

# Sequence to Sequence: Teacher Forcing

Ki Hyun Kim

[nlp.with.deep.learning@gmail.com](mailto:nlp.with.deep.learning@gmail.com)

# Review: Auto-regressive

- Inference

$$\hat{x}_t = \operatorname{argmax}_{x_t \in \mathcal{X}} \log P(x_t | \hat{x}_{<t}; \theta)$$

- Auto-regressive:

- 과거 자신의 상태를 참조하여 현재 자신의 상태를 업데이트.

$$\hat{x}_{t=1} = \operatorname{argmax}_{x_t \in \mathcal{X}} \log P(x_{t=1} | x_0; \theta) \text{ where } x_0 = \langle \text{BOS} \rangle.$$

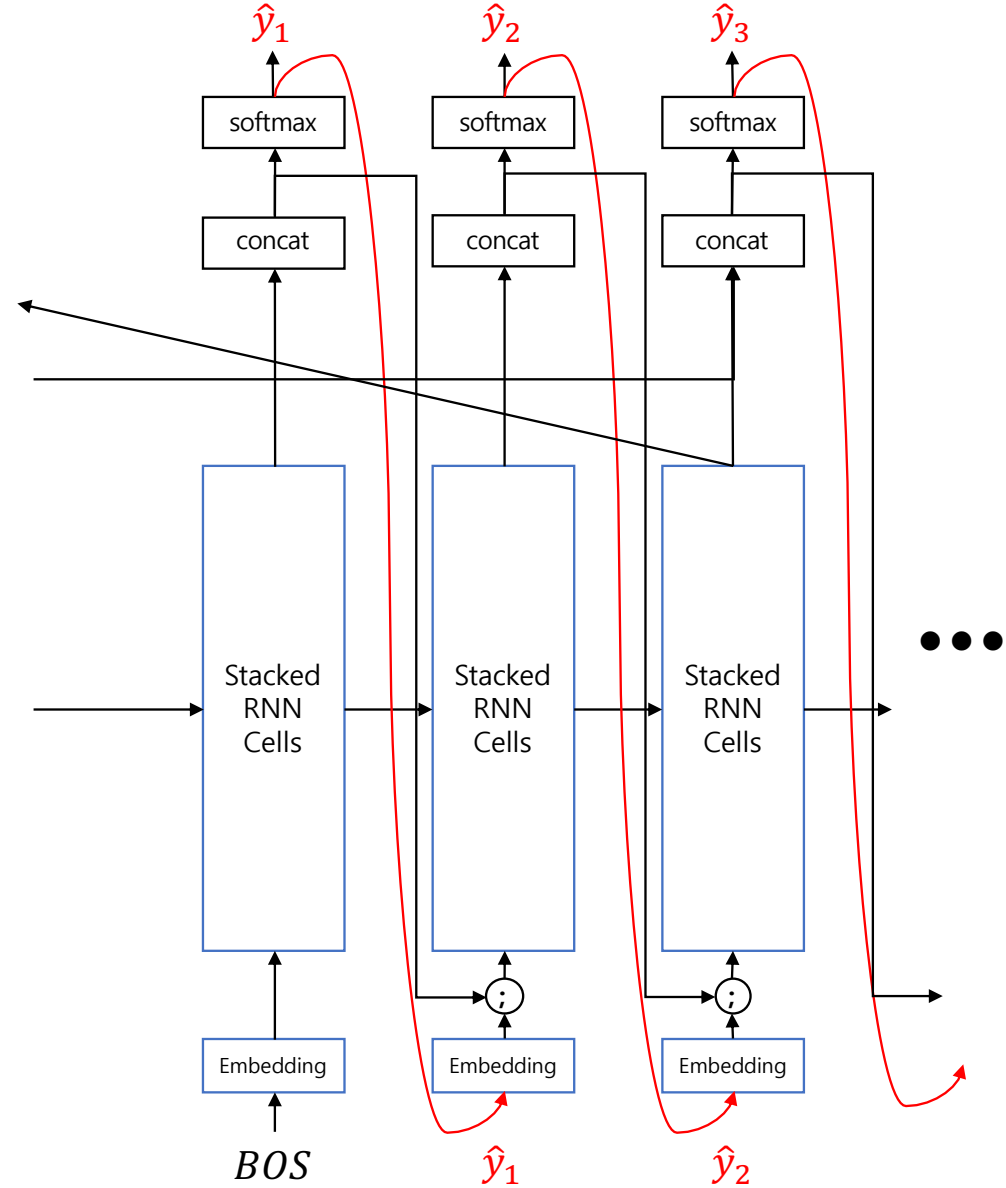
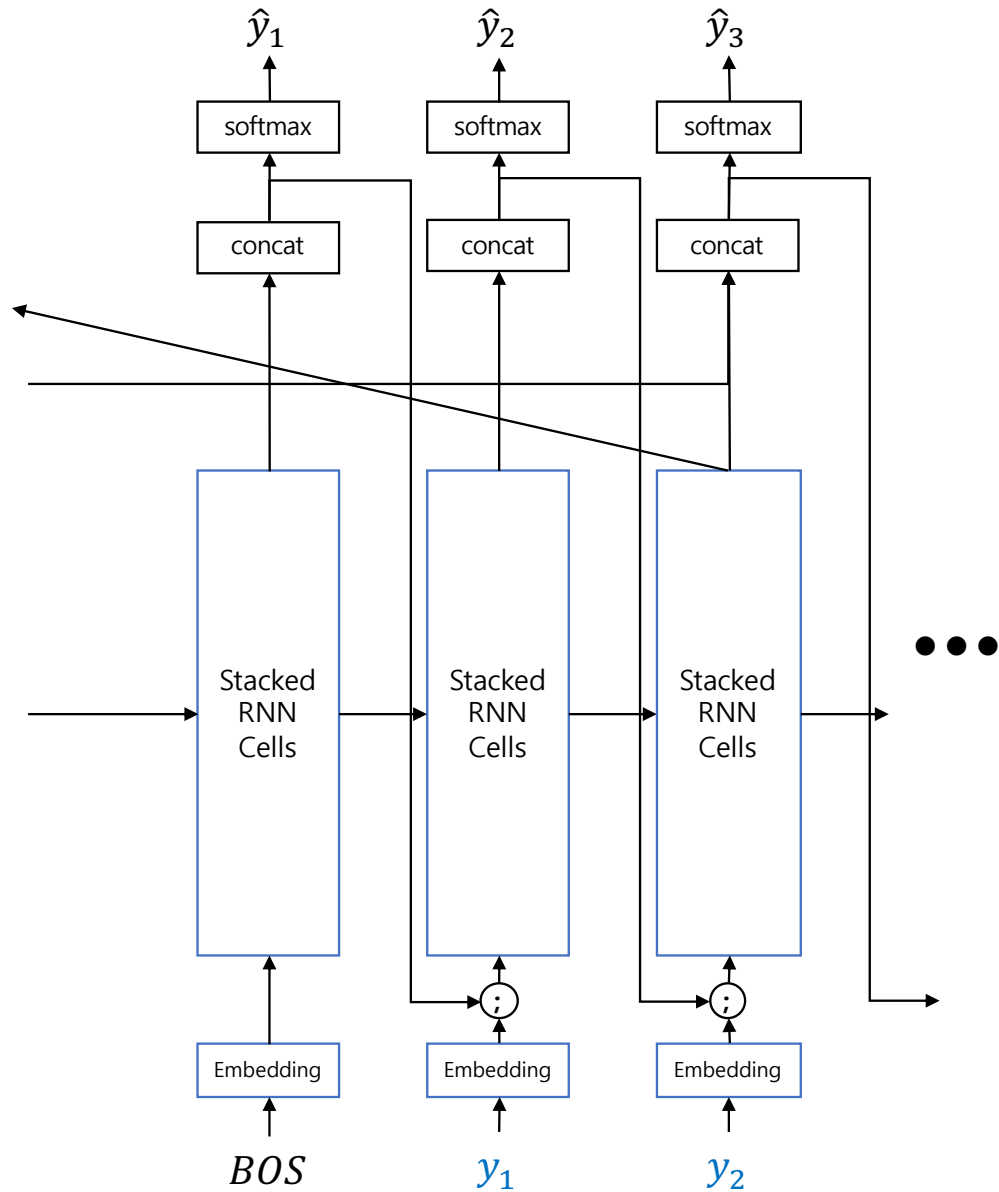
$$\hat{x}_{t=2} = \operatorname{argmax}_{x_t \in \mathcal{X}} \log P(x_{t=2} | x_0, \hat{x}_1; \theta)$$

$$\hat{x}_{t=3} = \operatorname{argmax}_{x_t \in \mathcal{X}} \log P(x_{t=3} | x_0, \hat{x}_1, \hat{x}_2; \theta)$$

...

$$\hat{x}_t = \operatorname{argmax}_{x_t \in \mathcal{X}} \log P(x_t | x_0, \hat{x}_{<t}; \theta)$$

# Training vs Inference



# Summary

- Auto-regressive task를 feed-forward할 때는  
보통 이전 time-step의 출력이 현재 time-step의 입력이 됨
- Teacher Forcing을 통해 Auto-regressive task에 대한  
sequential modeling을 할 수 있음
  - 하지만 training mode와 inference mode의 괴리(discrepancy)가 생김
- RL을 통해 이러한 괴리를 없애고 성능을 높일 수 있음
  - 이외에도 다양한 방법(e.g. professor forcing)들이 제안됨