

Hybrid Code Network를 이용한 한국어 식당 예약 시스템 모델*

이동엽 · 허윤아 · 임희석†

† 고려대학교 컴퓨터학과

Korean Restaurant Reservation System Model Using Hybrid Code Network

Dong-Yub Lee · Yun-A Hur · Heui-Seok Lim†

† Dept. of Computer Science and Engineering, Korea University

요 약

대화 시스템(dialogue system)은 텍스트나 음성을 통해 다양한 분야에서 특정한 목적을 수행할 수 있는 시스템이다. 대화 시스템을 구현하기 위한 방법으로 인공 신경망(neural network)을 기반으로한 end-to-end learning 방식이 제안되었다. End-to-end learning 방식을 이용한 식당 예약 시스템 모델의 학습을 위해 페이스북은 영어로 이루어진 식당 예약에 관련된 학습 대화 데이터셋(The 6 dialog bAbI tasks)을 구축하였다. 하지만 end-to-end learning 방식의 학습은 많은 학습 데이터가 필요하다는 단점이 존재하는데, 액션 템플릿(action template)의 정의를 통해 도메인 지식을 표현함으로써 일반적인 end-to-end learning 방식보다 적은 학습량으로 좋은 성능의 모델을 학습할 수 있는 Hybrid Code Network 구조를 제안한 연구가 있다. 본 논문에서는 Hybrid Code Network 구조를 이용하여 한국어 식당 예약 시스템을 구축할 수 있는 방법을 제안하고, 한국어로 이루어진 식당 예약에 관련된 학습 대화 데이터를 구축하는 방법을 제안한다.

1. 서 론

대화 시스템(dialogue system)은 텍스트나 음성을 통해 다양한 분야에서 특정한 목적을 수행할 수 있는 시스템이다. 대화 시스템을 구현하기 위한 전통적인 연구 방법으로는 slot-filling을 기반으로 한 방법[1,2,3]이 있다. Slot-filling은 대화의 각 상태를 slot으로 미리 정의하고 대화가 진행되는 동안 slot을 채우는 방식이다. 식당 예약 시스템의 경우 slot값은 식당의 위치나 가격 또는 요리의 종류 등이 될 수 있다. 이와 같이 해당 도메인(domain)에서 특정 목적을 수행하는데 필요한 속성 값들은 slot으로 정의될 수 있다. 하지만 slot-filling을 이용한 연구는 hand-craft된 자질(feature)를 활용하고 해당 도메인에 대해 많은 지식을 필요로 하기 때문에 새로운 도메인에 대한 확장이 어렵다는 단점이 있다. 이러한 단점을 극복하기 위한 방법으로 인공 신경망(neural network)을 기반으로한 end-to-end learning 방식이 제안되었다[4,5]. 해당 도메인에 대한 지식 없이 학습 데이터 자체만으로 모델의 학습이 가능하고 새로운 도메인에 대한 확장이 비

교적 쉽다는 장점으로 인해 end-to-end learning 방식은 파싱(parsing)[6]이나 기계 번역(machine translation)[7] 그리고 대화 시스템[8]에 적용된 연구가 있다.

기존의 end-to-end learning 방식은 간단한 대화(chit-chat)을 목적으로 한 대화 시스템들에 활용되었지만 최근 end-to-end learning 방식을 이용하여 목적 지향적 대화(goal-oriented dialog)인 식당 예약 시스템에 적용한 연구가 있다[9]. End-to-end learning 방식을 이용한 식당 예약 시스템 모델의 학습을 위해 페이스북은 영어로 이루어진 식당 예약에 관련된 학습 대화 데이터셋(The 6 dialog bAbI tasks)을 구축하였다. 하지만 end-to-end learning 방식의 학습은 많은 학습 데이터가 필요하다는 단점이 존재하는데, 액션 템플릿(action template)의 정의를 통해 도메인 지식을 표현함으로써 일반적인 end-to-end learning 방식보다 적은 학습량으로 좋은 성능의 모델을 학습할 수 있는 Hybrid Code Network 구조를 제안한 연구가 있다[10].

본 논문에서는 Hybrid Code Network 구조를 이용하여 한국어 식당 예약 시스템을 구축할 수 있는 방법을 제안하고, 한국어로 이루어진 식당 예약에 관련된 학습 대화 데이터셋을 구축하는 방법을 제안한다.

* "이 논문은 2017년도 정부(산업통상자원부)의 재원으로 한국산업기술평가관리원의 지원을 받아 수행된 연구임(No. 10079423)."

2. 학습 데이터 수집

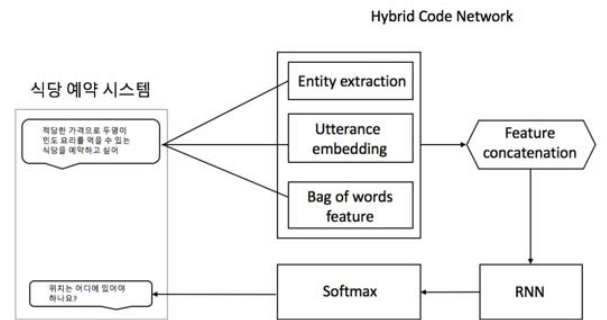
한국어로 이루어진 식당 예약에 관련한 학습 대화 데이터셋을 구축하기 위해 본 연구에서는 페이스북에서 구축한 영어로 이루어진 식당 예약에 관련한 학습 대화 데이터셋(The 6 dialog bAbI tasks)을 이용하였다. 영어로 이루어진 학습 데이터를 이용하여 아래의 표[1] 과 같이 발화의 목적에 따라 영어로 이루어진 각각의 발화(utterance)를 한국어로 번역하여 학습 데이터를 구성하는 것을 제안한다.

순서	발화 목적	영어	한국어
1	인사	hello	안녕
2	예약	can you make a restaurant reservation with (food, location, party_size, price)	(음식, 장소, 인원, 가격)으로 예약하고 싶어
3	예약 수정	instead could it be with (food, location, party_size, price)	~ 로 바꿔줘
4	예약 가능성	no this does not work for me	별로야, 다른거 보여줘
5	추가 변경 거부	no	아니
6	예약 확정	it's perfect	좋아, 훌륭해,
7	식당 장소 확인	can you provide the address	식당 주소 좀 알려줘,
8	식당 연락처 확인	may i have the phone number of the restaurant	식당 연락처 좀 알려줘
9	감사 인사	thank you	고마워
10	추가 도움 거부	no thank you	아니

[표 1] 발화의 목적에 따른 영어 - 한국어 번역

3. Hybrid Code Network

Hybrid Code Network는 액션 템플릿(action template)의 정의를 통해 도메인 지식을 표현함으로써 일반적인 end-to-end learning 모델에 비해 상대적으로 적은 양의 학습 데이터를 필요로 하고, end-to-end learning 방식의 학습이 가능하다는 장점을 가진 모델이다. 이 장에서는 Hybrid Code Network를 이용하여 한국어 식당 예약 시스템을 구축하는 방법을 제안한다. Hybrid Code Network의 구조는 그림[1]과 같다.



[그림 1] Hybrid Code Network 구조도

사용자가 발화를 입력하면 3 가지의 방법으로 자질(feature)을 형성할 수 있다. Entity extraction 모듈은 사용자의 발화로부터 <두명>, <인도 요리> 와 같이 식당 예약에 필요한 속성이 될 수 있는 개체(entity)들을 추출한다. 개체를 추출하기 위해 미리 각 속성들에 해당하는 개체 리스트들을 정의하고 발화로부터 문자열 매칭 알고리즘을 이용하여 발화속에 존재하는 개체들을 추출할 수 있다. Utterance embedding 모듈은 한국어로 이루어진 위키피디아 데이터를 학습시킨 word2vec 모델을 이용하여 사용자의 발화를 임베딩한 후 이를 자질로 구성한다. 마지막으로 사용자의 발화를 대상으로 bag of words 기법을 이용하여 특징을 구성한다. 3 가지의 방법으로 구성된 각 자질들은 연결(concatenation)되어 최종 자질을 구성한다. 구성된 최종 자질은 RNN(recurrent neural networks)의 입력으로 사용된다. RNN은 구성된 자질을 이용하여 은닉 상태(hidden state)를 계산하고 계산된 은닉 상태는 softmax layer의 입력으로 사용된다. 최종적으로 softmax layer는 사용자의 발화에 대한 응답을 생성한다.

4. 결론 및 논의

본 논문에서는 한국어 식당 예약 시스템을 구축하기 위해 식당 예약에 관련한 학습 대화 데이터셋을 수집하는 방법과 Hybrid Code Network 구조를 이용하여

수집된 데이터셋을 학습 시키는 방법을 제안하였다. 제안한 시스템은 일반적인 end-to-end learning 방식의 모델에 비해 상대적으로 적은 양으로 모델의 학습이 가능하고 좋은 성능을 보일것으로 기대된다.

향후 연구에서는 발화의 목적에 따른 한국어 발화를 더욱 다양하게 구성하여 사용자의 다양한 입력에도 식당 예약을 정확하게 진행할 수 있도록 학습 데이터를 수집을 할 것이다. 또한 word2vec이나 RNN 모델의 하이퍼 파라미터 및 가중치(weight) 초기화를 다양한 방법으로 실험함으로써 제안된 한국어 식당 예약 시스템 모델의 성능 향상을 연구할 필요가 있다.

참 고 문 헌

- [1] Lemon, O , Georgila, K, Henderson, J, and Stuttle, M. "An isu dialogue system exhibiting reinforcement learning of dialogue policies: generic slot-filling in the talk in-car system,"In Proceedings of the 11th. Conference of the European Chapter of the ACL:Posters & Demonstrations, pages 119 - 122. 2006
- [2] Wang, Z. and Lemon, O."A simple and generic belief tracking mechanism for the dialog state tracking challenge: On the believability of observed information," In Proceedings of the SIGDIALConference. 2013
- [3] Young, S, Gasic, M, Thomson, B , and Williams, J.D. "Pomdp-based statistical spoken dialog systems: A review,"Proceedings of the IEEE, 101(5), 1160 - 1179.2013
- [4] Ilya, S, James M, and Geoffrey E, H. "Generating text with recurrent neural networks,"In Proceedings of the 28th International Conference on Machine Learning (ICML), pages 1017 - 1024. 2011
- [5] Ilva, S, Oriol, V, and Quoc, VV L."Sequence to sequence learning with neural networks,"In Advances in Neural Information Processing Systems (NIPS), pages 3104 - 3112. 2014
- [6] Oriol, V, Łukasz, K, Terry, K, Slav P, Ilya, S, and Geoffrey, H. "Grammar as a foreign language," In Advances in Neural Information Processing Systems, pages 2755 - 2763. 2015
- [7] Dzmitry, B, Kyunghyun, C, and Yoshua, B. "Neural machine translation by jointly learning to align and translate," In ICLR. 2015
- [8] Iulian, V, S, Alessandro, S, Yoshua, B, Aaron C, C, and Joelle P. "Building end-to-end dialogue systems using generative hierarchical neural network models,"In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, pages 3776 - 3784. 2016
- [9] Bordes,A, and Weston,J, "Learning End-to-End Goal-Oriented Dialog," ArXiv e-prints. 2016
- [10] Jason D Williams, Kavosh Asadi, and Geoffrey Zweig. 2017. Hybrid code networks: practical and efficient end-to-end dialog control with supervised and reinforcement learning. In ACL.