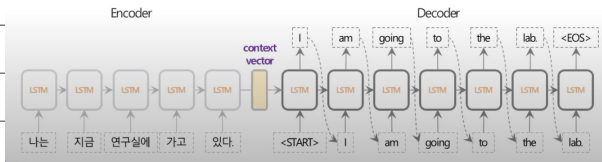


## Attention

### Seq-2-Seq 간단설명

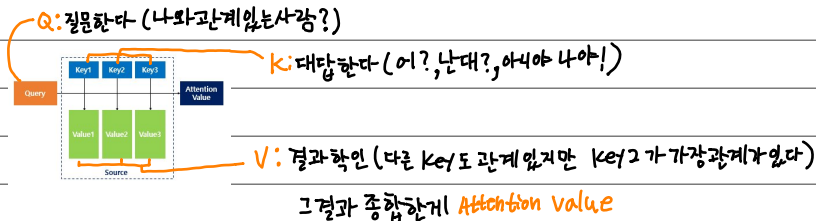
- NLP의 기본 모델 Seq 2Seq에서는 인코더가 입력 시퀀스를 (context인 Vector로 압축하고  
디코더가 이 벡터를 통해서 출력 시퀀스를 만들어 낸다
- 문제점 발생
  - 1) context vector의 사이즈가 정해져 있고 모든 정보를 저장하고 있기 때문에 정보손실 발생
  - 2) context vector로 부터 너무 멀리 있는 input 데이터는 미약시 Vanishing gradient, 즉 Long-term Problem이 존재한다.



## Attention

- 기본 아이디어는 디코더에서 출력 단어를 예측하는 매 시점 (step)마다, 인코더의 입력 시퀀스를 다시 참고하는 것이다.
- 위에서 참고할때 동일한 비중으로 참고하는것이 아닌, 예측단어와 관련있는 입력 단어를 더욱 집중해서 본다.
- 결과: Long-term Problem문제를 해결하였다.

### 어텐션만 표현



### 적용하면

