

Chapter 07. 맥락을 파악하는 Attention 기법

STEP1. Attention 신경망

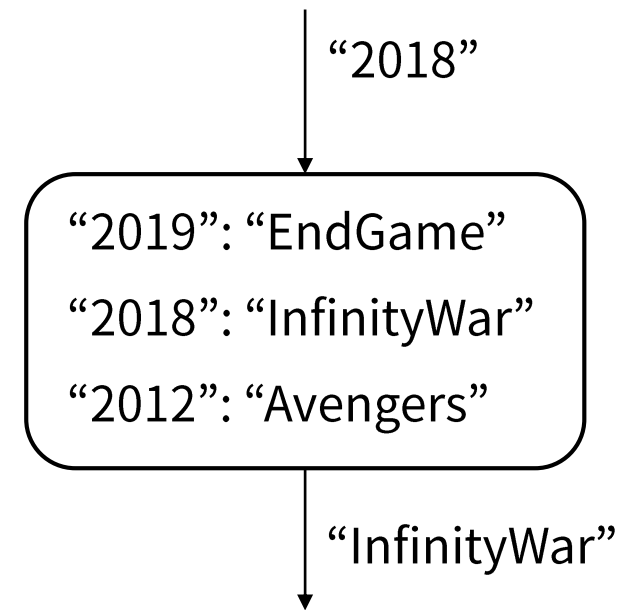
Query, Key-Value

Query : 질의. 찾고자 하는 대상

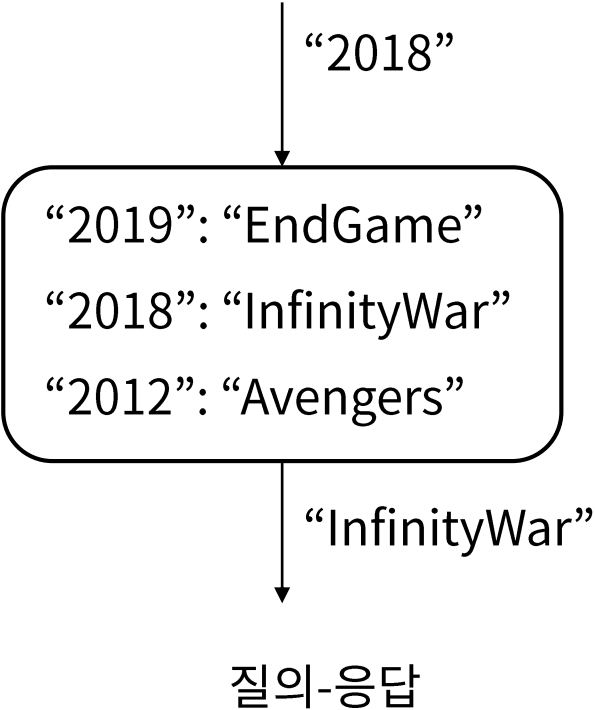
Key : 키. 저장된 데이터를 찾고자 할 때 참조하는 값

Value : 값. 저장되는 데이터

Dictionary : Key-Value Pair로 이루어진 집합



Querying



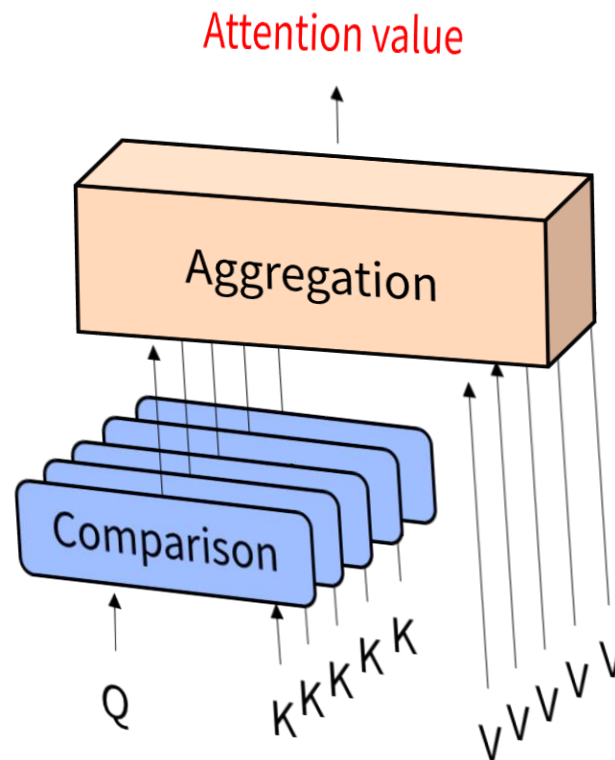
	"2018"
"2019"	0
"2018"	1
"2012"	0

비교

- 0 "2019": "EndGame"
- 1 "2018": "InfinityWar"
- 0 "2012": "Avengers"

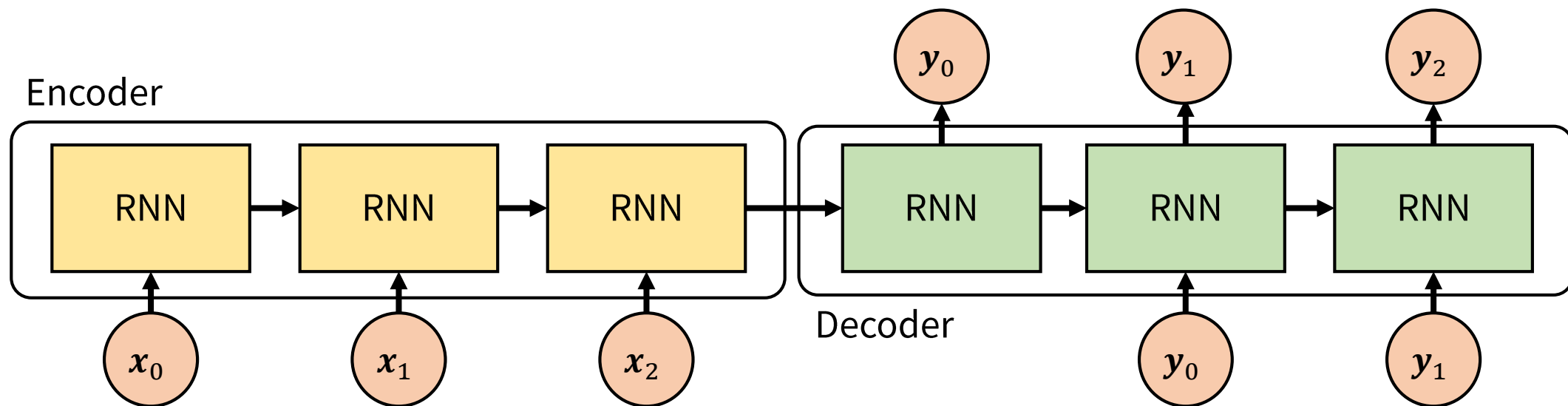
값 출력

Attention mechanism



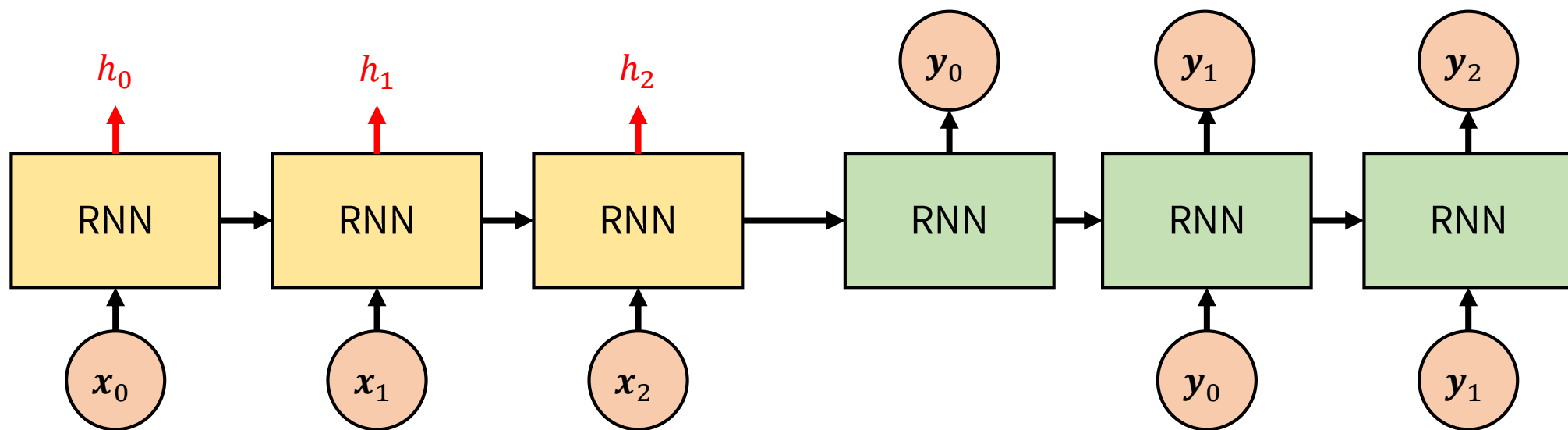
Q 에 대해 어떤 K 가 유사한지 비교하고, 유사도를 반영하여 V 들을 합성한 것이 Attention value이다.

Seq2seq



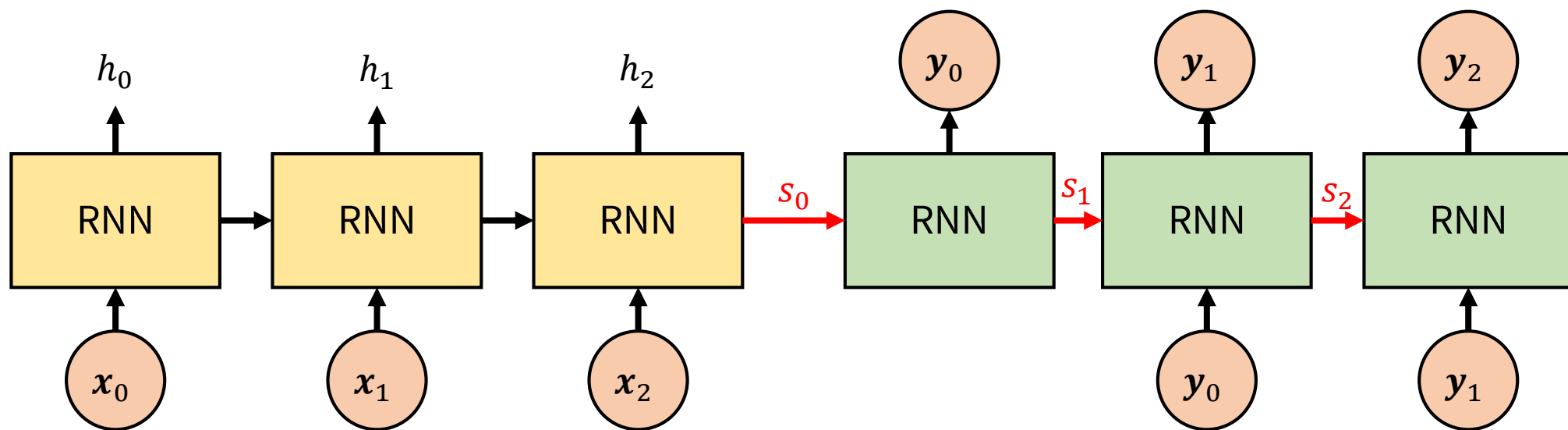
Attention mechanism이 Seq2seq 모델에 어떻게 적용 가능한지 알아보자.

Seq2seq – Key-Value



대부분의 attention network에서는 key와 value를 같은 값을 사용한다.
Seq2seq에서는 **Encoder의 hidden layer들을 key와 value로 사용한다.**

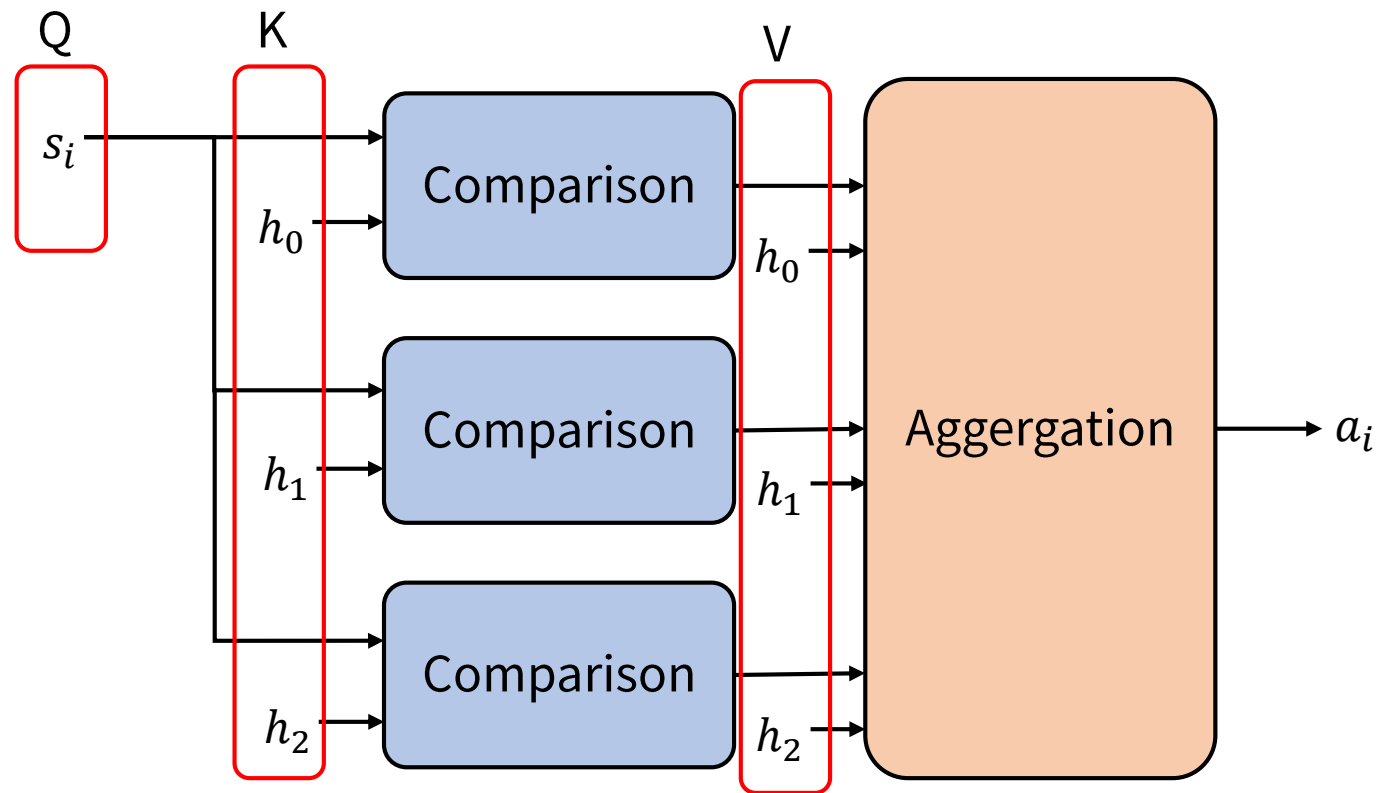
Seq2seq – Query



Seq2seq에서는 **Decoder의 hidden layer들을 Query로 사용**한다.

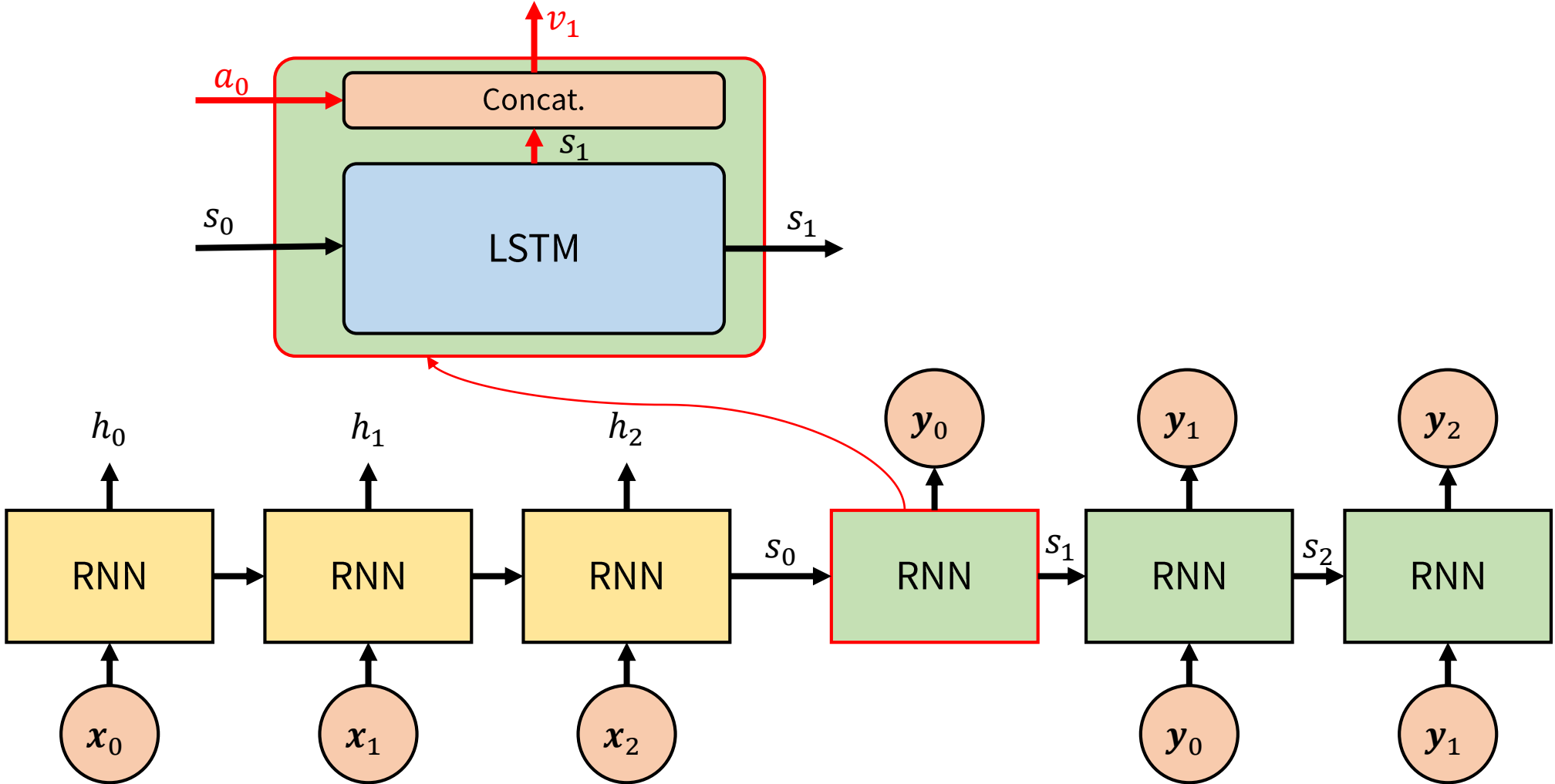
주의할 점은, Encoder와 달리 **하나 앞선 time-step**의 hidden layer를 사용한다는 점.

Seq2seq – Attention mechanism



i 번째 decoder에 대해서 a_i 의 attention value를 얻는다.

Seq2seq – Application



RNN으로 Hidden state를 입력하기 전에, **attention value**를 **concatenate**하여 **입력**한다.