페르소나 대화모델에서 일관된 발화 생성을 위한 연구

문현석○, 이찬희, 임희석*

고려대학교 컴퓨터학과

glee889@korea.ac.kr, chanhee0222@korea.ac.kr, limhseok@korea.ac.kr

Personality Consistent Dialogue Generation in No-Persona-Aware System

Hyeonseok Moon○, Chanhee Lee, Heuiseok Lim*

Dept. of Computer Science and Engineering, Korea University

요 약

일관된 발화를 생성함에 있어 인격데이터(persona)의 도입을 이용한 연구가 활발히 진행되고 있지만, 한국어 데이터셋의 부재와 데이터셋 생성의 어려움이 문제점으로 지적된다. 본 연구에서는 인격데이터를 포함하지 않고 일관된 발화를 생성할 수 있는 방법으로 다중 대화 시스템에서 사전 학습된 자연어 추론(NLI) 모델을 도입하는 방법을 제안한다. 자연어 추론 모델을 이용한 관계 분석을 통해 과거 대화 내용 중 발화 생성에 이용할 대화를 선택하고, 자가 참조 모델(self-attention)과 다중 어텐션(multi-head attention) 모델을 활용하여 과거 대화 내용을 반영한 발화를 생성한다. 일관성 있는 발화 생성을 위해 기존 NLI데이터셋으로 수행할 수 있는 새로운 학습모델 nMLM을 제안하고, 이 방법이 일관성 있는 발화를 만드는데 기여할 수 있는 방법에 대해 연구한다.

주제어: Multi turn Dialogue, NLI, Pretraining

1. 서론

발화 생성 모델이란 메세지와 그에 대한 반응을 학습하여 주어지는 메세지에 대한 적절한 대답을 생성해내는 모델을 말한다. 다중대화(multi turn dialogue)시스템이란 메세지와 반응 뿐 아니라, 같은 인물이 했던 과거 대화까지 훈련데이터에 포함시키는 모델을 의미한다.

발화생성 모델은 기본적으로 seq2seq모델을 통해 학습한다. 하지만 아무런 외부 정보 없이 seq2seq모델로 발화생성을 학습시킬 때, 여러 한계점이 존재한다는 사실이 연구를 통해 알려졌다[1, 2]. 단순히 메세지와 그에 대한 반응만으로 학습을 진행했을 때, 발생하는 문제점으로 4가지를 지적할 수 있다. 첫번째로는 일관되지 않는 발화 생성에 대한 문제[1], 두번째로는 과거 대화를 모두 훈련에 반영할 수 없어 생기는 문제[1], 그리고 세번째로는 발화에 생략된 정보를 다루는 문제이다[2].

이 문제를 해결하기 위해 제안된 방법 중 하나는 발화자의 특성을 담은 여러 평문들로 이루어져 있는 인격데이터(persona)를 추가하는 것이다[1]. 다만 인격데이터가 추가되어 있는 발화 데이터셋이 충분히 존재하지 않는다는 문제점이 있다. 특히 제대로 준비된 인격데이터가 포함된 한국어 발화 데이터셋은 현재 존재하지 않는 것으로 보인다. X-persona에서 발표한 데이터셋[3] 중, 외국어 데이터셋(multilingual)에 한국어 데이터셋이 존재하기는 하지만, 이는 영어 데이터셋을 번역하여 만들어졌기에 실제 대화에서 쓰이지 않는 번역체와 한국 정서에 맞지 않는 대화 내용이 다수 포함되어 있어 실제 사용하기에는 무리가 있다.

본 연구에서는 다중 대화 시스템에서의 발화 생성 모델을 기반으로 하여, 인격 데이터를 도입하지 않고 위에서 언급된 문제를 해결할 수 있는 방안에 대해 탐구하고자 한다. 이전에 했던 발화와 모순되지 않는 발화는 과거에 비슷한 질문을 받았을 때 보였던 반응과 유사한 반응을 보이는 발화를 의미한다. 과거에 받았던 질문에 대한 발화자의 답변은, 그 발화자의 특성을 나타내는 하나의 데이터로 인식될 수 있다. 만약 과거 대화 중에서 받은 질문과 연관성이 높은 질문을 선택할 수 있고, 과거에 보였던 반응과 일관된 발화를 만들어낼 수 있다면, 인격 데이터를 추가하지 않고도 앞서 제기한 문제를 해결할 수 있을 것이다. 이 때 과거 대화 중에서 현재 발화와 연관성이 높은 대화를 선택하는 방법으로 자연어 추론(Natural Language Inference, NLI)모델을 활용할 수 있다. 최근 연구에서 문장 간의 연관성을 파악하는 데에 NLI학습의 영향이 다수 연구되고 있고[4], 문장 간의 관계를 이용하는 모델에서 NLI를 사전 학습시키면, 전체 학습에 긍정적인 영향을 미친다는 연구가 발표된 바 있

* 교신저자(Corresponding author).

다[4, 5, 6]. 본 연구에서는 사전 학습된 자연어 추론 모델을 활용하여 현재 발화 생성과 연관이 있는 과거 대화 내용을 선택하는 방법과, 이를 활용하여 과거 발화와 연관된 발화를 만들어내는 방법을 제안한다.

2. 관련 연구

최근 일관된 발화를 생성하는 모델은 과거 대화 데이터(Dialogue history)와 인격데이터의 활용방법을 중심으로 활발하게 연구되고 있다[7, 8, 9]. 특히 인격데이터의 사용은 일관된 발화 생성에 있어 매우 중요하다고 볼 수 있다[10]. 인격데이터를 활용하지 않고 설계한 발화생성 모델에 인격데이터를 추가했을 때, 그 성능이 전반적으로 향상되었다는 연구결과도 발표된 바 있다[11]. 그럼에도 불구하고 인격데이터가 포함되지 않은 연구(no persona aware)는 인격데이터가 포함된 연구(persona aware)보다 더 활발히 이루어지고 있다[12]. 이에 대한 대표적인 이유로, 인격데이터가 포함된 양질의 데이터셋을 만들어 내기 어렵다는 점을 지적할 수 있다[13].

인격데이터를 포함하지 않고 발화시스템에서 발생하는 문제를 해결하는 방법으로, 다중 대화 시스템에서 발화를 생성하는 것을 생각해볼 수 있다. 일반적으로 사람이 대화할 때, 예전에 했던 모든 대화내용을 생각하면서 말을 만들어내지는 않는다. 즉, 발화는 과거의 모든 대화가 아닌, 일부 대화에만 영향을 받아 생성되기에 모든 과거 대화내용을 발화생성에 이용하기보다, 관련이 깊은 일부 대화만을 반영하여 발화를 만드는 데에 집중해야 한다. 다중 대화 발화생성 모델에서는 발화생성에 이용할 대화를 선택하는 방법과 그 데이터를 활용하는 방법이 집중적으로 연구되고 있다.

과거 대화 데이터를 활용한 발화생성 방법으로, 먼저 과거 대화들 중 선택한 한 문장과 직전 질문을 연결(concatenate)하여 직접적으로 발화 생성에 이용한 연구사례가 있다[14]. 이후 전체 과거 대화 문장들을 직전 질문과 연관 지어 발화를 생성하는 방법에 대한 연구도 진행되었다[15]. 이 연구에서는 과거 대화를 이용하는 데 생각할 수 있는 방법 9가지를 제시하고, 각각에 대한 실험을 통해 가장 좋은 성능을 내는 방법을 밝혔다. 논문에서 발표한 바에 따르면, 과거 대화에 가중치를 주고 연결(weighted concatenation)한 방법이 가장 높은 성능(BLEU score)을 보였다.

과거 대화에 가중치를 두어 연결하는 것 이상으로, RNN을 통해 과거 문장들 간의 연관성을 추가로 학습에 이용한 연구사례도 있다[16]. 다만 과거 대화 내용을 각각 순차적으로 발화 생성에 이용하여 병렬처리가 불가능하고, 순차적으로 정보가 전달되는 RNN 구조의 특성상, 특정 대화 내용에 집중하기 어렵다는 점이 지적된다[17]. 이에 어텐션(attention)구조를 추가하여 대화 내용들 중 발화 생성에 더 중요한 대화를 선택하고, 문장 내 단어 수준에서도 발화 생성에 중요한 부분을 확인하는 HRAN(Hierarchical Recurrent Attention Network)모델이 발표되었다[17]. 이후 HRAN에서 과거 대화와 현재 질문 간의 연관성을 확인할 때 cosine유사도를 평가기준으로 삼는 부분에 문제를 제기하고 자가 참조 어텐션(self-attention) 구조[18]를 도입하여 문장 간의 유사도나 중요도를 파악하는 연구도 발표된 바 있다[19].

과거 대화 내용 중에서 현재 발화와 연관된 대화 내용을 찾는 데에 자연어 추론이 학습된 모델을 이용한 연구사례도 최근 발표된 바 있다[4]. 이 논문에서는 자연어 추론이 학습된 모델을 통해 과거 발화 내용과 생성할 발화 간의 관계를 파악하고, 만약 두 문장이 서로를 함의(entailment)하는 관계라면, 해당 과거 발화를 현재 발화를 생성하는 데 핵심문장(key turn)으로 간주하는 방법을 제시한다. 본 논문에서는 과거 대화가 현재 질문이나 발화를 함의 하는 관계인 경우일 뿐 아니라 중립, 모순되는 경우에도 이를 발화 생성에 이용하는 방법을 제시한다.

3. 제안하는 모델

본 연구에서 제안 발화 생성은 크게 3가지 단계로 분류된다. 첫번째로 자연어 추론이 사전 학습된 모델을 이용하여 현재의 질문과 과거 대화에 존재하는 질문간의 관계를 분석한다. 과거에 받았던 질문과 유사한 질문을 받았다면 과거에 했던 대답과 유사한 발화를 생성해야 하고, 과거에 받았던 질문과 반대되는 질문을 받았다면 과거에 했던 대답과 반대되는 발화를 생성해야 한다. 따라서 과거와 현재의 질문이 함의, 또는 모순관계에 더 가까운 대답을 선택하여 발화 생성에 참고한다. 두번째 단계에서는 첫번째 단계에서 선택한 발화를 참고하여 발화를 생성하고, 마지막 단계에서는 생성한 발화가 과거 대화와 일관성을 갖도록 발화를 수정한다.

두번째와 마지막 단계는 GDR모델[7]에서 제시된 발화생성 방법을 참고한다. GDR는 인격 데이터를 활용한 단일 발화 생성(single turn dialogue) 모델로, BERT[20]을 통해 발화를 생성한 후, 생성된 발화가 인격 데이터와 일관되도록 발화 수정작업을 거친다. 본 논문에서는 인격데이터 대신 과거 대화를 활용하고, 발화의 수정 단계에서 사전 학습된 모델이 활용될 수 있는 방안을 추가로 제시한다.

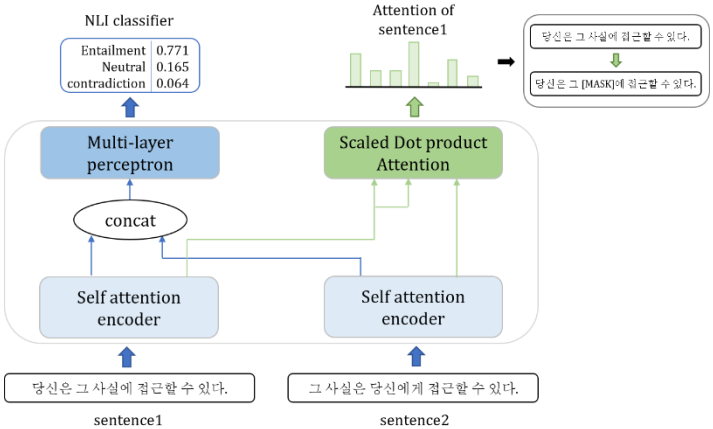

그림 1. NLI모델과 어텐션 기반 masking.

3.1 사전 학습된 자연어 추론(NLI) 모델

자연어 추론 모델은 전제(premise)와 가설(hypothesis) 두 문장 간의 관계를 함의(entailment), 중립(neutral), 모순(contradiction) 세 가지 중 하나로 분류하는 작업을 학습한다[21].

현재 NLI훈련을 위한 한국어 데이터셋은 카카오브레인에서 발표한 KorNLI[22]데이터셋이 있고, github에 공개된 KoBERT[1]와 KoELECTRA[2] 모델을 이용했을 때, 약 80%정도의 NLI 작업 정확도를 얻었다는 결과가 있다. 그림 1의 좌측 모델은 BERT모델[20]을 활용한 NLI 분류모델의 예시를 보여준 것이다. 먼저 두 문장을 자가 참조 어텐션 모델을 통해 각각 구조화된 벡터 u와 v로 만든다. 이후 u − v 의 절댓값과 원소 단위 곱(element-wise product)을 통해 얻어진 u ∗ v를 u, v와 함께 연결(concatenate)하여 [u; |u − v|; u ∗ v; v] 형태의 벡터를 얻는다. 이 벡터는 완전연결층을 연결한 모델(multi-layer perceptron)을 거쳐 최종적으로 3개 분류군에 대한 확률값을 얻게 된다[5]. 이 과정을 통해 학습된 자가 참조 어텐션 모델은 문장 간의 관계 파악을 위한 구조화 단계에 이용될 뿐 아니라, 3.3장 에서 사용될 자가 참조 어텐션 구조의 사전학습 모델로도 이용된다.

3.2 자연어 추론을 활용한 과거 대화 선택

가장 먼저 과거 대화록 중 현재 발화 생성에 참고할 대화를 선택한다. 이 과정은 과거 대화가 많이 축적되어 모든 대화를 전부 이용하기 어려운 상황[1]에서, 훈련에 사용할 대화를 선택하는 적절한 기준이 될 수 있다.
과거 대화록에 있는 질문과 답변을 각각

$$(Hq_1, Hq_2, \dots, Hq_n),\ (Ha_1, Ha_2, \dots, Ha_n)$$

이라고 하고, 발화 생성의 대상이 되는 직전 질문(query)을 Q라고 하자. 이 단계에서는 사전 학습된 모델을 통해 각 Hq_i,와 Q간의 관계를 분석한다. 이를 Nq_i라 하자. Nq_i는 함의, 중립, 모순에 해당하는 0과 1사이의 값(각각 eq_i, nq_i, cq_i)으로 구성된 3*1벡터이다.

$$Nq_i = NLI(Hq_i, Q) = \begin{pmatrix} eq_i \\ nq_i \\ cq_i \end{pmatrix}$$

이 때, 함의에 해당하는 값이 높게 나온다면 모델을 통해 생성해야 할 발화는 Ha_i,와 유사한 반응을 보여야 하고, 모순에 해당하는 값이 높게 나온다면 생성해야 할 발화는 Ha_i,와 반대되는 반응을 보여야 한다. 만약 중립에 해당하는 값이 높게 나온다면 이 대화는 모델을 통해 생성할 발화에 크게 영향을 미치지 않는다고 생각할 수 있다. 즉, $eq_i - nq_i$ 의 값이 크거나 $cq_i - nq_i$ 값이 클수록 Hq_i와 Ha_i는 생성할 발화와 관계가 큰 대화내용이라 생각할 수 있다. 이에 따라 $\max(eq_i - nq_i, cq_i - nq_i)$을 평가기준으로 삼아, 그 값이 큰 순서대로 k개의 대화를 선택한다. 이렇게 선택한 과거 대화의 질문들을 $(hQ_1, hQ_2, \dots, hQ_k)$ 라 하고, 그에 해당하는 대답들을 $(hA_1, hA_2, \dots, hA_k)$라고 하자.

3.3 과거 대화를 활용한 발화생성

이 단계에서는 전 단계에서 선택한 과거 대화내용과 현재 질문을 이용하여 일차적으로 발화를 생성한다. 그 과정을 요약하면 그림2와 같다.

먼저 문장 내에서 단어 간의 관계를 얻기 위해 자가 참조 모델[18]을 이용한다. 이 때, 각 문장들은 임베딩 층을 (embedding) 거친 이후, 위치 임베딩을 더한 결과물로 간주한다[20]. 이 모델(SAE)을 이용하여 과거 대화의 질문 hQ_i를 구조화(encode)시킨 결과물을 OQ_i라 하고, hA_i를 구조화한 결과물을 OA_i, 직전 질문 Q를 구조화한 결과물을 OQ라고 하자.

$$OQ_i = SAE(hQ_i)$$
$$OA_i = SAE(hA_i)$$
$$OQ = SAE(Q)$$

본 연구에서는 '과거에 받았던 질문과 그에 대한 발화자의 대답'을 하나의 인격 데이터와 같게 볼 것이다. 질문과 대답을 연관 지어 하나의 구조화된 출력물을 만들기 위해 다중 어텐션 모델(Multi-head Attention)[18]을 사용한다. 이를 $MHA(Q,K,V)$ 라고 하자. Q,K,V 는 각각 query, key, value를 의미한다.

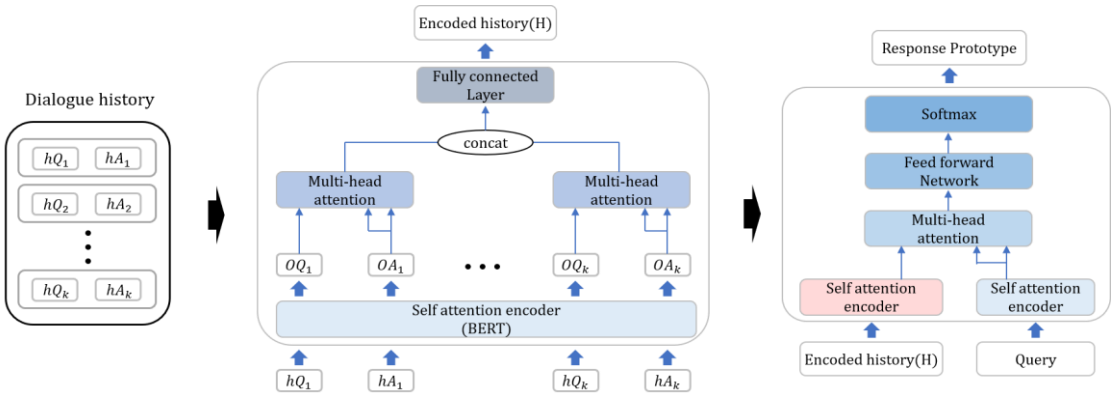

그림 2. 1차 발화생성모델.

[1] https://github.com/SKTBrain/KoBERT

[2] https://github.com/monologg/KoELECTRA

위 과정을 통해 얻은 한 쌍의 OQ_i, OA_i는 다중 어텐션 모델을 거쳐 OP_i가 된다. 만들어진 $(OP_1, OP_2, \dots, OP_k)$는 하나의 결과물로 연결되고 완전연결 레이어(Fully connected)를 거쳐 과거 대화에 대한 정보를 모두 담고 있는 결과물 H가 된다.

$$OP_i = MHA(OQ_i, OA_i, OA_i)$$
$$H = \text{FullyConnected}([OP_1; OP_2; \dots; OP_k])$$

구조화된 정보들의 단순 연결을 통해 만들어진 H는, 대화 내용들 간의 연관성을 담고 발화생성에 더 큰 영향을 줄 대화내용을 결정하기 위해 자가 참조 모델을 거친다. 이 자가 참조 모델은 앞서 OQ_i를 만들기 위해 이용된 자가 참조 모델과 다른 가중치를 가진다. 이후 OQ와 함께 다중 어텐션 모델을 거치고, Feedforward모델과 Softmax를 통해 1차 발화(prototype) $\dot{Y}$을 생성한다.

$$\dot{Y} = MHA(SAE(H), OQ, OQ)$$
$$\dot{Y} = Softmax(FFN(\dot{Y}))$$

3.4.1 nMLM(NLI Masked Language Model)

발화 재생성 모델에 들어가기에 앞서 NLI dataset을 활용한 MLM을 사전학습모델을 제시하려 한다. MLM은 BERT논문[20]에서 제시된 학습 모델로, 일부가 가려진 문장을 원래 문장으로 복원하는 작업을 훈련시킨다. 본 논문에서 진행할 nMLM에서는, 일부가 가려진 문장 외에 하나의 문장을 더 추가하고, 가려진 문장과 추가된 문장 간의 관계를 추가하여 학습을 진행한다.

nMLM의 훈련 데이터 생성방법은 그림1의 오른쪽 그림과 같다. 자연어 추론 훈련데이터에서 nMLM의 훈련데이터를 만들기 위하여, 먼저 자가 참조 모델을 거쳐 각 문장의 구조화된 벡터를 얻는다. 이 때 3.1 과정에서 자연어 추론을 학습했던 모델을 사전 학습 모델로 하여 본 훈련을 진행한다. 이 때, 자가 참조 모델의 은닉층 개수를 n_{hidden}, 모델이 입력으로 받을 수 있는 토큰의 최대 개수를 n_{max}라고 하자.

자연어 추론 훈련데이터에 존재하는 i번째 데이터의 전제, 가설, 관계를 각각 p_i, h_i, idx_i라 하고 자가 참조 모델을 거친 p_i와 h_i를 op_i, oh_i라 하자. 여기서 op_i, oh_i는 $n_{max} * n_{hidden}$의 구조를 가진다. op_i와 oh_i는 서로의 scaled dot product attention[18]값을 구하는 데에 이용된다.

$$Ap_i = \frac{op_i\, oh_i^T}{\sqrt{n_{hidden}}}\, oh_i$$

Ap_i는 op_i에 대한 oh_i의 어텐션 값을 의미한다. $n_{max} * n_{hidden}$의 구조를 가지는 Ap_i에서 각 행의 평균을 구함으로써 oh_i에 대한 최종 어텐션 값을 얻는다. 여기서 높은 값을 가진 토큰 순서대로 oh_i 전체 토큰 수의 10%만큼 [mask]토큰으로 대체하여 일부가 가려진 전제 문장 mh_i를 얻는다. 이 과정을 통해 $([mh_i, op_i, idx_i], oh_i)$형태의 데이터셋을 얻을 수 있다.

위 과정을 통해 얻은 데이터셋으로 nMLM을 훈련하는 과정은 그림3과 같다. nMLM에서는 복구할 문장과의 관계를 알고 있는 문장을 활용하여, 일부가 가려진 문장을 원래 문장으로 복구하는 작업을 훈련한다. 가장 먼저 문장 내에서 단어 간의 관계 파악을 위해 3.1과정에서 활용된 자가 참조 어텐션 구조를 이용한다.

$$Omh_i = SAE(mh_i)$$
$$Oop_i = SAE(op_i)$$

Omh_i와 Oop_i는 mh_i와 op_i를 각각 자가 참조 모델을 거쳐 얻은 구조화된 결과물이다. 이 두 결과물은 다중 어텐션 모델을 거쳐 최종 결과물을 만들어낸다. 과정 3.2에서 과거 대화를 선택할 때, 선택한 대화에는 함의관계에 있는 대화에 더해 모순관계에 있는 대화도 포함되어 있으므로, 이 과거 대화를 이용하여 빈 단어를 유추하는 훈련에는 idx_i도 함께 이용되어야 한다. idx_i는 완전연결 레이어를 통해 Oop_i와 크기를 맞추고, Oop_i와 곱해져 MHA 의 query가 되는 방식으로 훈련에 이용된다.

$$Oidx_i = FullyConnected(idx_i)$$
$$out_i = MHA(Oop_i * Oidx_i, Omh_i, Omh_i)$$
$$\widetilde{out_i} = Softmax(FFN(out_i))$$

이 방법으로 생성된 $\widetilde{out_i}$와 oh_i간의 cross-entropy를 최소화 하는 방향으로 학습을 진행한다.

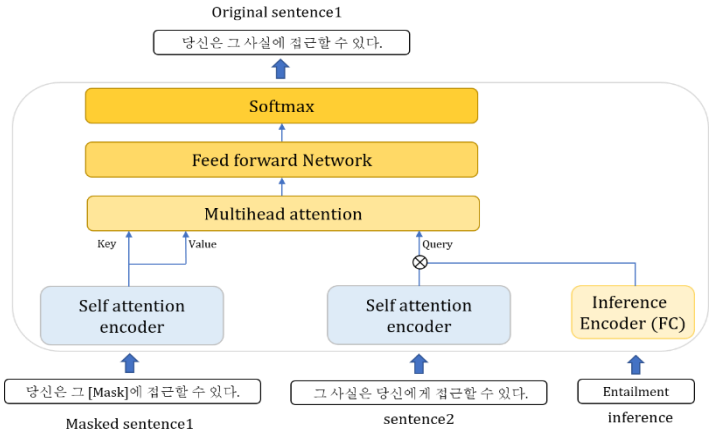

그림 3. nMLM.

3.4.2 발화의 제거와 재생성

이 단계에서는 3.3과정에서 일차적으로 생성한 발화 $\check{Y}$에서 과거 대화와 일관성이 맞지 않는 부분을 제거하고, 일관성을 갖도록 발화를 수정하는 과정을 거친다.

$\check{Y}$에서 일관되지 않는 부분을 제거하는 과정은 그림 4와 같다. $\check{Y}$에서 과거 대화와 일관되지 않은 부분을 표시하기 위하여, 3.4.1에서 Ap_i,를 만들어내는 과정과 유사하게, 3.2과정에서 선택한 문장 $(hA_1, hA_2, \dots, hA_k)$ 각각에 대한 $\check{Y}$의 어텐션 값을 구한다.

$$OhA_i = \text{SAE}(hA_i)$$
$$O\check{Y} = SAE(\check{Y})$$
$$Ay_i = \frac{OhA_i\, O\check{Y}^T}{\sqrt{n_{hidden}}}\, O\check{Y}$$

만약 hA_i와 Q가 서로를 함의 하는 관계라면, Ay_i에서 높은 값을 가지는 부분은 일관된 발화를 만드는 데 중요한 역할을 하는 부분이라 생각할 수 있다. 반면 hA_i와 Q가 서로 모순되거나 중립적인 관계라면, Ay_i에서 높은 값을 가지는 부분은 일관된 발화를 만드는 데 방해가 되는 부분으로 이해할 수 있다.

이에 따라 함의관계일때 높은 어텐션을 가지는 부분은

빼고, 중립이나 모순관계일때 높은 어텐션을 가지는 부분을 더해서 $\check{Y}$에서 대체할 토큰를 찾는다. 3.2에서 구한 Nq_i를 통해 계산한 $cq_i + nq_i - eq_i$를 Ay_i에 곱해주고, 그 평균값을 구한다면, 이때 가장 높은 값을 가지는 부분이 일관된 발화를 만들지 못하게 하는 부분이라 생각할 수 있다.

$$\widetilde{Ay_i} = (cq_i + nq_i - eq_i) * Ay_i$$
$$Attention\_final = Average(\widetilde{Ay_i})$$

$Attention_final$의 값이 큰 토큰 순서대로 $\check{Y}$ 전체 토큰 길이의 10%만큼을 [mask]토큰으로 대체하여 일부가 가려진 결과물 MY를 얻는다.

위 과정을 통해 얻은 MY는 과거 대화에서의 질문들 $(hQ_1, hQ_2, \ldots, hQ_k)$과 함께 사전 학습된 nMLM을 거쳐 최종 결과물을 만들어낸다. nMLM을 통해 발화를 재생성하기 위해선 관계를 알고 있는 문장이 필요하다. 여기서, 일관된 발화시스템에서 두 질문간의 관계는 두 질문에 대한 대답 간의 관계와 같다는 가정을 할 필요가 있다. 이 가정 하에, $(hQ_1, hQ_2, \ldots, hQ_k)$ 와 Q 간의 관계를 $(hA_1, hA_2, \ldots, hA_k)$와 이상적인 결과물 Y간의 관계와 같다고 생각하고, 사전 학습된 자연어 추론 모델을 통해 $(hQ_1, hQ_2, \ldots, hQ_k)$와 Q간의 관계$(idx_1, idx_2, \ldots, idx_k)$를 구한다.

최종적으로 다음 과정을 각 i에 대해 k번 시행하는 것으로, 최종 결과물 $output_k$를 얻는다. $output_0 = MY$라 한다.

$$output_i = \text{nMLM}(output_{i-1}, hA_i, idx_i)$$

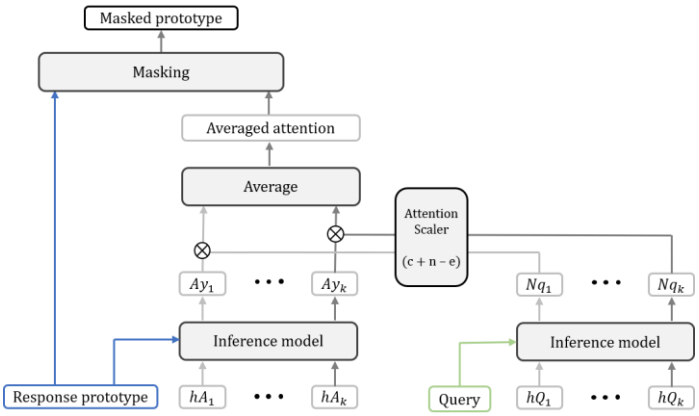

그림 4. 발화 제거(masking) 과정.

3.5 실험 설계

본 모델의 정확도는 BERT-score[23]를 통해 평가할 예정이다. 정답이 정해진 발화시스템이 아니고, 같은 단어라도 현재 질문이나 과거 대화내용 등, 상황에 따라 다른 해석이 요구되기 때문에 정확도를 평가할 때 문맥적 의미도 함께 고려하는 평가기준이 필요하다. 단, 평가기준뿐만 아니라 본 연구에서 제시된 자연어 추론, nMLM, 발화생성 세가지 모델에서도 모두 사전 학습된 BERT를 요구하기 때문에 신뢰할 만한 정확도를 보이는 BERT모델을 도입해야 할 필요가 있다. 한국어 모델의 경우 SKTBrain에서 발표한 KoBERT[3]를 사전학습 모델로 활용했을 때, 77.91%의 KorNLI[22] 정확도를 얻은 바 있다.

자연어 추론과 nMLM 학습의 경우, KorNLI[22]데이터셋을 사용하면 될 것으로 생각된다. 본 연구에서 제시된 nMLM의 경우 하나의 자연어추론 훈련 데이터당 두개의 nMLM 훈련 데이터를 얻을 수 있다.

발화생성 모델의 훈련데이터로 생각할 수 있는 것은 aihub에서 발표한 일상 대화 데이터[4]이다. 일상에서 이루어지는 대화들을 담고 있는 데이터셋으로, 서로 다른 두 발화자가 주고받은 대화 약 35,000 문장으로 구성 되어 있다.

4. 결론

본 연구는 다중 대화시스템 연구에서 일관된 대화를 만들어 내는 데에 사전 학습된 외부 모델을 도입하는 새로운 방법을 제시했다는 점에서 의의가 있다. 자연어 추론 모델을 이용하여 과거 대화들 중에서 중요한 부분을 선택하는 방법에 대해서는 논의된 바 있지만[4], 분석된 관계가 중립이거나 모순인 문장들을 활용하는 방법, 그리고 발화자의 대화 내용 뿐 아니라 상대방의 대화내용도 활용하는 방법을 제시했다는 점에서 논의에 진전이 있었다고 볼 수 있다.

향후 연구 방향으로 크게 두가지를 제시할 수 있다. 먼저 본 연구는 실험을 통해 입증가능한 결과를 내놓지 못했다는 부분에서 명확한 한계가 존재하므로, nMLM과 전체 모델이 실제 학습을 통하여 성능 평가가 이루어져야 할 필요가 있다. 또한 3.4.2에서 가정하고 넘어갔던 부분이 실험을 통해 입증되어야 한다. 일관된 발화시스템에서 두 질문간의 관계는 두 질문에 대한 대답 간의 관계와 같다는 것을 확인할 모델과 적절한 평가기준이 제시되어야 할 필요가 있다.

5. 감사의 말

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기술기획평가원의 지원을 받아 수행된 연구(No. 2020-0-00368, 뉴럴-심볼릭(neural-symbolic) 모델의 지식 학습 및 추론 기술 개발)와 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT명품인재양성사업의 연구결과로 수행되었음 (IITP-2020-0-01819)

참고문헌

[1] ZHANG, Saizheng, et al. Personalizing dialogue agents: I have a dog, do you have pets too?. arXiv preprint arXiv:1801.07243, 2018.
[2] SU, Hui, et al. Improving multi-turn dialogue

[3] https://github.com/SKTBrain/KoBERT

[4] https://aihub.or.kr/aidata/85/download

modelling with utterance ReWriter. arXiv preprint arXiv:1906.07004, 2019.
[3] LIN, Zhaojiang, et al. XPersona: Evaluating Multilingual Personalized Chatbot. arXiv preprint arXiv:2003.07568, 2020.
[4] LI, Junlong; ZHANG, Zhuosheng; ZHAO, Hai. Multi-choice dialogue-based reading comprehension with knowledge and key turns. arXiv preprint arXiv:2004.13988, 2020.
[5] CONNEAU, Alexis, et al. Supervised learning of universal sentence representations from natural language inference data. arXiv preprint arXiv:1705.02364, 2017.
[6] SUBRAMANIAN, Sandeep, et al. Learning general purpose distributed sentence representations via large scale multi-task learning. arXiv preprint arXiv:1804.00079, 2018.
[7] SONG, Haoyu, et al. Generate, Delete and Rewrite: A Three-Stage Framework for Improving Persona Consistency of Dialogue Generation. arXiv preprint arXiv:2004.07672, 2020.
[8] MESGAR, Mohsen, et al. Generating Persona-Consistent Dialogue Responses Using Deep Reinforcement Learning. arXiv preprint arXiv:2005.00036, 2020.
[9] KOTTUR, Satwik; WANG, Xiaoyu; CARVALHO, Vítor. Exploring Personalized Neural Conversational Models. In: IJCAI. p. 3728-3734. 2017.
[10] ZHENG, Yinhe, et al. Personalized dialogue generation with diversified traits. arXiv preprint arXiv:1901.09672, 2019.
[11] MAZARÉ, Pierre-Emmanuel, et al. Training millions of personalized dialogue agents. arXiv preprint arXiv:1809.01984, 2018.
[12] HAN, Qiang. Improvement of a dedicated model for open domain persona-aware dialogue generation. arXiv preprint arXiv:2008.11970, 2020.
[13] ZHENG, Yinhe, et al. A Pre-Training Based Personalized Dialogue Generation Model with Persona-Sparse Data. In: AAAI. p. 9693-9700. 2020.
[14] MOU, Lili, et al. Sequence to backward and forward sequences: A content-introducing approach to generative short-text conversation. arXiv preprint arXiv:1607.00970, 2016.
[15] TIAN, Zhiliang, et al. How to make context more useful? an empirical study on context-aware neural conversational models. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). p. 231-236. 2017.
[16] SERBAN, Iulian V., et al. Building end-to-end dialogue systems using generative hierarchical neural network models. arXiv preprint arXiv:1507.04808, 2015.
[17] XING, Chen, et al. Hierarchical recurrent attention network for response generation. arXiv preprint arXiv:1701.07149, 2017.
[18] VASWANI, Ashish, et al. Attention is all you need. In: Advances in neural information processing systems. p. 5998-6008. 2017.
[19] ZHANG, Hainan, et al. Recosa: Detecting the relevant contexts with self-attention for multi-turn dialogue generation. arXiv preprint arXiv:1907.05339, 2019.
[20] DEVLIN, Jacob, et al. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805, 2018.
[21] FYODOROV, Yaroslav; WINTER, Yoad; FRANCEZ, Nissim. A natural logic inference system. In: Proceedings of the 2nd Workshop on Inference in Computational Semantics (ICoS-2). 2000.
[22] HAM, Jiyeon, et al. KorNLI and KorSTS: New Benchmark Datasets for Korean Natural Language Understanding. arXiv preprint arXiv:2004.03289, 2020.
[23] ZHANG, Tianyi, et al. Bertscore: Evaluating text generation with bert. arXiv preprint arXiv:1904.09675, 2019.