



seq2seq

Sang Yup Lee

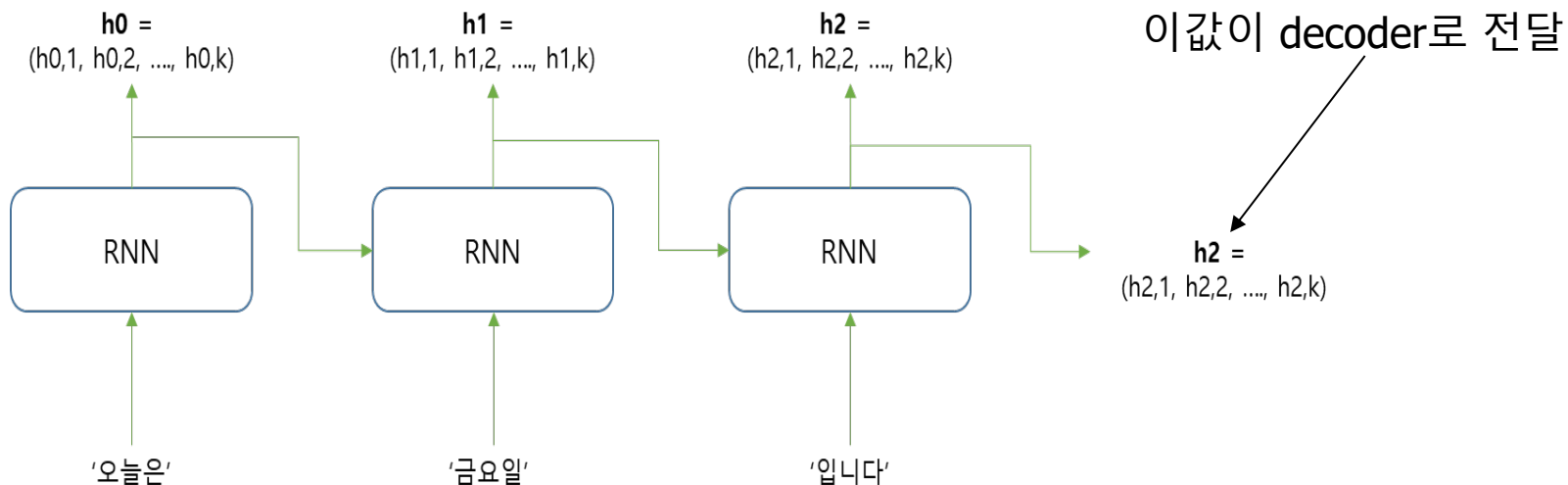


seq2seq

- seq2seq (sequence to sequence)
 - 주요 목적
 - To convert sequences from one domain (e.g. sentences in Korean) to sequences in another domain (e.g. the same sentences translated to English), 번역이 대표적인 예
 - 여기서는 번역을 예로 설명
 - Encoder-decoder 구조
 - Encoder의 역할
 - 입력된 텍스트 데이터를 숫자 형태로 혹은 벡터 형태로 변환
 - Decoder의 역할
 - Encoder에 의해 숫자로 변경된 정보를 다른 형태의 텍스트 데이터로 변환
 - Encoder와 decoder 를 위해 순환신경망 기반 모형 사용 가능 (예, RNN)
 - 첫번째 RNN이 encoder 역할, 두번째 RNN이 decoder 역할

seq2seq

- 예: '오늘은 금요일입니다' \Rightarrow 'Today is Friday'
 - Encoder의 역할: '오늘은 금요일입니다' 를 벡터형태로 변환



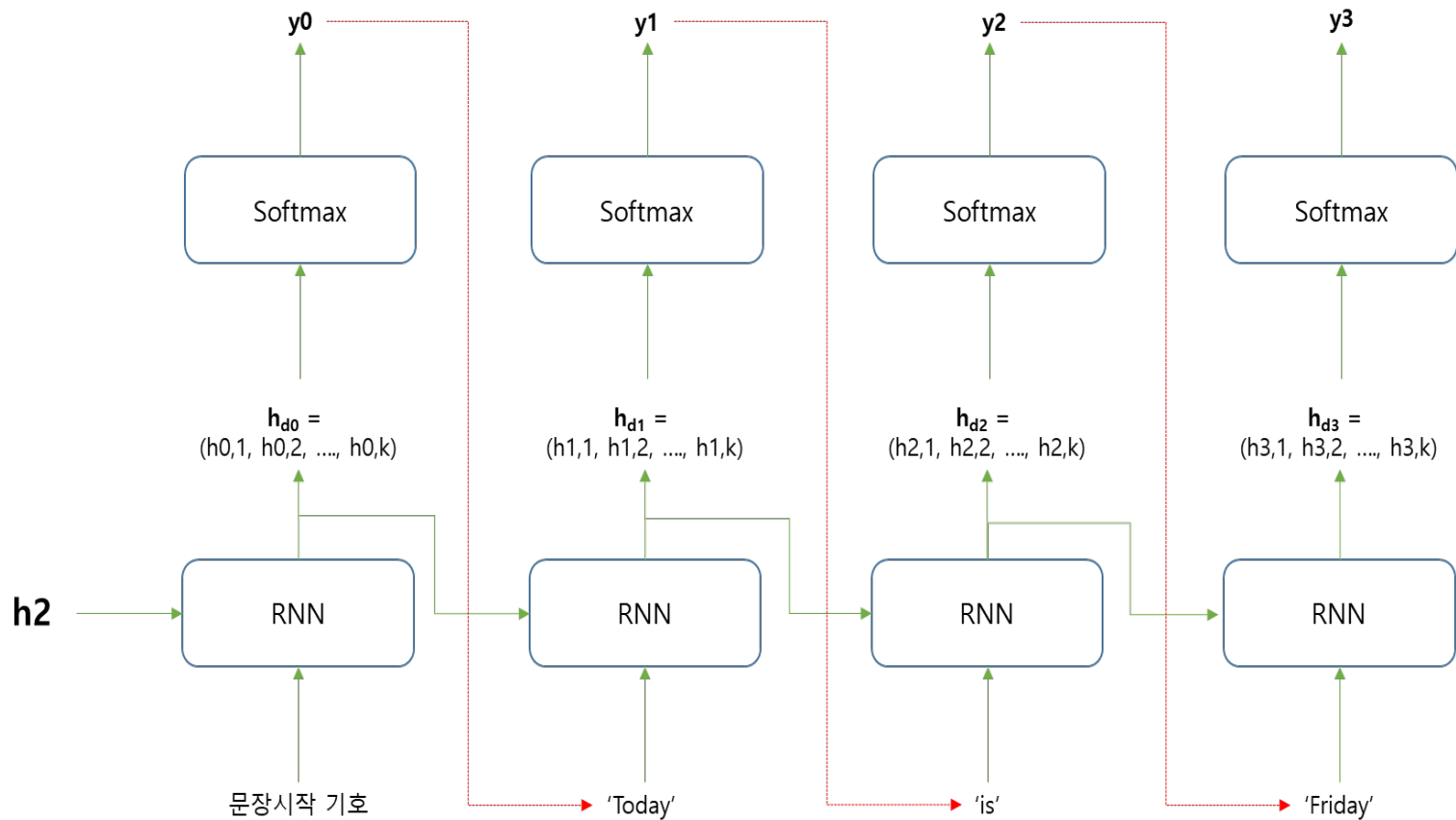


seq2seq

- Decoder의 역할
 - Encoder가 전달하는 벡터 정보를 받아서 영어 문장으로 변환
 - 이러한 번역을 위해서 decoder에 해당하는 RNN은 언어 모델(language model)로 사용
 - 즉, 주어진 단어들 정보를 사용하여 다음 단어를 예측
 - 새로운 도메인에서의 시퀀스 데이터 생성 (즉, 영어 문장 생성)

seq2seq

■ Decoder 구조





seq2seq

- Decoder 구조 (cont'd)
 - 출력층의 구조 in detail
 - If the size of the vocabulary (# of words in the corpus) = N , then the number of output nodes = N
 - Each output node returns the probability of each word in the corpus is the next appearing word