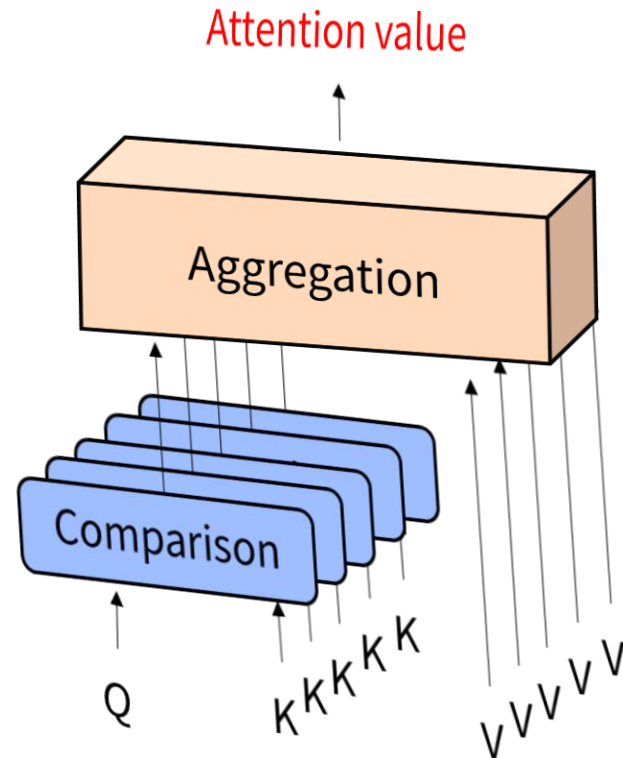


Chapter 07. 맥락을 파악하는 Attention 기법

# STEP2.

# Attention 신경망의 수식적 이해

# Attention mechanism



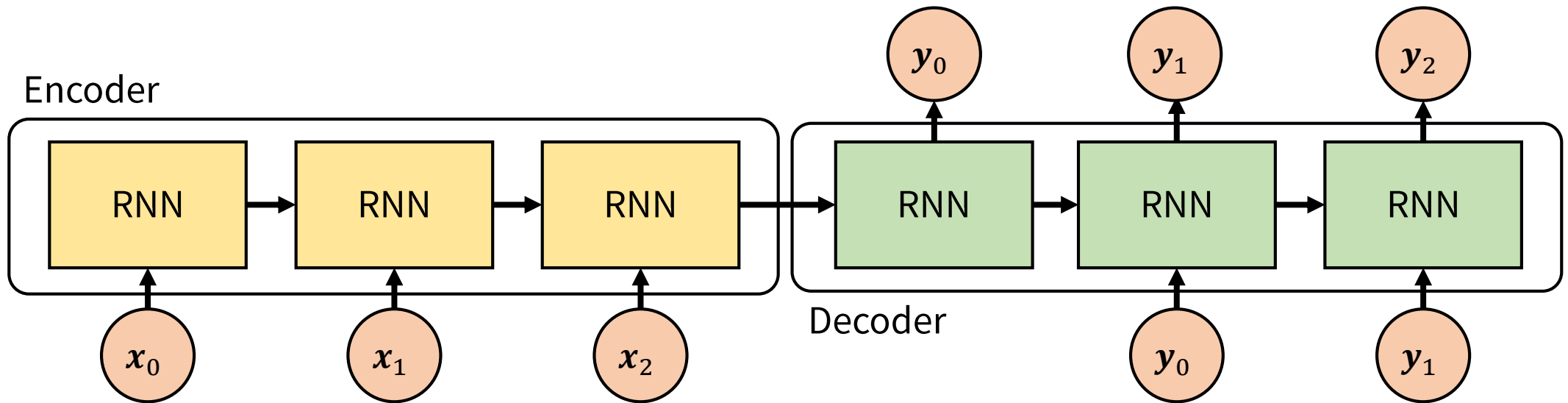
$$\mathbf{q} \in \mathbb{R}^n, \mathbf{k}_j \in \mathbb{R}^n$$

$$\text{Compare}(\mathbf{q}, \mathbf{k}_j) = \mathbf{q} \cdot \mathbf{k}_j = \mathbf{q}^T \mathbf{k}_j$$

$$\text{Aggregate}(\mathbf{c}, V) = \sum_j c_j \mathbf{v}_j$$

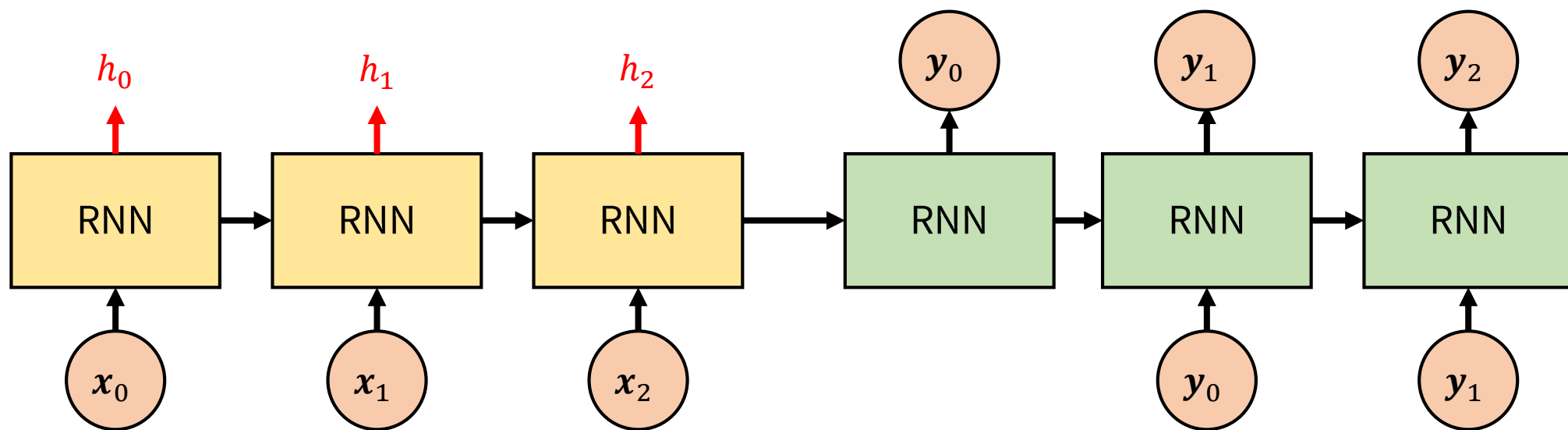
Compare 함수로는 Dot-Product (Inner Product)가 많이 쓰이며,  
Aggregation은 weighted sum을 많이 사용한다.

# Seq2seq



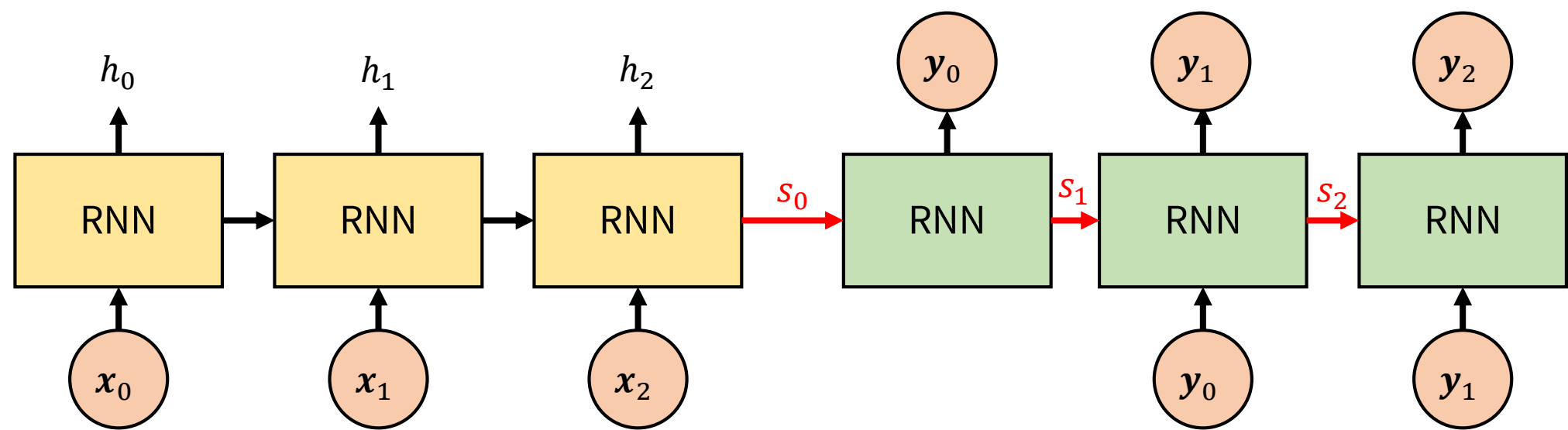
Seq2seq 모델을 다시 간단히 살펴 보면서 수식을 적용해 보자.

# Seq2seq – Key-Value



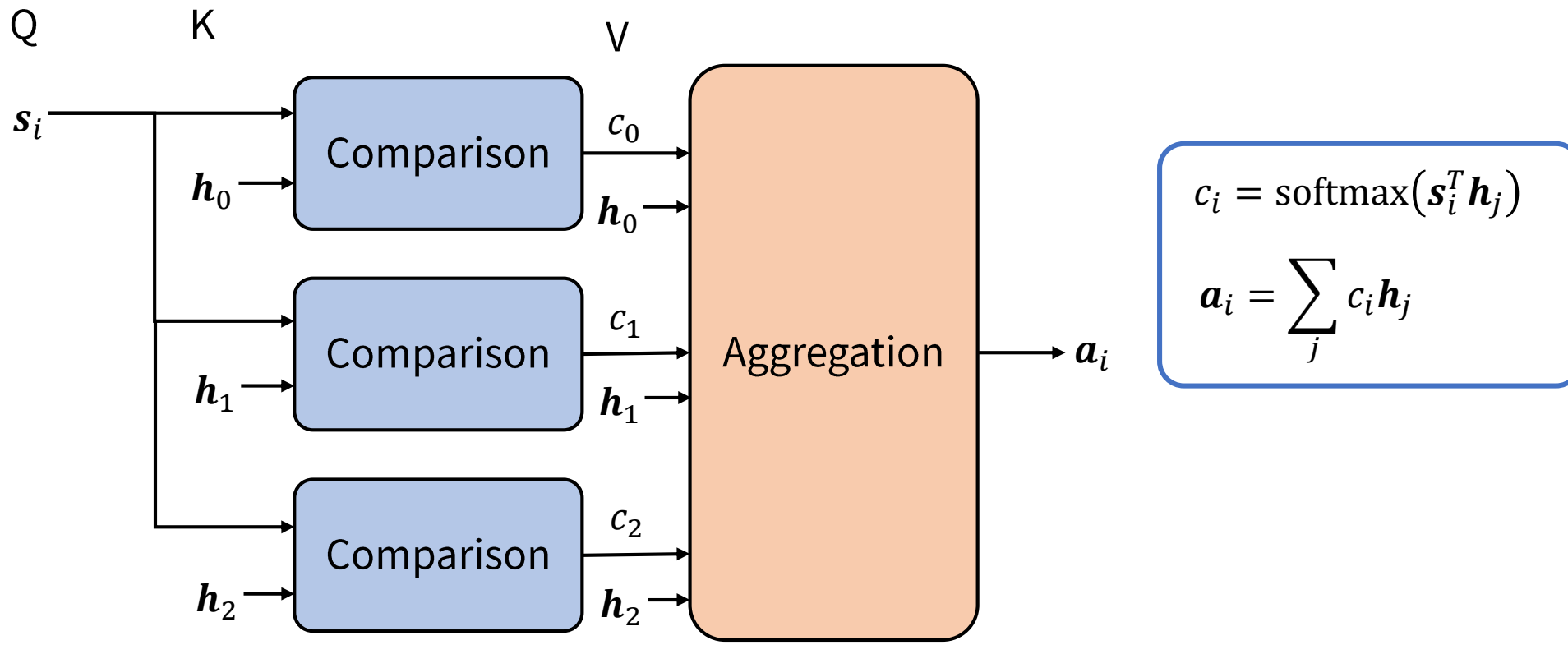
Seq2seq에서는 Encoder의 hidden layer들을 key와 value로 사용한다.

# Seq2seq – Query



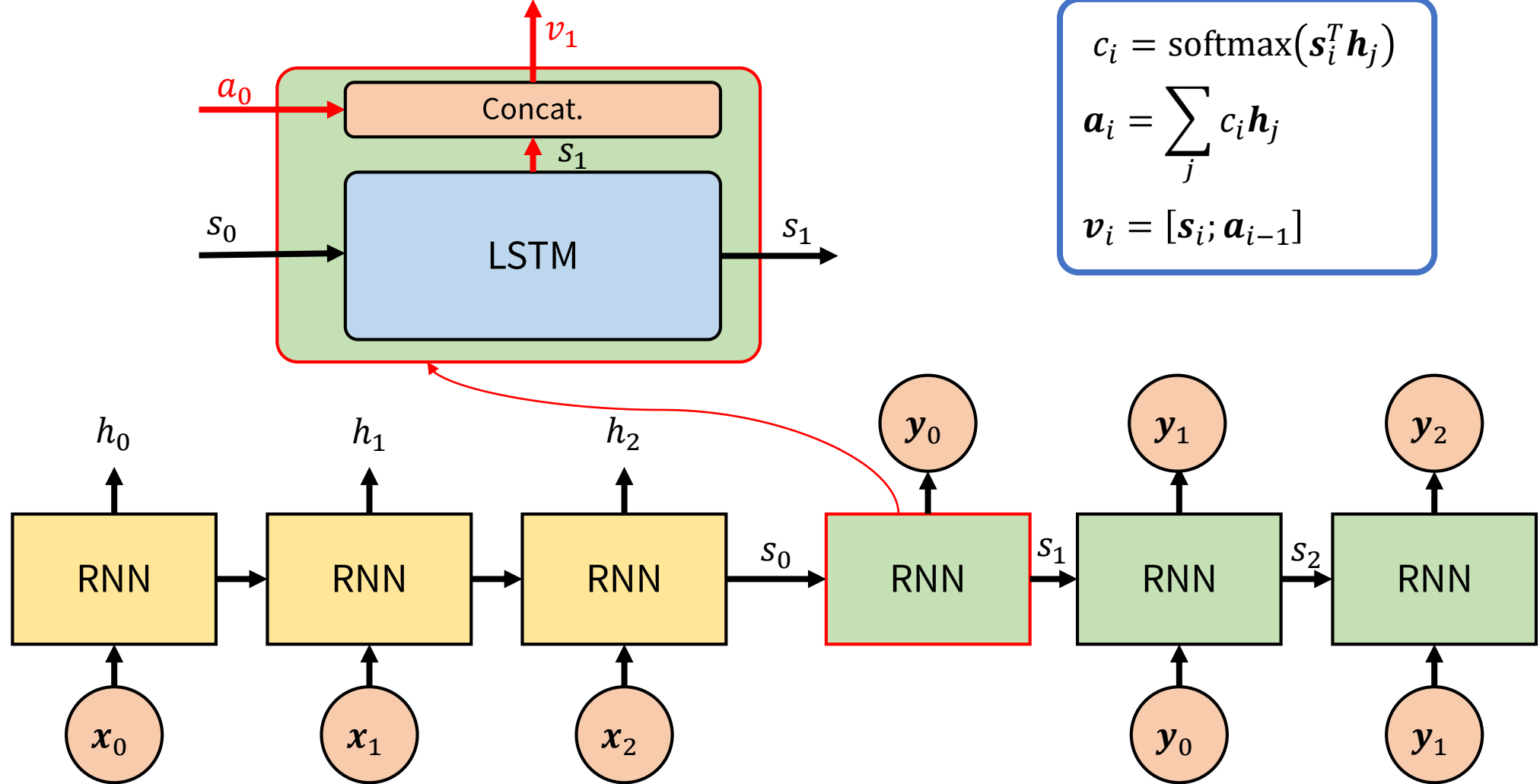
Seq2seq에서는 **Decoder의 hidden layer들을 Query로 사용한다.**

# Seq2seq – Attention mechanism



블록도에 비해 수식이 오히려 간단하다. 비교 함수와 결합 함수의 의미를 잘 이해하자.

# Seq2seq – Application



Hidden state에 attention value를 concatenate까지 하면 모든 수식적 표현이 끝난다.