용 결과

'이전 문장 자질 집합 2개(F_{n-1} , F_{n-2}), 현재 문장 자질 (F_n), 화자 변경 정보'를 사용한 것을 모델 7, '이전 화행 2개(S_{n-1} , S_{n-2}), 이전 문장 자질 집합 2개(F_{n-1} , F_{n-2}), 현재 문장 자질(F_n), 화자 변경 정보'를 사용한 것을 모델 8이라고 하자.

모델 7은 실제 이전 화행을 적용했을 때(모델 A)보다 향상된 성능을 가져왔으며 이는 이전 발화의 문장 자질이 이전 화행보다 유용한 정보로 활용될 수 있는 가능성을 보여준다. 물론 이전 화행에 의존적인 'Introducing-oneseft', 'Opening', 'Acknowledge'와 같은 화행에 대해서는 성능이 감소되는 결과를 보였다. 모델 A에 제안한 방법론, 즉 이전 문장 자질을 적용한 모델 8은 모델 A에 비해 전반적으로 높은 성능을 보였으며 이와 같은 결과는 본 코퍼스에 화행 이외의 정보로 이루어진 대화 패턴이 많이 존재한다는 것을 의미한다.

표 12 이전 문장 자질 적용 성능 평가 결과

화행	F-Measure (%)		
	모델A	모델7	모델8
Introducing.	95.7	92.5	92.5
Opening	94.1	90.6	90.6
Closing	100.0	100.0	100.0
Accept	94.3	93.5	95.3
Acknowledge	90.7	79.6	92.3
Ask-confirm	61.9	85.1	86.9
Ask-if	86.7	93.5	96.3
Ask-ref	96.4	96.6	97.0
Request	98.9	95.4	98.3
Expressive	84.5	90.2	90.7
Correct	100.0	100.0	100.0
Reject	92.9	100.0	96.6
Promise	25.0	94.1	94.1
Inform	89.6	88.9	94.8
Offer	95.1	91.5	91.5
Suggest	92.3	92.3	92.3
Response	95.7	95.8	97.5
정확도(accuracy)	91.62	92.49	95.33

4.4 통합 및 순차 처리 적용 결과

기존의 방법론(모델 A)에 본 논문에서 제안한 두 가지 방법론(이후 후보 화행, 이전 문장 자질)을 통합하여

표 13 통합 및 순차 처리 적용 성능 평가 결과

화행	F-Measure (%)			
	모델A	모델B	모델C	모델D
Introducing.	95.7	95.7	92.5	95.7
Opening	94.1	94.1	88.9	94.1
Closing	100.0	100.0	100.0	100.0
Accept	94.3	97.5	99.4	99.4
Acknowledge	90.7	94.8	97.6	97.6
Ask-confirm	61.9	77.4	90.2	90.2
Ask-if	86.7	89.9	96.2	96.2
Ask-ref	96.4	100.0	100.0	100.0
Request	98.9	96.6	99.4	99.4
Expressive	84.5	100.0	98.2	98.2
Correct	100.0	100.0	100.0	100.0
Reject	92.9	100.0	96.6	96.6
Promise	25.0	100.0	100.0	100.0
Inform	89.6	94.4	97.2	97.2
Offer	95.1	93.5	93.3	95.1
Suggest	92.3	89.5	89.5	92.3
Response	95.7	95.8	99.1	99.1
정확도(accuracy)	91.62	95.27	97.70	97.97

적용한 것을 모델 C, 화행별로 모델 A, 모델 B, 모델 C를 순차적으로 적용한 것을 모델 D라고 하자.

본 실험 결과를 통해 알 수 있듯이, 현재 발화에 대한 화행 결정은 제안한 방법론을 단독으로 적용했을 때보다 두 가지 방법론을 함께 적용했을 경우, 보다 높은 성능을 보였으며 현재 문장 자질에 영향을 많이 받거나 이전 문장 자질 정보가 존재하지 않는 몇몇 화행에 대해서는 성능 하락을 보였다. 모델 D는 이러한 문제점을 보완한 순차 처리 방법을 적용한 것이다. 제안한 방법을 통해 얻어진 최종 성능 평가 결과는 기존의 실험 결과보다 6.35% 향상된 97.97%의 정확도를 보였다.

5. 결론

대화 에이전트를 구현하기 위해서는 자연언어처리 기술에 바탕을 둔 다양한 연구 분야들의 적용이 필요하며, 각각의 분야들은 자신의 성능을 최대한 발휘함으로써 성공적인 대화 에이전트를 구현할 수 있다. 본 논문은 대화 처리 분야에서 중요한 위치를 차지하고 있는 화행 분석에 대하여 다루었으며, 현재 발화 이후에 나타날 수 있는 후보 화행들을 추정하고, 이전 발화의 문장 자질을 이전 화행과 함께 사용하며, 화행의 종류에 따라 그에 알맞은 확률 모델을 단계별로 적용함으로써 기존의 연구 방법과 차별화를 둠과 동시에 성능 향상을 꾀했다. 실제로, 제안한 방법을 통해 얻어진 결과는 기존의 결과보다 6.35% 향상된 97.97%의 성능을 보였으며 이는 현재 화행 결정에 대한 이후 화행의 유용성과 다양한 대

화 패턴을 나타낼 수 있는 문장 자질의 명확성에 의한 것으로 판단된다. 또한 본 방법론은 언어에 의존적이지 않기 때문에 다른 언어의 화행 분석에도 활용될 수 있을 것이다.

하지만 앞의 실험 결과에 대한 분석에서 언급했듯이 실험에 사용된 코퍼스가 특정 도메인, 즉 항공 예약 도메인에 한정되어있고 비교적 적은 양의 발화로 이루어져있기 때문에 대화의 흐름이 강한 방향성을 가지게 되었다. 발화를 이루고 있는 형태소 및 화행 결정에 큰 영향을 미치는 단서 단어도 화자의 다양한 발화들을 포괄할 수 있을 만큼 다양하지 못해서 해당 실험에 대한 기본 성능이 높은 것으로 여겨진다. 다시 말하면, 한정된 자원으로 인한 대용량 대화 코퍼스에서의 성능 평가를 수행하지 못해 제시한 방법론에 대한 보편적인 효용성을 보일 수 없었다. 향후에는 대용량 대화 코퍼스를 사용하여 제시한 방법론의 유용성을 검증하고 이후 화행을 보다 정확하게 추정할 수 있는 방법론이 모색되어야 할 것이다. 그리고 입력된 발화로부터 추출할 수 있는 정보들을 보다 유용하게 활용할 수 있도록 문장 자질의 개선도 수행되어야 할 것이다.

참 고 문 헌

[1] J. Allen, et al., "The TRAINS Project: A case study in building a conversational planning agent," *Journal of Experimental & Theoretical Artificial Intelligence*, Vol.7, No.1, pp. 7-48, 1995.

[2] L. Lambert and S. Carberry, "A tripartite plan-based model of dialogue," *Proceedings of the 29th annual meeting on Association for Computational Linguistics*, pp. 47-54, 1991.

[3] J. Chu-Carroll and S. Carberry, "Response generation in collaborative negotiation," *Proceedings of the 33rd annual meeting on Association for Computational Linguistics*, pp. 136-143, 1995.

[4] C. Doran, et al., "Comparing Several Aspects of Human-Computer and Human-Human Dialogues," *Current and New Directions in Discourse and Dialogue*, pp. 133-159, 2003.

[5] A. Stolcke, et al., "Dialogue act modeling for automatic tagging and recognition of conversational speech," *Computational Linguistics*, Vol.26, No.3, pp. 339-373, 2000.

[6] E. Ivanovic, "Dialogue act tagging for instant messaging chat sessions," *Proceedings of the ACL Student Research Workshop*, pp. 79-84, 2005.

[7] 김경선, 서정연, "자질 선택 기법을 이용한 한국어 화행 결정," *정보과학회논문지: 소프트웨어 및 응용*, Vol.30, No.3, pp. 278-284, 2003.

[8] M. Core and J. Allen, "Coding Dialogs with the DAMSL Annotation Scheme," *Proceedings of the AAI Fall Symposium on Communicative Action*

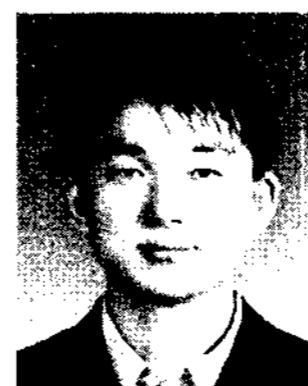
in *Humans and Machines*, pp. 28-35, 1997.

- [9] SRI ATIS Collection, <http://www.ai.sri.com/~communic/amex/amex.html>
- [10] 이현정, 이재원, 서정연, "자동통역을 위한 한국어 대화 문장의 화행 분석 모델," *정보과학회논문지(B)*, Vol.25, No.10, pp. 1443-1452, 1998.
- [11] K. Samuel, et al., "Dialogue Act Tagging with Transformation-Based Learning," *Proceedings of the 36th annual meeting on Association for Computational Linguistics*, Vol.2, pp. 1150-1156, 1998.
- [12] W. Choi, et al., "An Integrated Dialogue Analysis Model for Determining Speech Acts and Discourse Structures," *IEICE Transactions on Information and Systems*, Vol.E88-D, No.1, pp. 150-157, 2005.
- [13] S. Lee and J. Seo, "Korean Speech Act Analysis System Using Hidden Markov Model with Decision Trees," *International Journal of Computer Processing of Oriental Languages*, Vol.15, No.3, pp. 231-243, 2002.
- [14] I. Witten and E. Frank, "Weka - Machine Learning Algorithms in Java," *Data Mining : Practical Machine Learning Tools and Techniques with Java Implementations*, pp. 264-321, 2000.
- [15] N. Reithinger and M. Klesen, "Dialogue Act Classification Using Language Models," *Proceedings of EuroSpeech-97*, pp. 2235-2238, 1997.
- [16] T. Fukada, et al., "Probabilistic Dialogue Act Extraction for Concept Based Multilingual Translation Systems," *Proceedings of the International Conference on Spoken Language Processing*, Vol. 6, pp. 2771-2774, 1998.
- [17] N. Webb, et al., "Dialogue Act Classification Based on Intra-Utterance Features," *Proceedings of the AAI Workshop on Spoken Language Understanding*, 2005.



김 세 종

2006년 충남대학교 전기정보통신공학부 컴퓨터전공 학사. 2008년 포항공과대학교 컴퓨터공학과 석사. 2008년~현재 포항공과대학교 컴퓨터공학과 박사과정. 관심분야는 자연언어처리, 의미분석, 담화 분석 등



이 용 훈

2002년 숭실대학교 컴퓨터학부 학사. 2004년 포항공과대학교 컴퓨터공학과 석사. 2004년~현재 포항공과대학교 컴퓨터공학과 박사과정. 관심분야는 자연언어처리, 구문분석, 의미분석 등



이 종 혁

1980년 서울대학교 수학교육학과 학사

1982년 한국과학기술원 전산학과 석사

1988년 한국과학기술원 전산학과 박사

1989년~1991년 일본전기(NEC) 중앙연

구소 초청연구원. 1991년~현재 포항공

과대학교 컴퓨터공학과 교수. 1998년~

1999년 미국 CRL/NMSU(뉴멕시코주립대학) 방문교수. 관

심분야는 자연언어처리, 기계번역, 정보검색 등