**Masked Hierarchical Transformer**

**Who Did They Respond to? Conversation Structure Modeling Using Masked Hierarchical Transformer(AAAI 2020)**

대화는 문어체(위키피디아, 책 등)와 많은 차이가 있다.

자주 사용되는 어휘, 어투 등 발화의 형태적 차이 + 두 명 이상의 화자가 서로 상호작용 하면서 생기는 구조적 차이

Transformer 구조 기반의 BERT 기반은 많은 SOTA!

하지만 문어체 데이터로 주로 pre-training이 있기 때문에 대화에 적용하기 애매

=> 그 부분을 해결하려고 했다.

**Main Idea**

Pre-trained BERT를 문장 인코더로 이용 -> 추가적인 트랜스포머 인코더는 문장의 구조 파악으로 이용 => 대화 구조 모델링

즉 문장 인코딩 결과를 이용하여 다시 트랜스포머 인코더를 거쳐 컨텍스트 인코딩을 함 => 2단계의 계층적 구조

컨텍스트 인코더를 학습할 때, 대화에서 각 발화의 의존 관계를 파악할 수 있도록 하는 Masking 방법 제시

**Conversation Disentanglement**

대화는 여러 개의 “thread”로 구성 (단톡 방을 생각)

“특정 발화”가 “이전의 대화” 중 어떤 발화의 대답인지 “Reply-to” 관계를 파악하는 것이 목적

이를 통해 대화 구조를 계층적인 그래프의 형태로 분리 가능

이전 논문에서는 (문장1, 문장2)의 문장 쌍들에 대해 문장1의 대답이 문장2인지 구분하는 이진 분류 문제로 해결 -> 히스토리 중 문장1 이외에는 다른 발화 고려 x (문장 쌍이니까)

이 방식대로라면 아래 그림의 utterance4가 utterance3에 대해 “reply-to” 형성 가능  
(이 쌍만 봤을 때 문맥이 맞긴 하니까)

하지만 utterance4는 utterance2에 답변하는 것이 였음.(인간은 쉽게 알아도 기계는 어렵)

따라서 모든 히스토리를 고려하는 방식을 제시한다.

