

N423

• 언어 모델 (Language Model)

- 통계 기반 언어모델을 이해하고 설명할 수 있습니다.
- 통계 기반 언어모델의 한계를 이해하고 이를 극복하기 위해 등장한 신경망 언어 모델의 장점을 설명할 수 있습니다.

• 순환 신경망 (Recurrent Neural Network, RNN)

- RNN의 구조와 작동 방식을 이해하고 설명할 수 있습니다.
- RNN의 장점과 단점을 설명하고 이해할 수 있습니다.

• LSTM & GRU

- LSTM과 GRU가 고안된 배경과 구조를 연관지어 설명할 수 있습니다.
- 두 방법의 차이에 대해서 설명할 수 있습니다.

• Attention

- Attention이 탄생하게 된 배경에 대해서 설명할 수 있습니다.
- Attention의 장점에 대해서 설명하고 Attention 으로도 해결할 수 없는 RNN의 구조적 단점에 대해서도 이해할 수 있습니다.

* 언어 모델

단어 시퀀스에서 각 단어의 확률을 계산하는 모델 < ① 통계적 언어 모델 (전통적 방식)
② 신경망 언어 모델

② 신경망 언어 모델

- 희소기반 대신 word2vec 이나 fasttext 사용

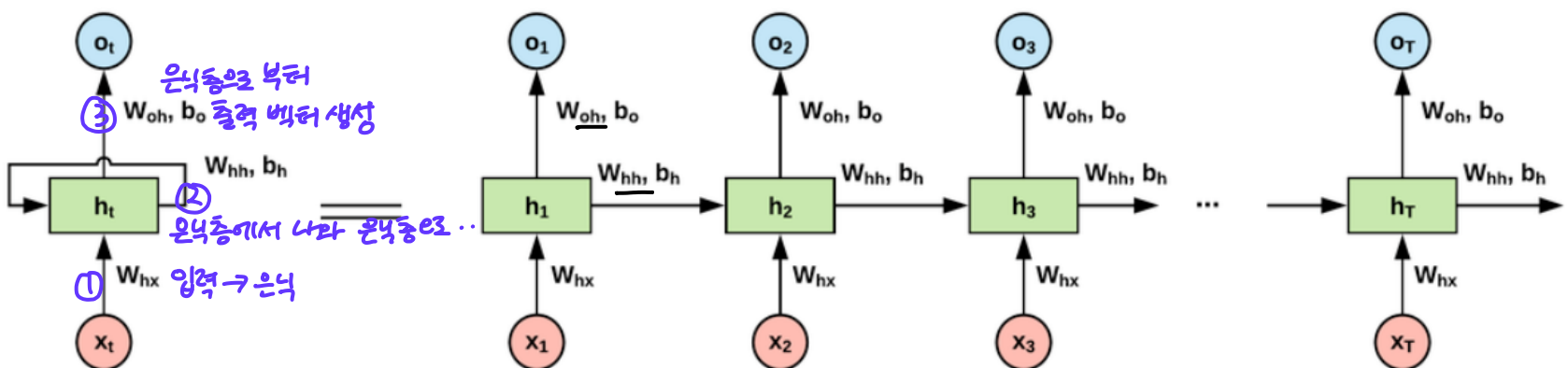
∴ 말뭉치에 등장 x 라도, 의미적 & 문법적 유사하면 선택 가능

* RNN (순환 신경망)

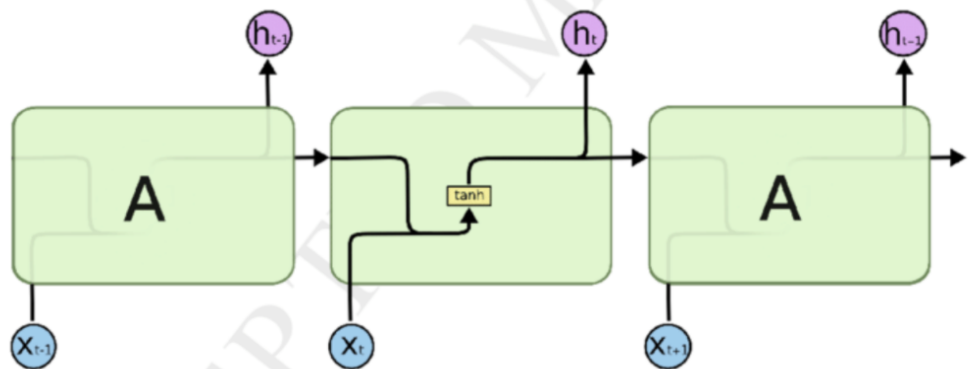
RNN이란? 연속형 데이터 처리를 위해 고안된 신경망 구조

Sequential data : 주식가격, 자연어 등... 어떤 순서로 오는지 따라 의미 달라짐

Non-Sequential data : 이미지 등...



$$h_t = \tanh \left(\underbrace{h_{t-1} W_h}_{\text{전역 가중치}} + \underbrace{x_t W_x}_{\text{현재 가중치}} + \underbrace{b}_{\text{bias}} \right)$$



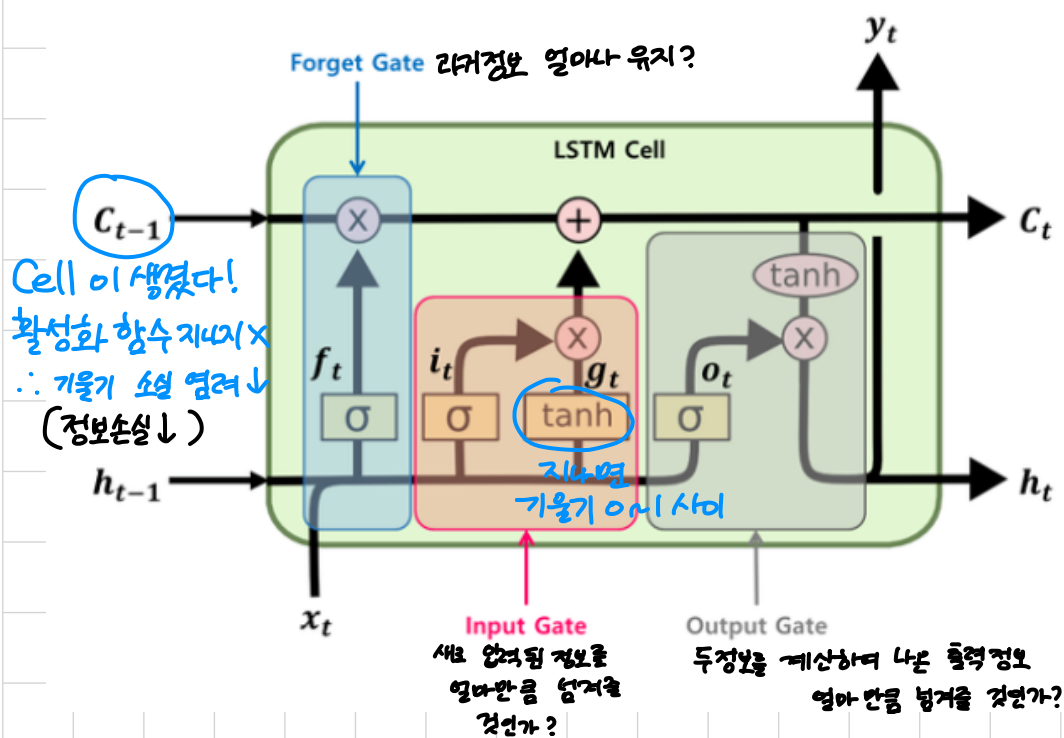
RNN 특징

- 모델 간단, 어떤길이의 Sequential 데이터 처리 가능

-(단점) 병렬화 (Parallelization) 불가,

가울기 폭발 (Exploding Gradient)
가울기 소멸 (Vanishing Gradient)) 역전파과정에서
→ 장기 의존성 문제

* LSTM

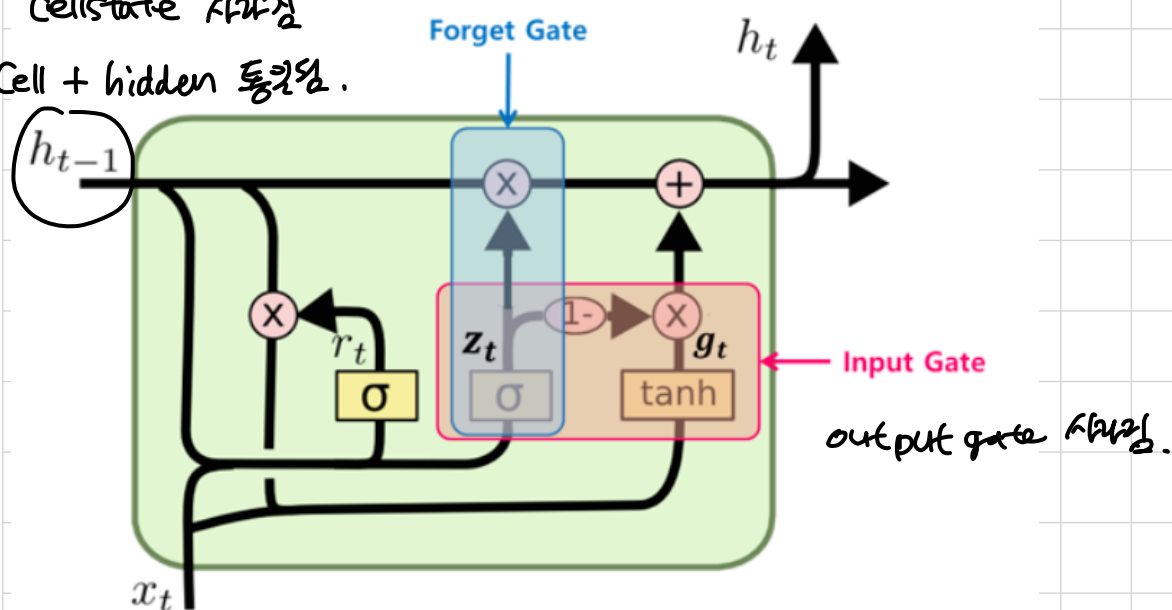


* GRU

LSTM의 간소화 버전

Cellstate 사라짐

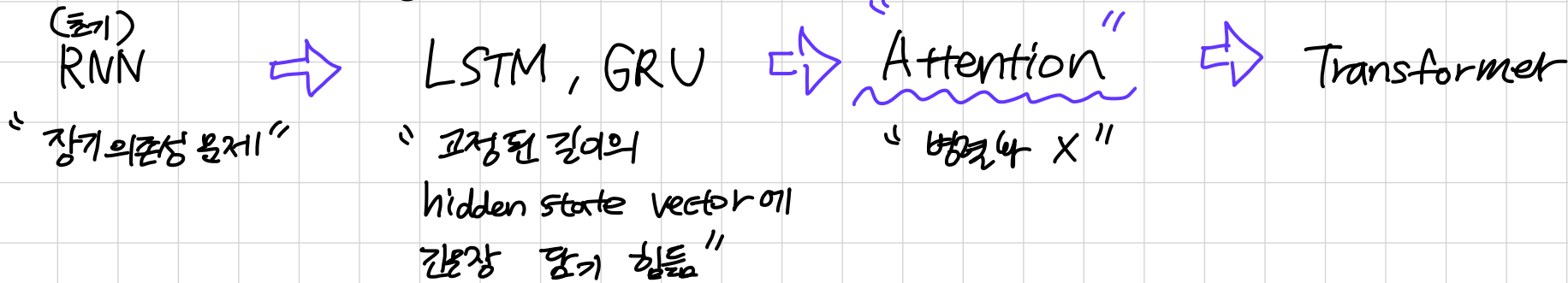
Cell + hidden 통칭함.



* Attention

RNN 구조에 Attention 적용하기

Attention의 등장.

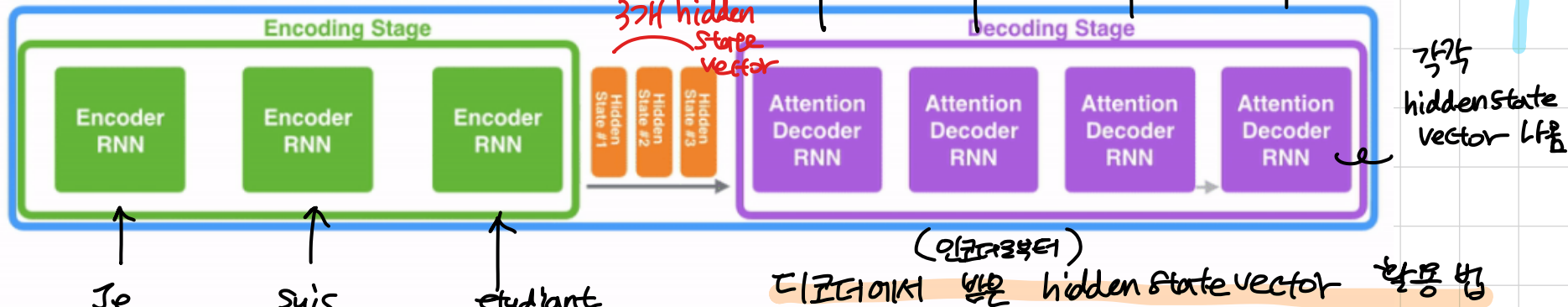


Attention 특징

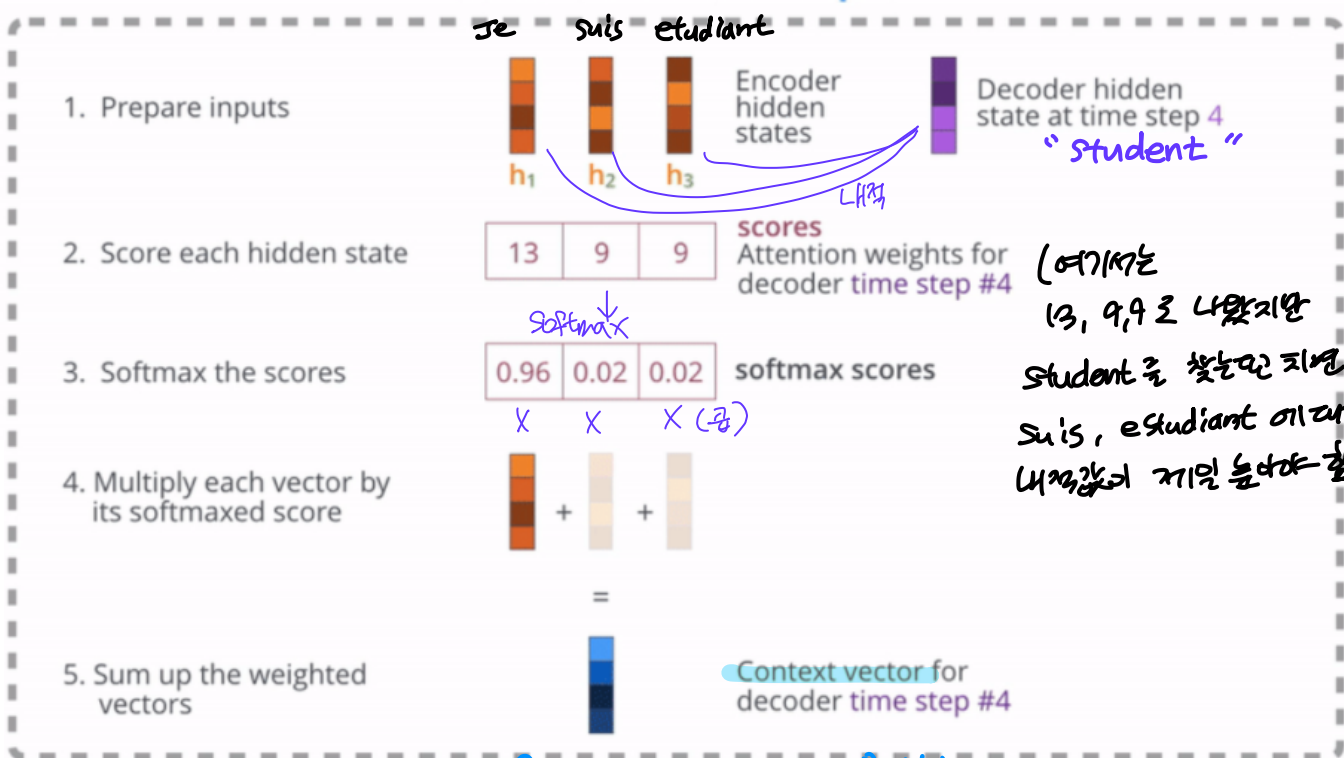
각 인코더의 Time-step 마다 생성되는 hidden-state vector 각각 입력단어 N개 \sim N개의 hiddenstatevector \rightarrow 디코더에 넘겨줌

Neural Machine Translation

SEQUENCE TO SEQUENCE MODEL WITH ATTENTION

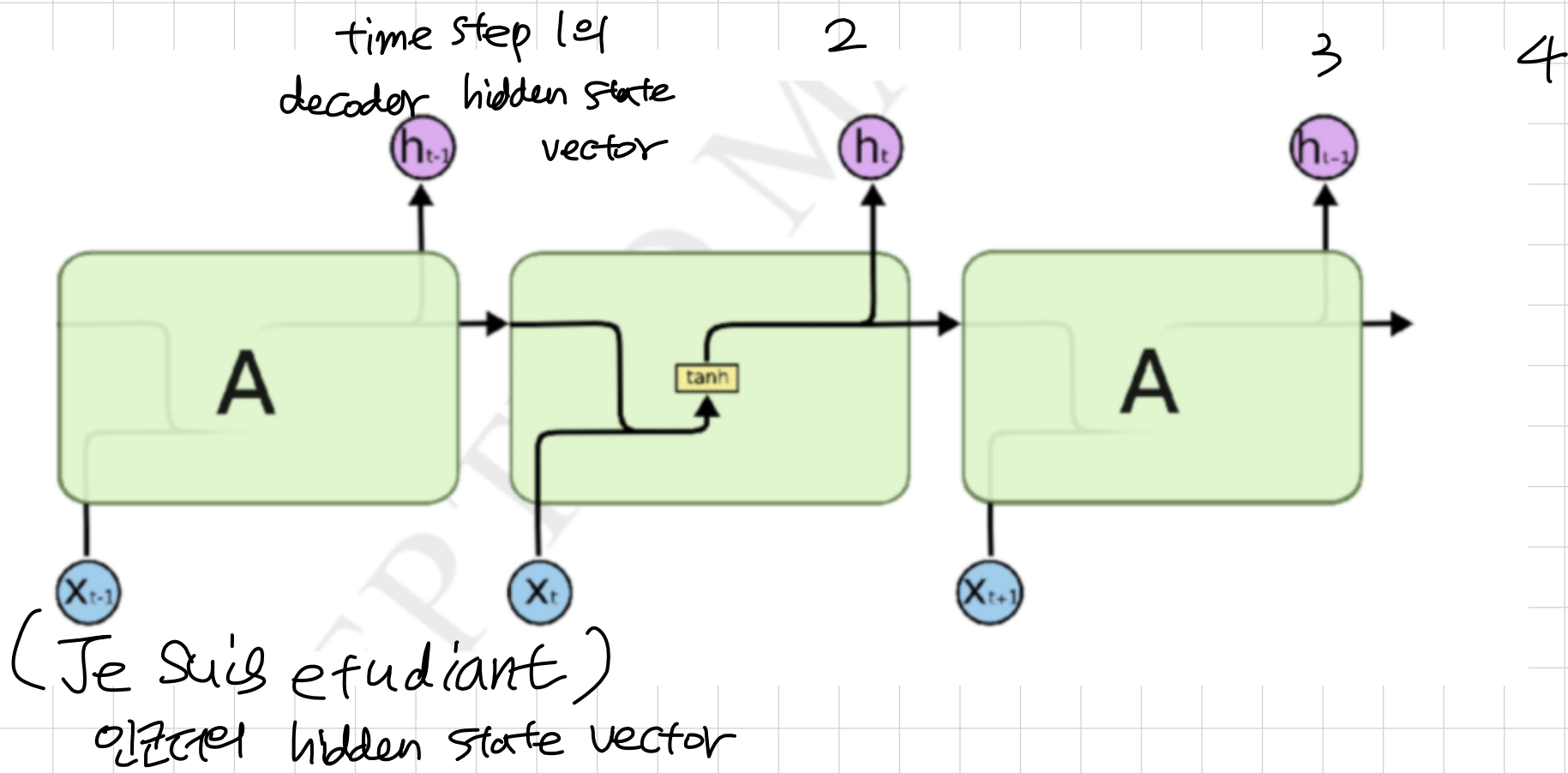


Attention at time step 4

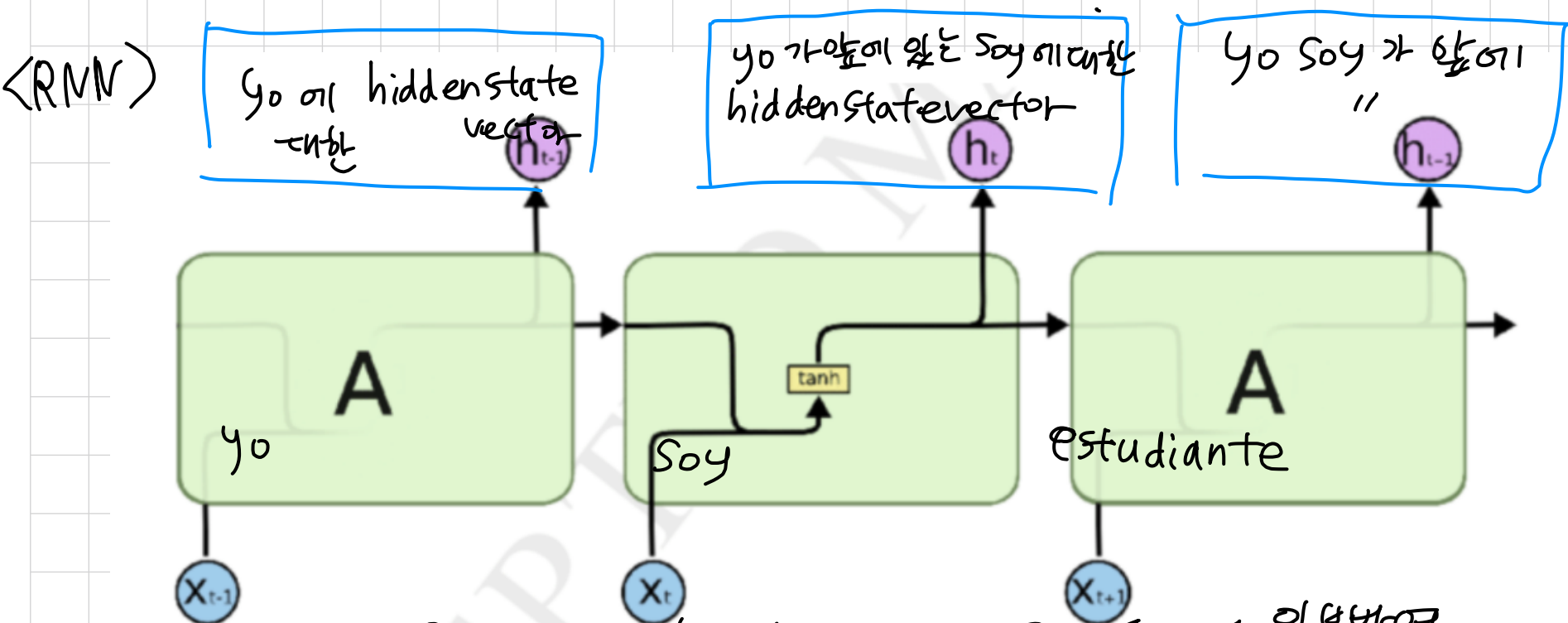


6. * Context vector 와 디코더 hidden state vector로 최종 출력단어 결정

→ 디코더의 RNN 이라는...?

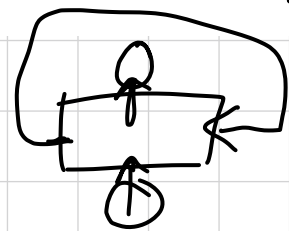


인코더의 RNN 이라는...? (yo soy estudiante → Je suis etudiant)



* 특히 각 수를 이전 입력값들이 다들지 되어서 일부분만.

그래서 sequential data 처리 가능.



이전. 이동하기는 가능!