

순차적 구문 분석 방법을 반영한 포인터 네트워크 기반의 한국어 의존 구문 분석기

(한장훈, 박영준, 정영훈, 이인권, 한정욱, 박서준)

목차

Task

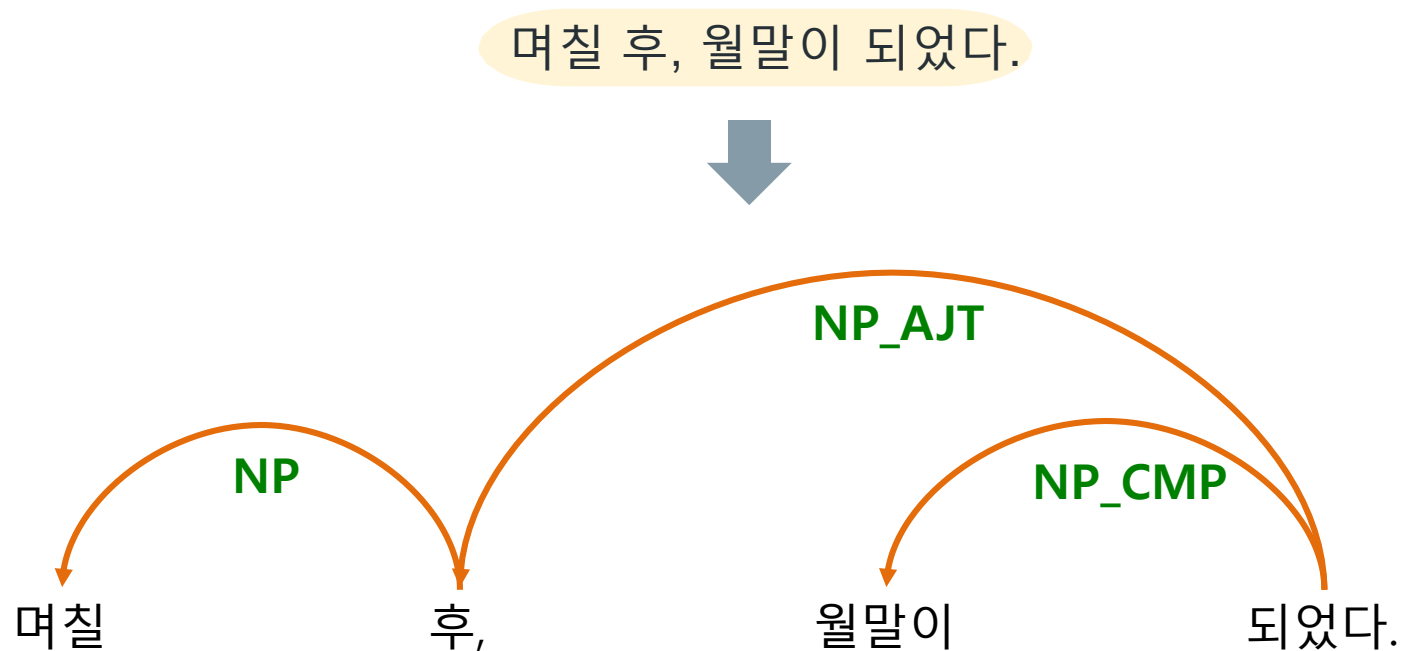
모델 설명

실험 결과 및 분석

결론

의존 구문 분석

의존 구문 분석이란 문장 성분 간의 관계를 파악함으로써
문장의 구조를 이해하는 작업이다.



모델



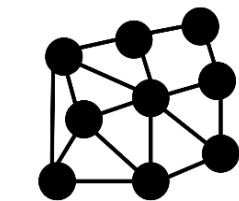
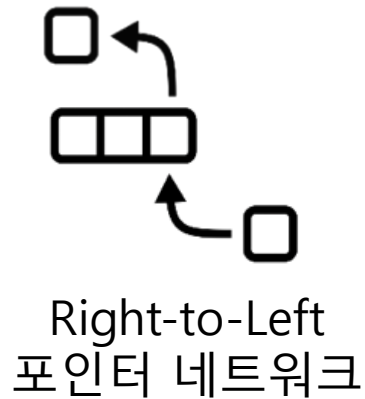
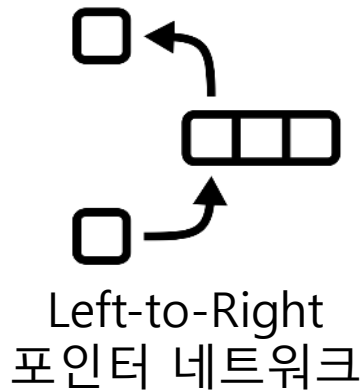
Embedding



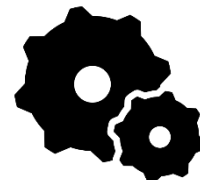
ELMo



BERT



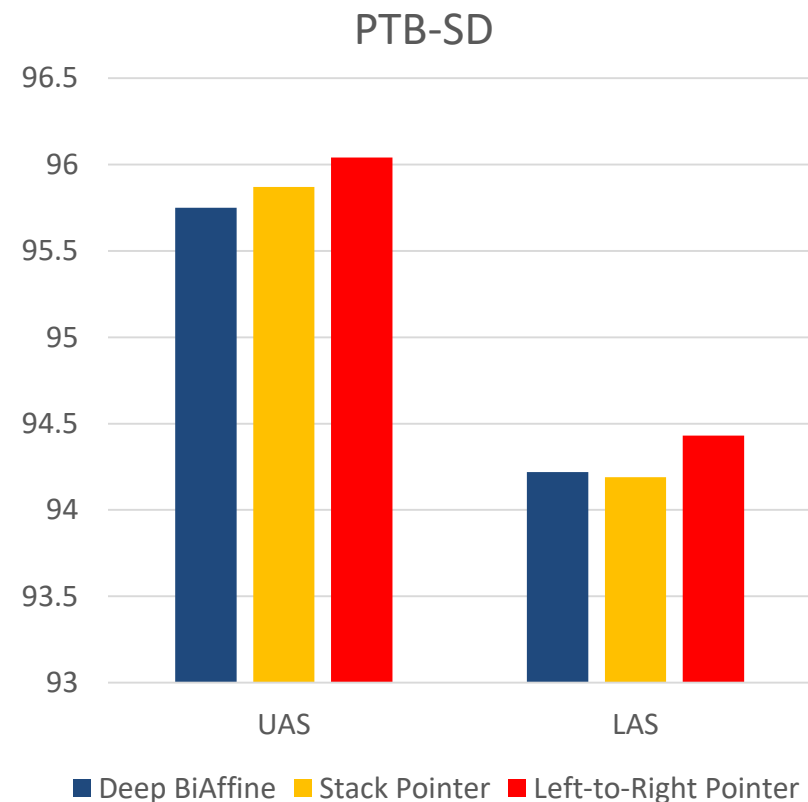
Deep Biaffine



앙상블

최종 결과

Left-to-Right 포인터 네트워크 선정 이유



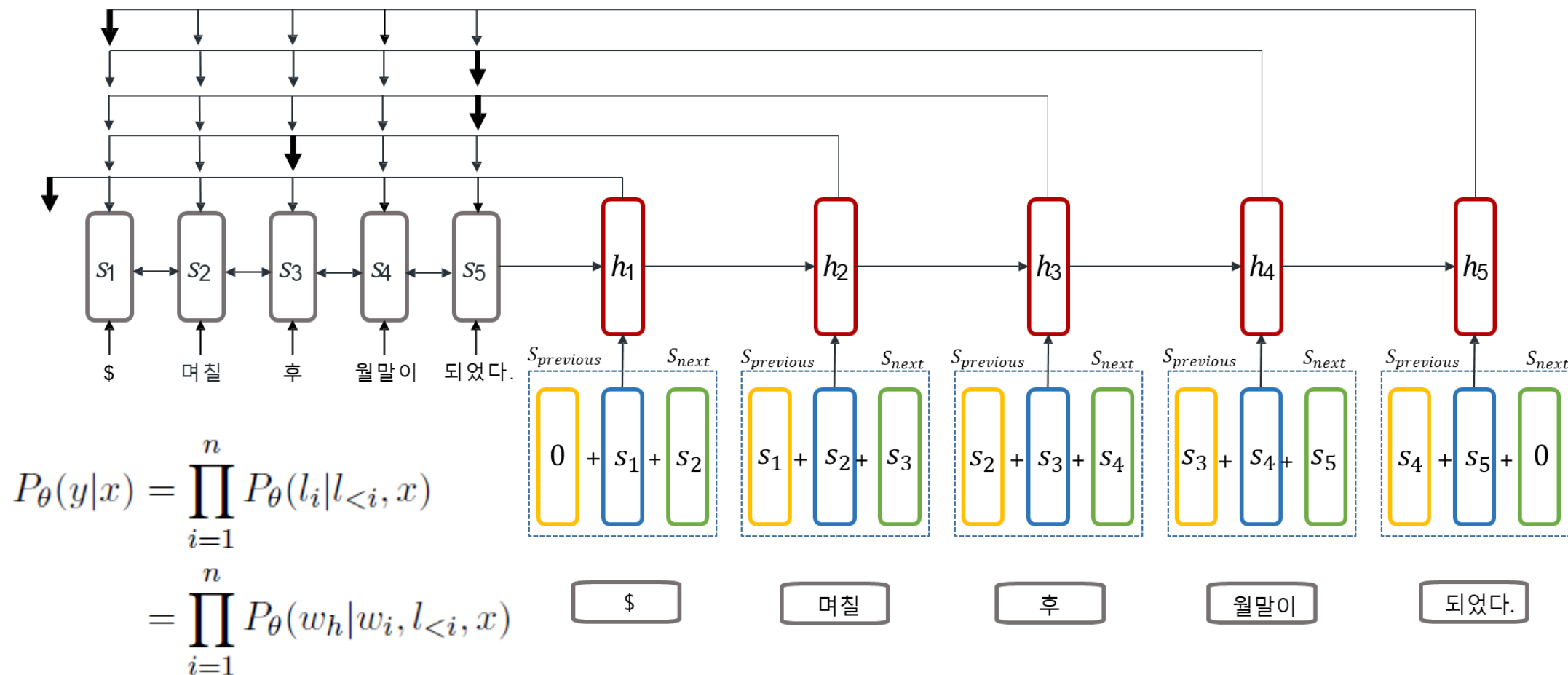
- ✓ BiAffine과 Stack Pointer보다 좋은 성능
- ✓ 스택포인터 보다 빠른 속도 ($2N-1 \rightarrow N$)

Timothy, Dozat, et al. "Deep BiAffine Attention for Neural Dependency Parsing" arXiv:1611.01734v3 [cs.CL] (2017).

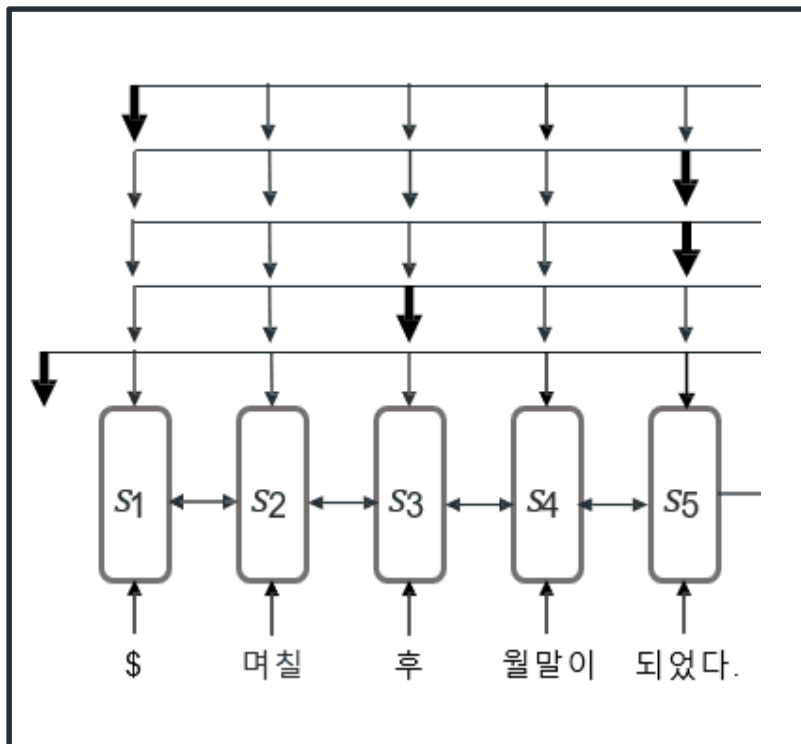
Ma, Xuezhe, et al. "Stack-Pointer Networks for Dependency Parsing." arXiv preprint arXiv:1805.01087 (2018).

Fernández-González, Daniel, et al. "Left-to-Right Dependency Parsing with Pointer Networks" arXiv preprint arXiv:1903.08445 (2019).

Left-to-Right 포인터 네트워크



Left-to-Right 포인터 네트워크 - 인코더



✓ Encoder

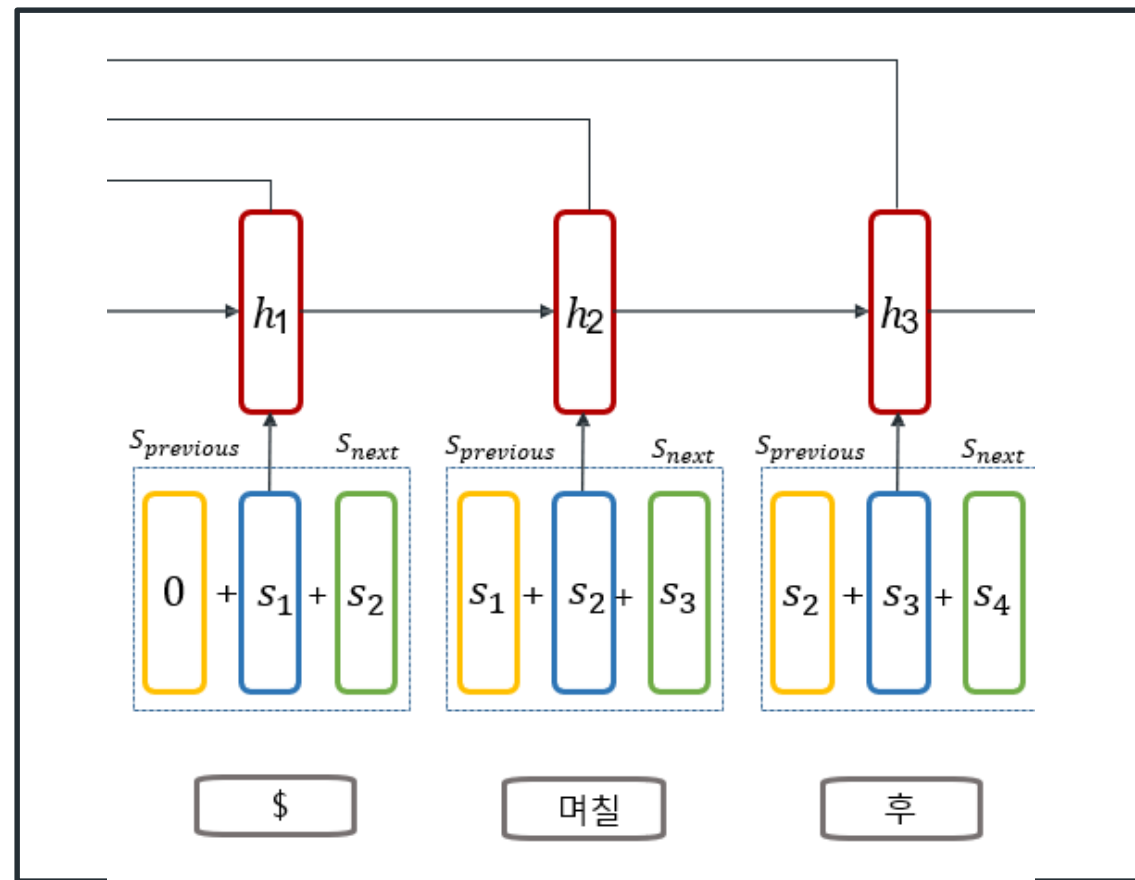
✓ BERT + ELMo 임베딩

✓ Bi-LSTM

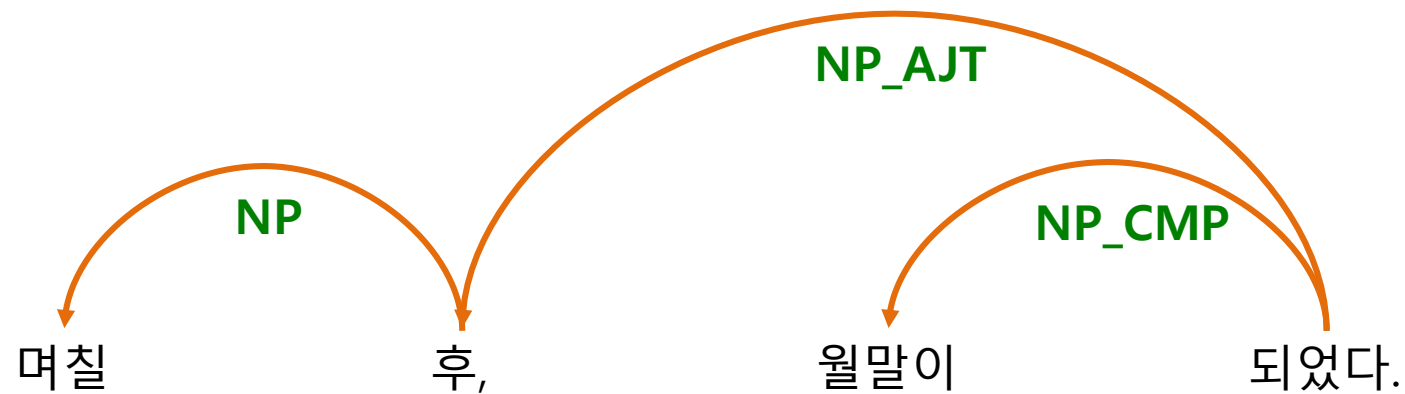
✓ Bayesian approximation을
이용한 dropout

Left-to-Right 포인터 네트워크 - 디코더

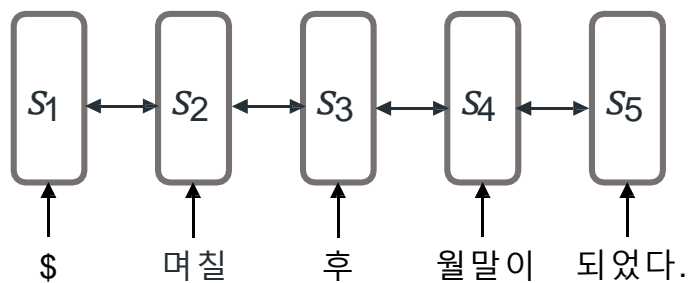
- ✓ Decoder
 - ✓ Uni-LSTM
 - ✓ 현재 어절에 대한 인코더 hidden state 사용
 - ✓ Extra Features로 previous, next 사용
 - ✓ 포인터 네트워크
 - ✓ Cycle이 있는지 체크



며칠 후, 월말이 되었다.



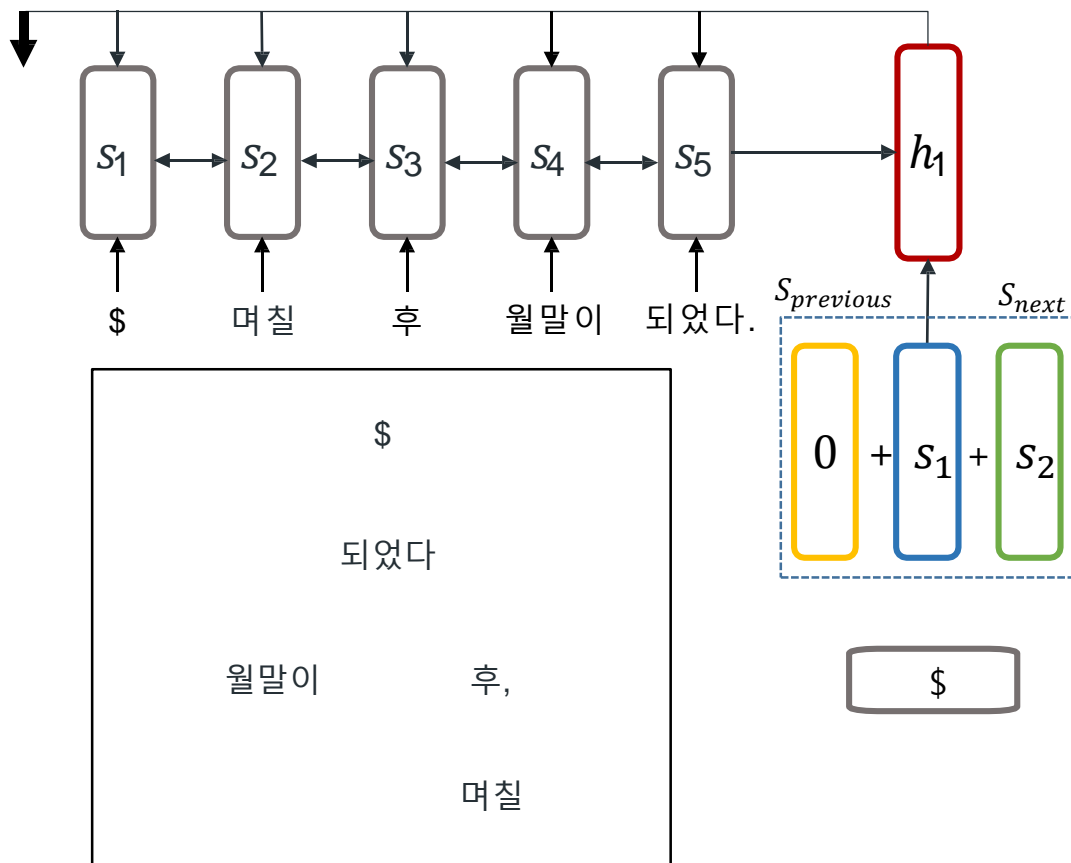
인코더

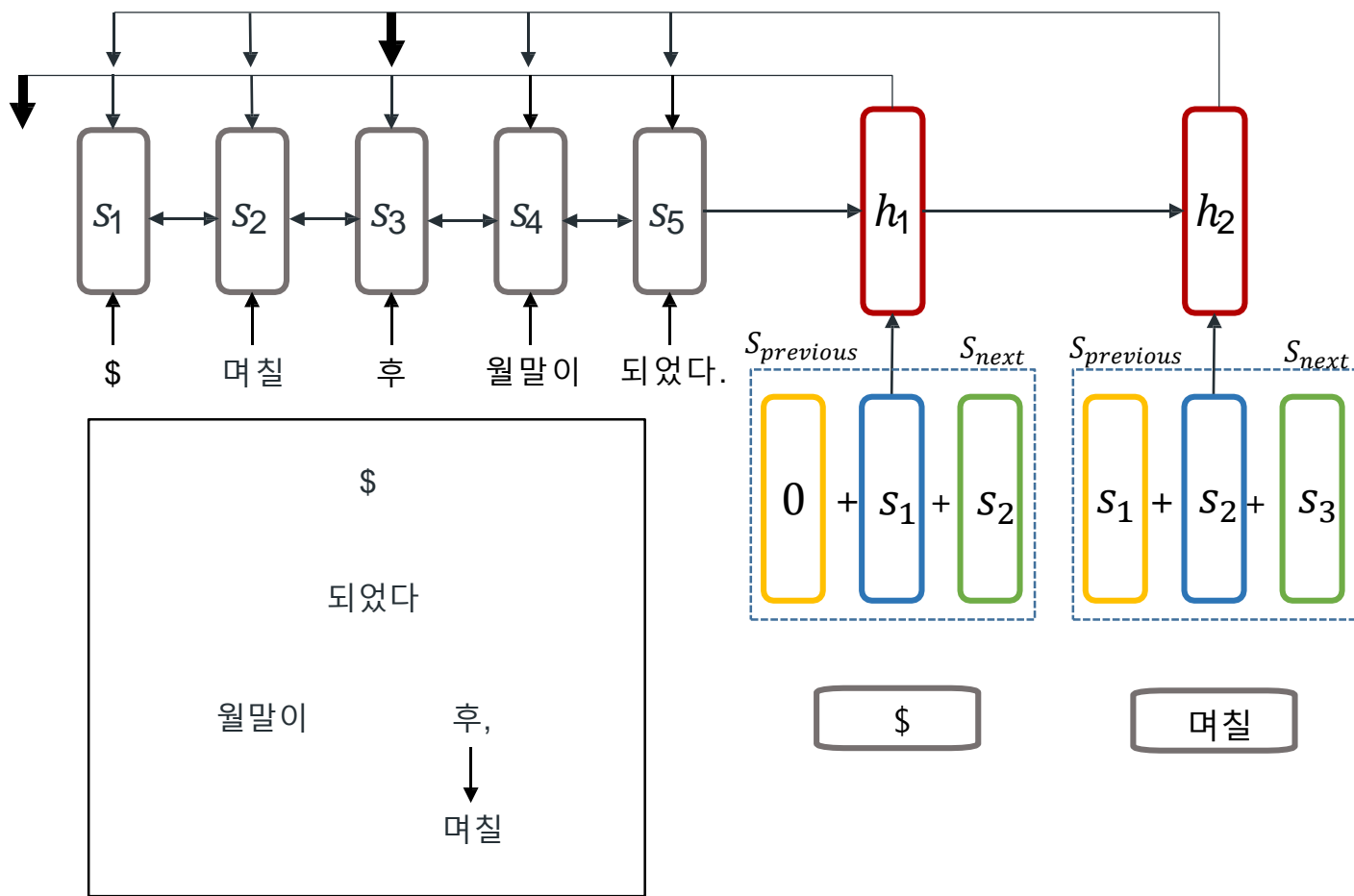


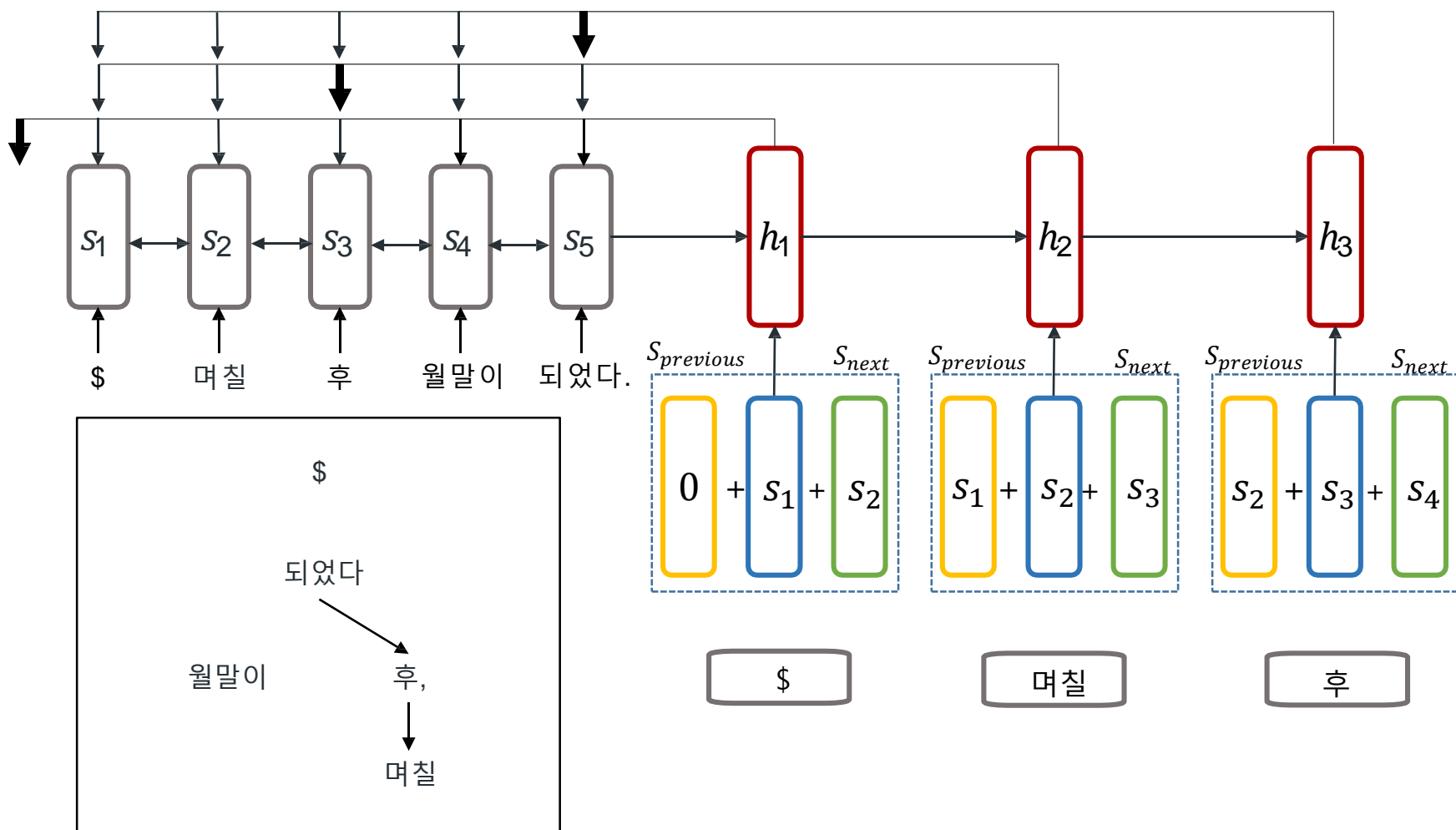
BiLSTM: S_i

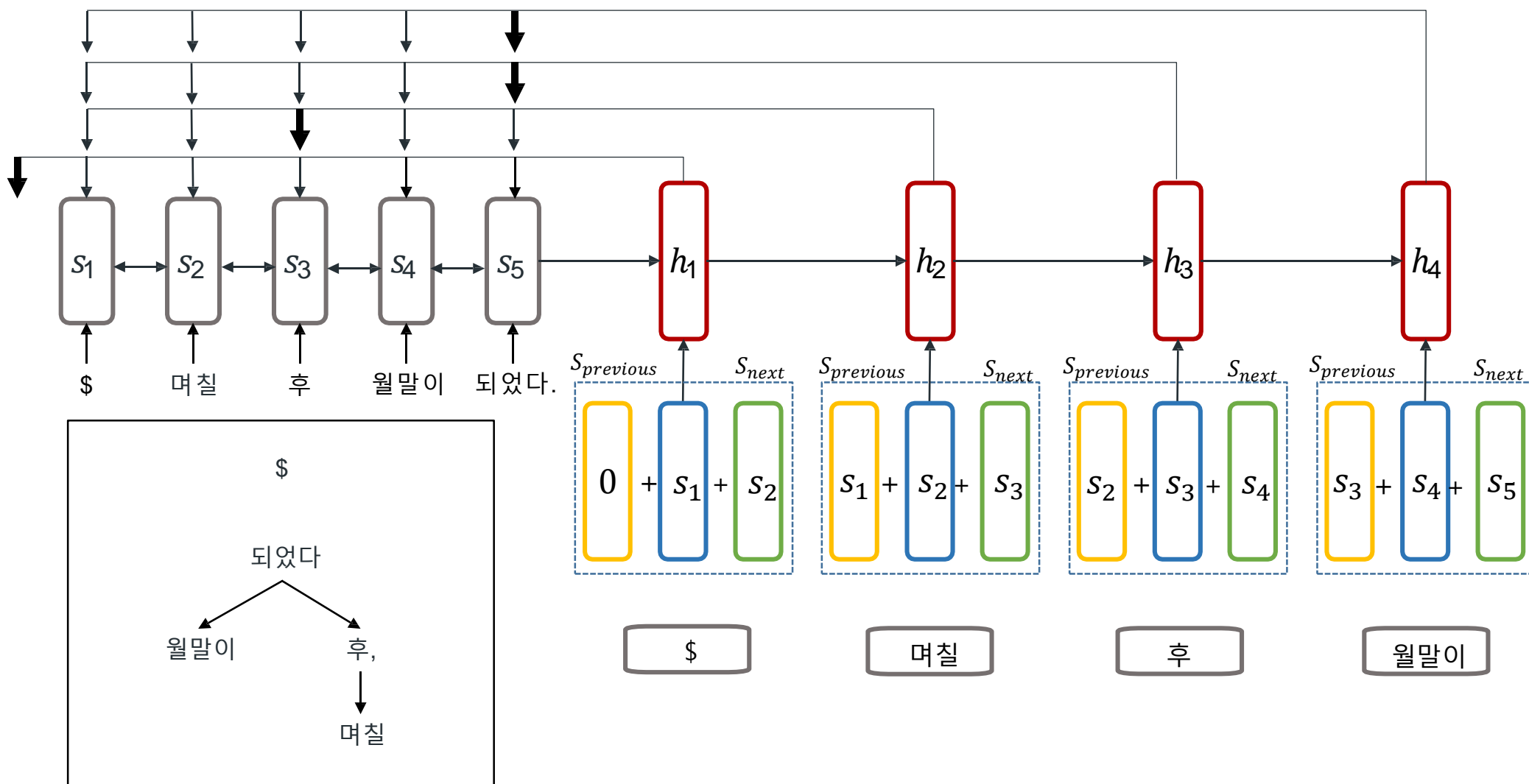
embeddings: x_i

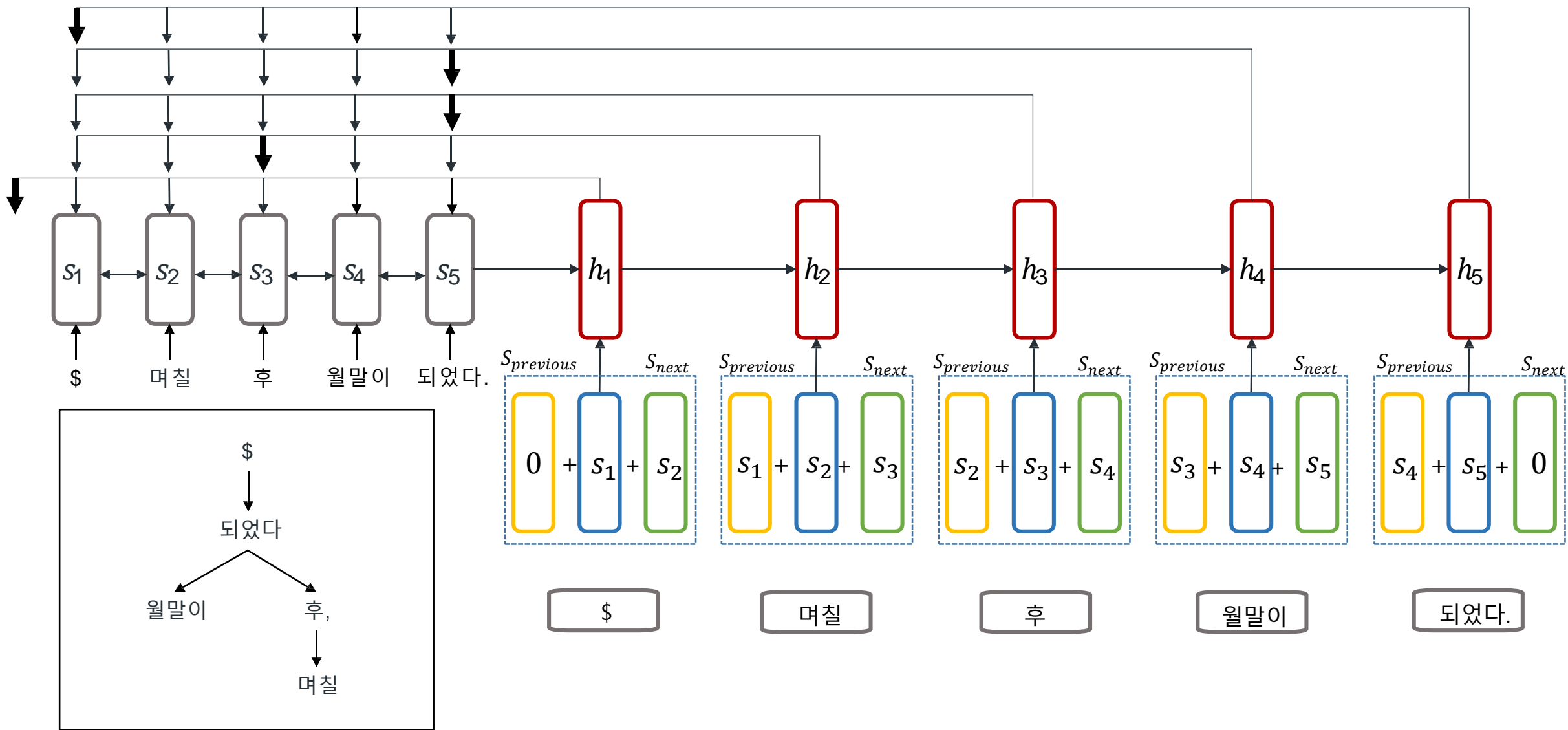
디코더



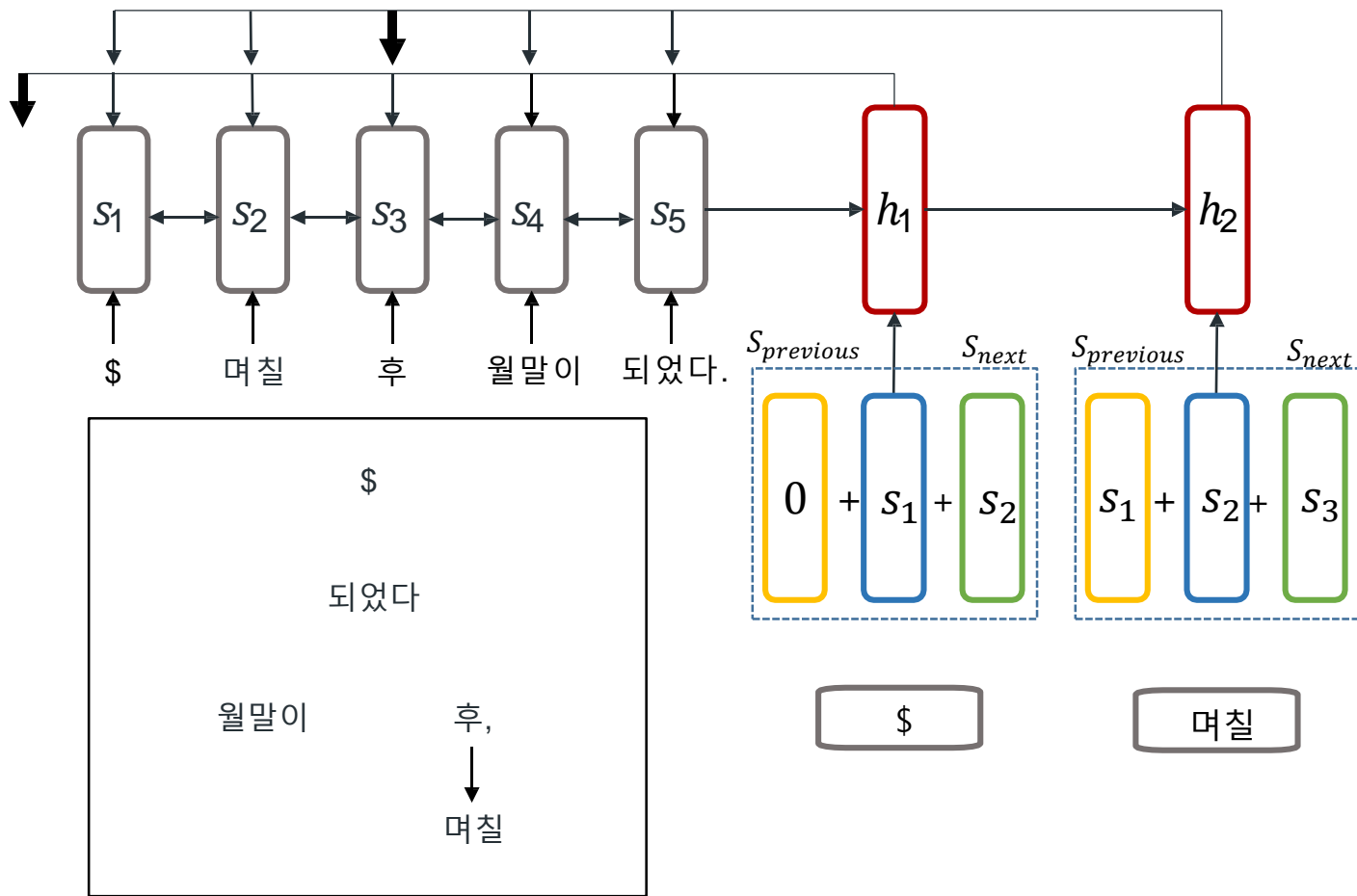




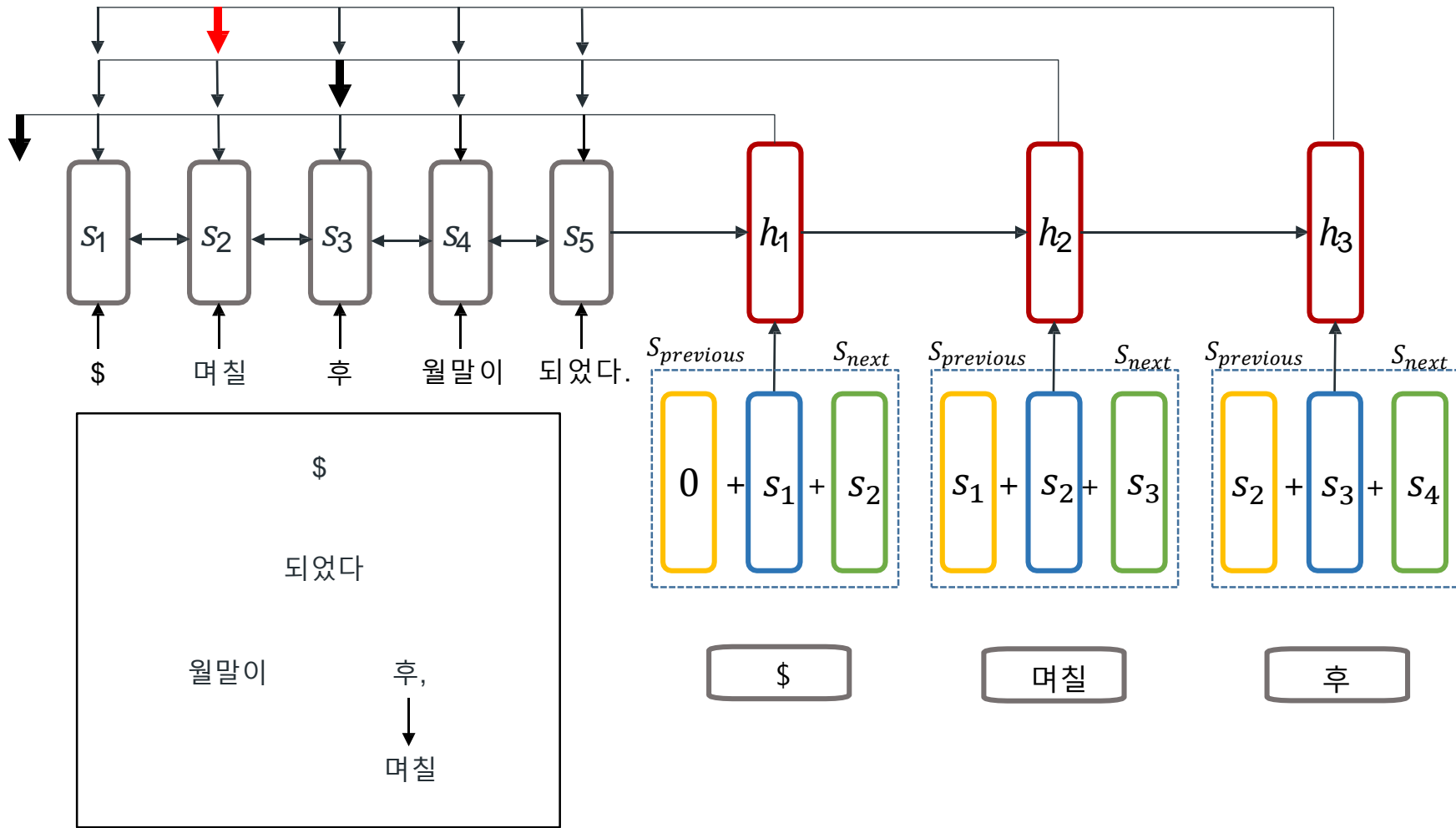




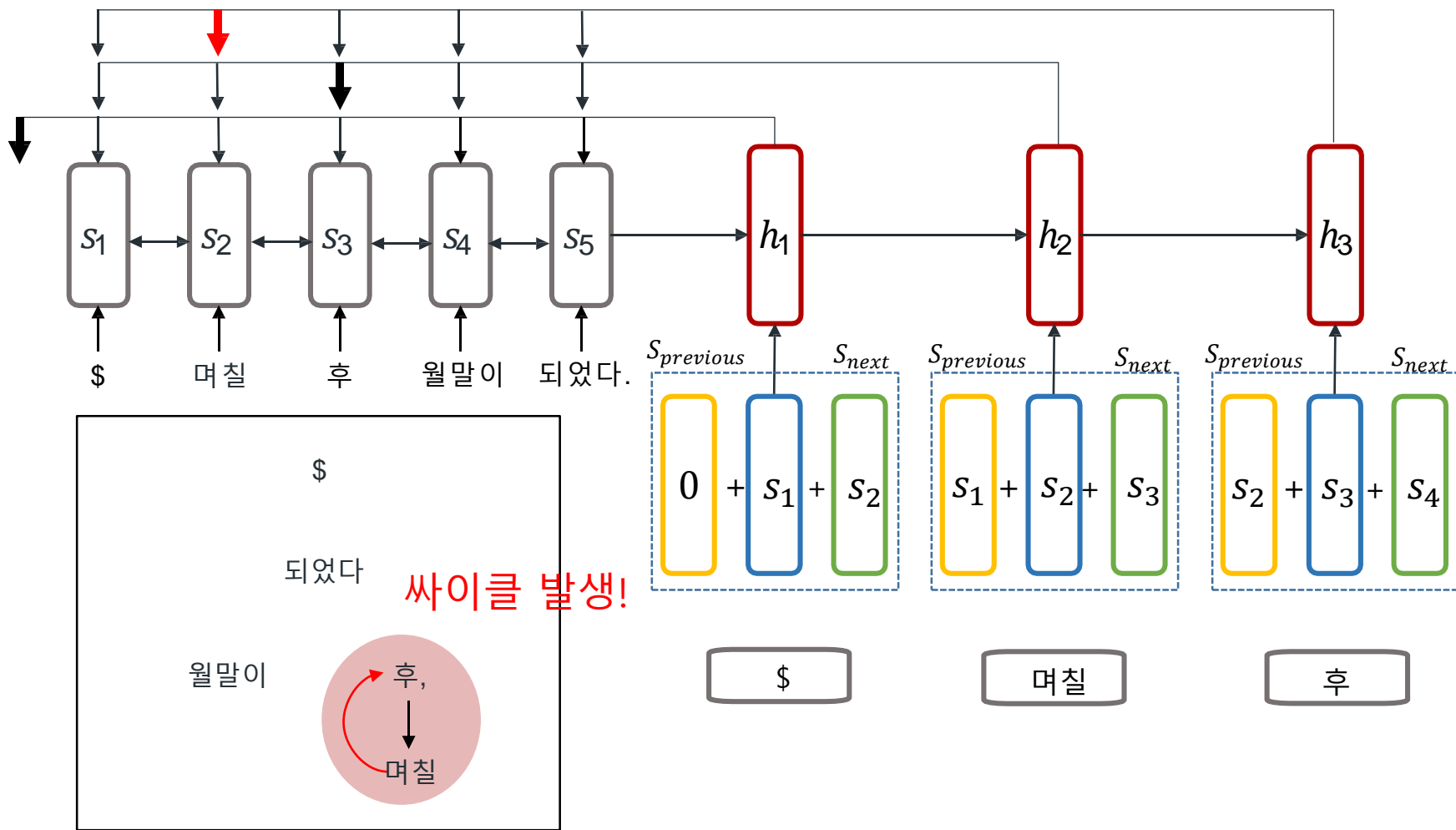
Cycle 체크



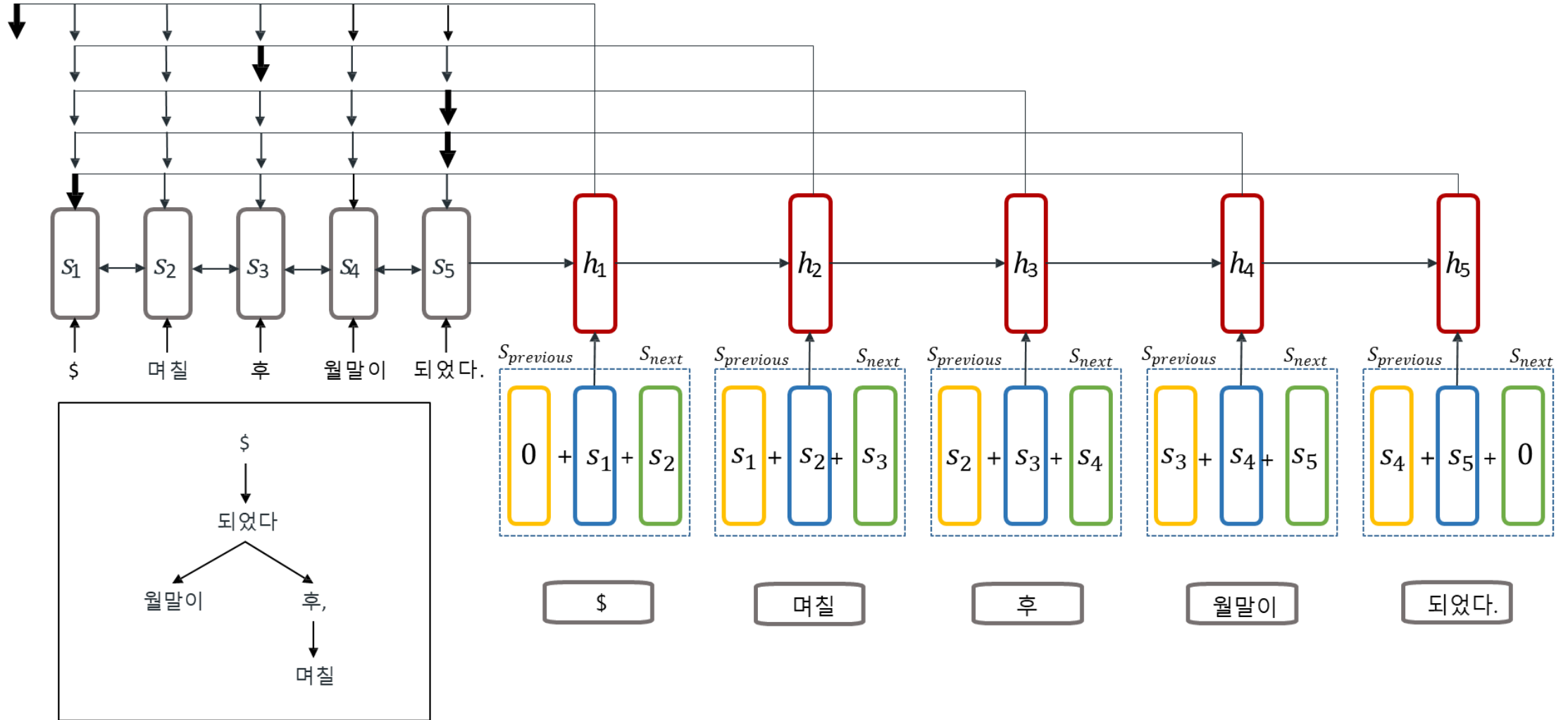
Cycle 체크



Cycle 체크



Right-to-Left 포인터 네트워크



Training details

✓Objective function

$$\begin{aligned} P_{\theta}(y|x) &= \prod_{i=1}^n P_{\theta}(l_i | l_{<i}, x) \\ &= \prod_{i=1}^n P_{\theta}(w_h | w_i, l_{<i}, x) \end{aligned}$$

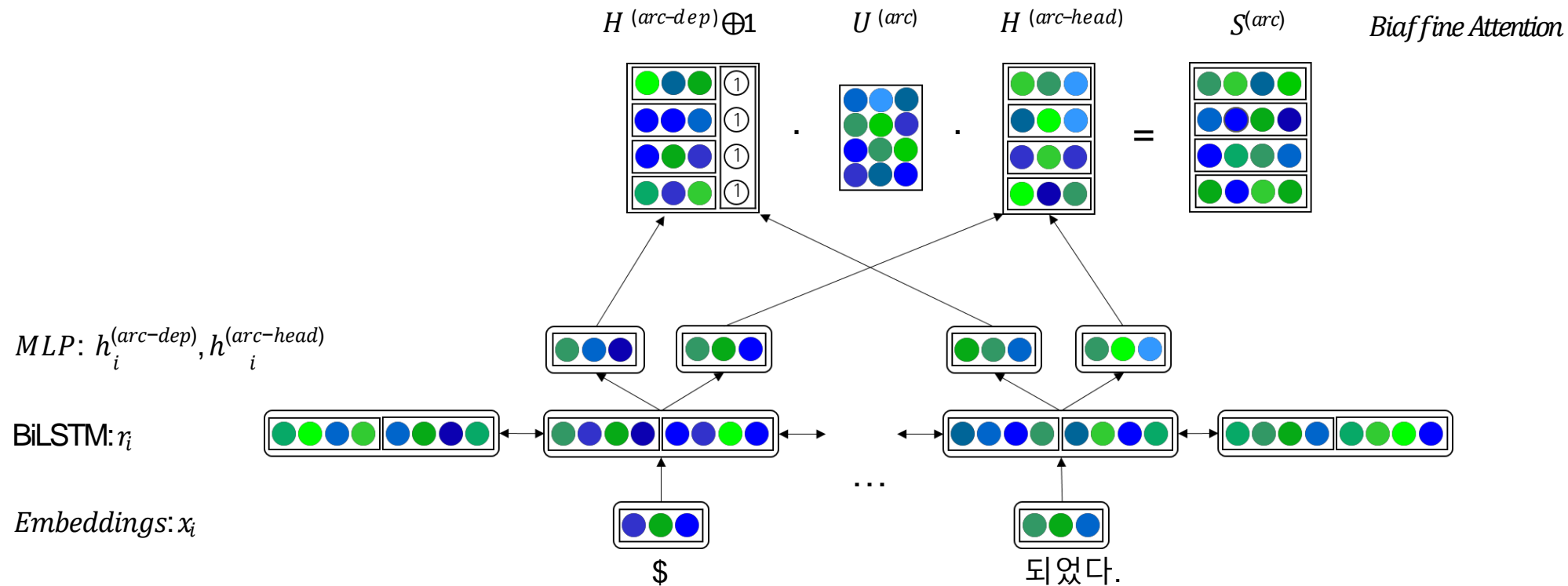
✓Dropout: 0.33

✓Optimizer: BertAdam, lr=2e-5

✓ELMo 임베딩

✓BERT 임베딩 (ETRI 공공 인공지능 오픈 API·DATA 중 한국어 BERT 언어모델)

Deep Biaffine Network



- ✓ Embedding, BiLSTM, MLP, Biaffine Attention으로 구성
- ✓ 각 어절에 대한 의존소, 지배소 표상 생성
- ✓ Biaffine attention을 통해 의존 관계 및 관계명 예측

ELMo

순방향 언어 모델

역방향 언어 모델

열/VV

었/EP

다/EC

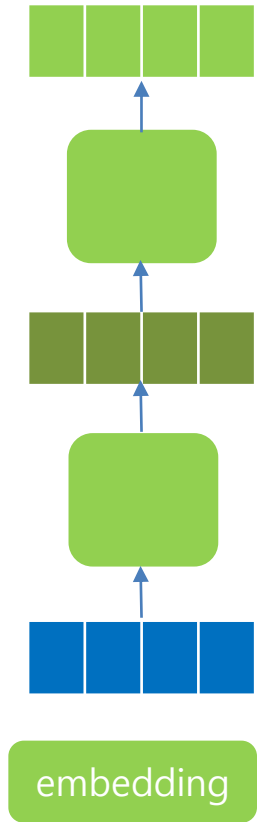
열/VV

었/EP

다/EC

ELMo

순방향 언어 모델

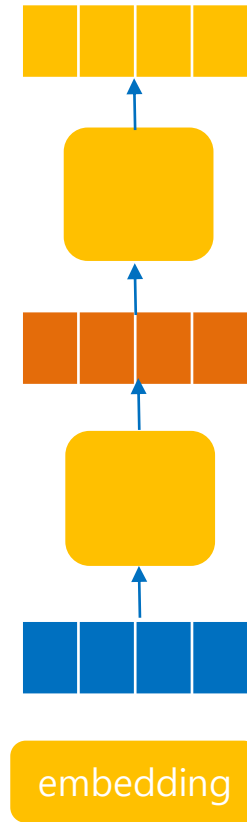


열/VV

있/EP

다/EC

역방향 언어 모델



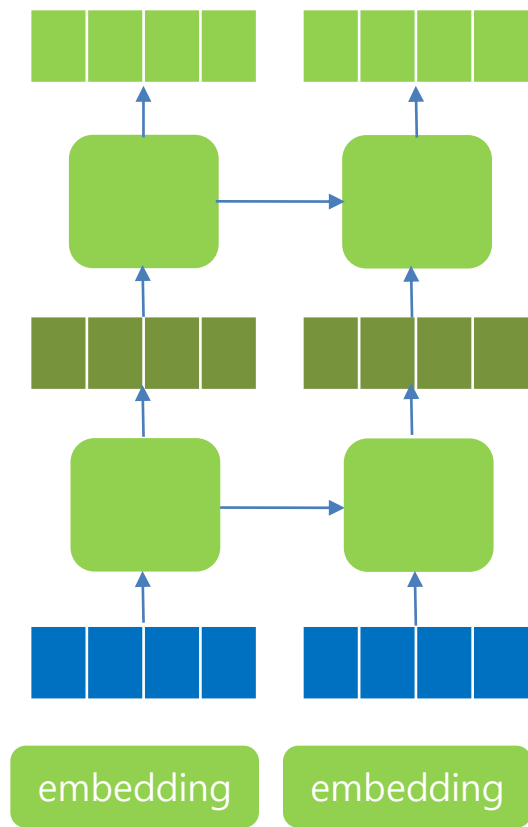
열/VV

있/EP

다/EC

ELMo

순방향 언어 모델

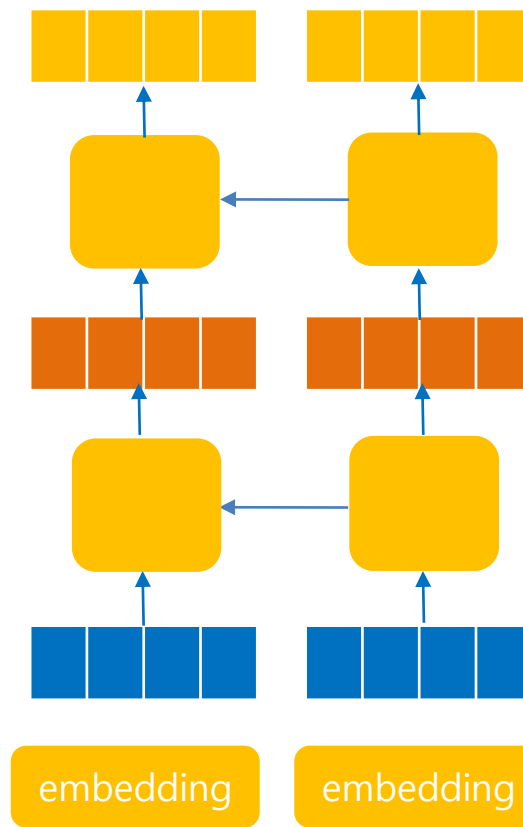


열/VV

있/EP

다/EC

역방향 언어 모델



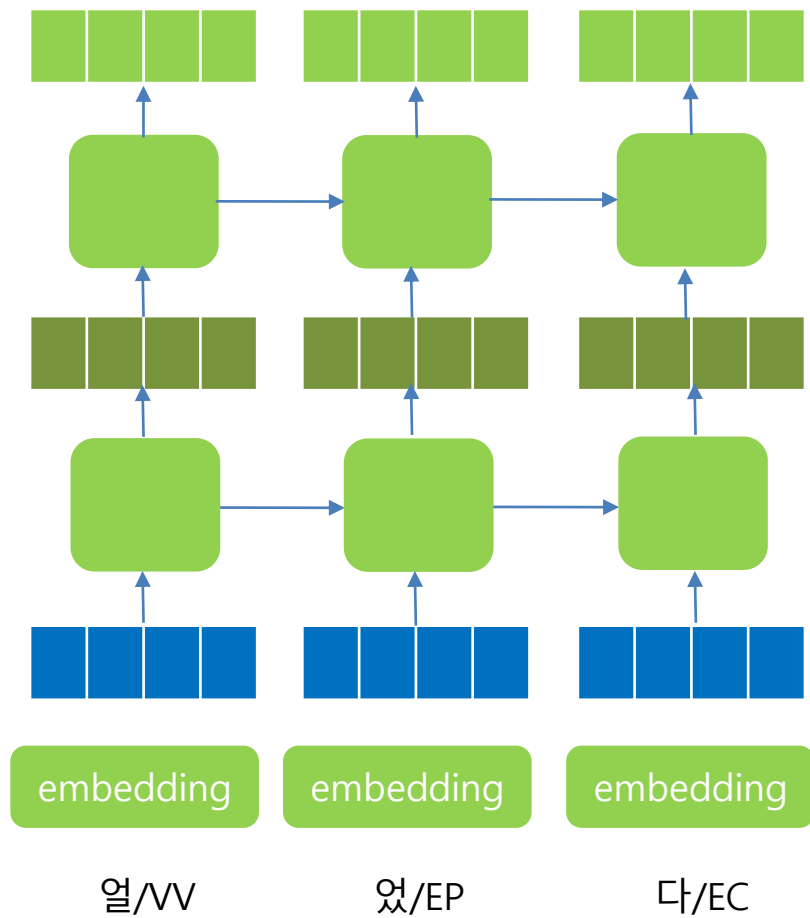
열/VV

있/EP

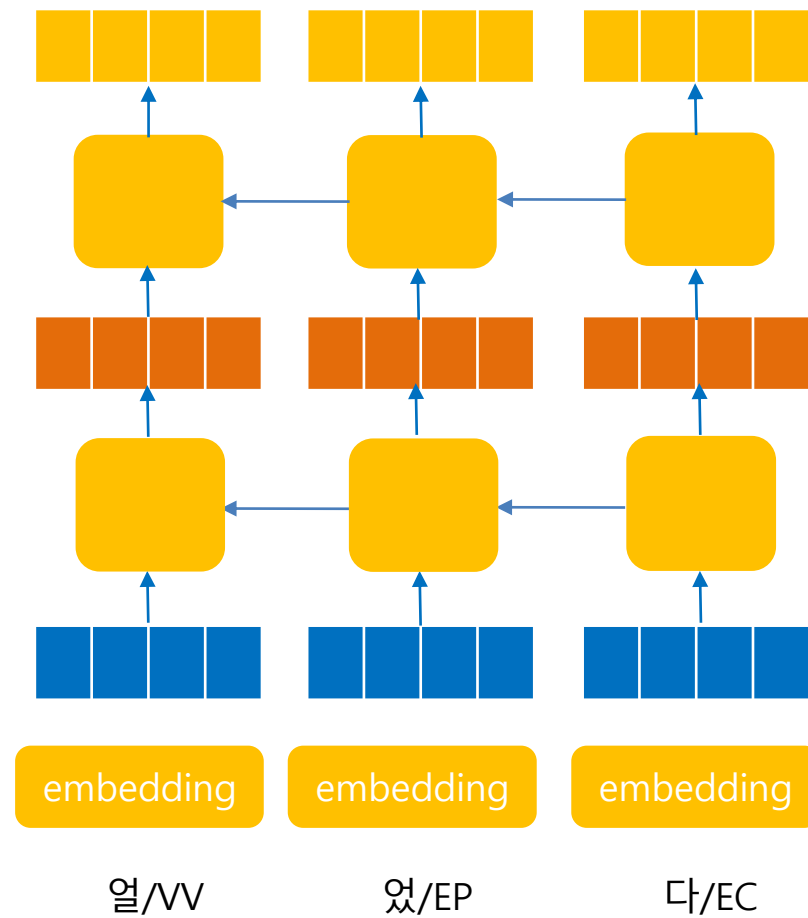
다/EC

ELMo

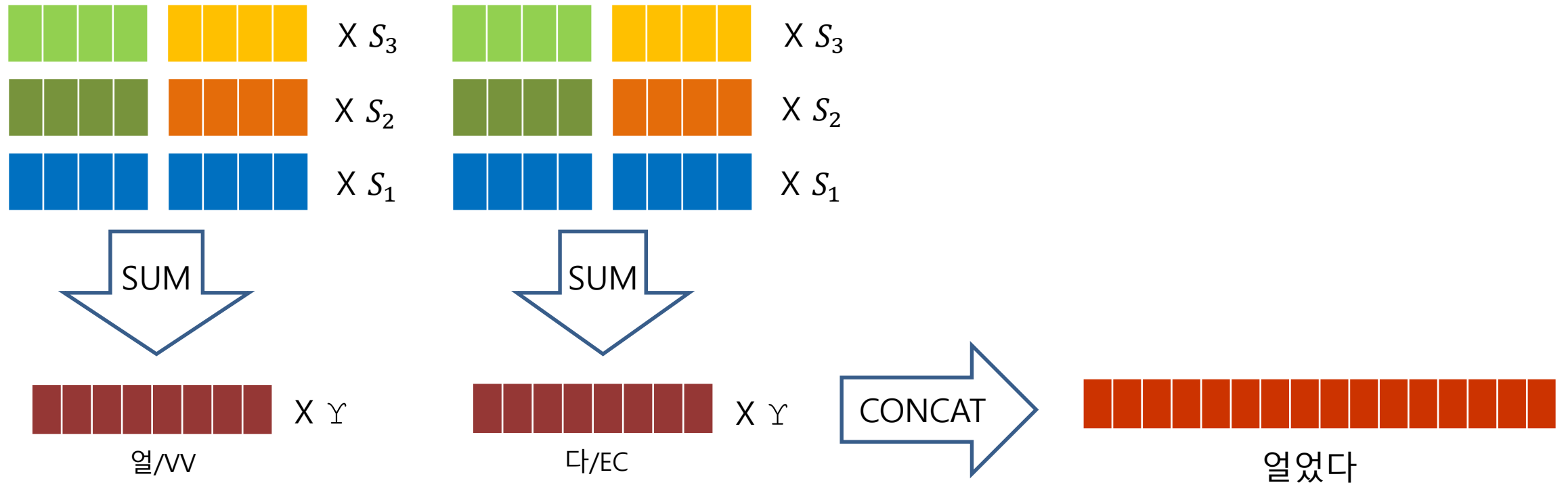
순방향 언어 모델



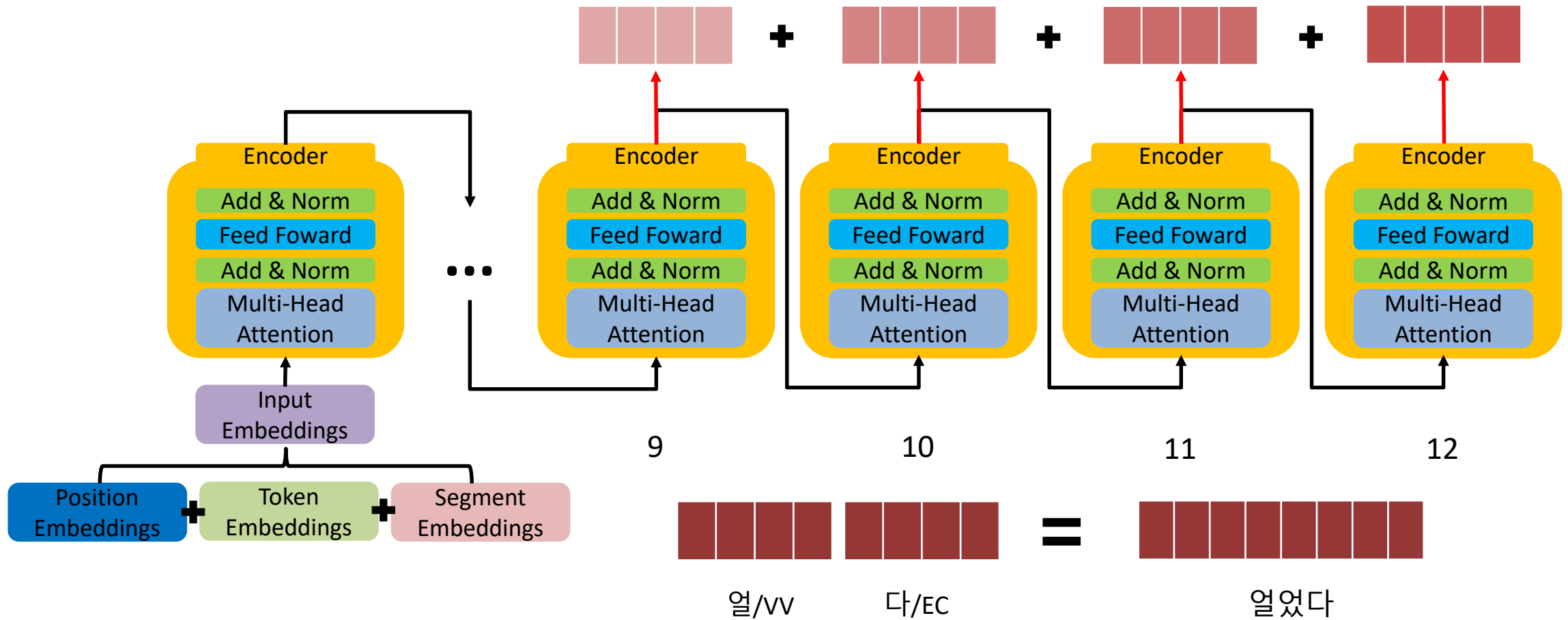
역방향 언어 모델



ELMo



BERT



앙상블

- ✓ 세 가지 모델
- ✓ 같은 모델에 대해서는
다른 임베딩 사용
- ✓ 서로 다른 오답

모델
BiAffine + BERT
BiAffine + ELMo + BERT
LR_Parser + BERT
LR_Parser + ELMo + BERT
RL_Parser + BERT

실험 결과 및 분석

모델	UAS	LAS
BiAffine + BERT	94.07	91.92
BiAffine + ELMo + BERT	94.10	92.03
LR_Parser + BERT	94.14	92.07
LR_Parser + ELMo + BERT	94.23	92.12
Ensemble_1	94.43	92.40

모델	UAS	LAS
BiAffine + BERT	94.07	91.92
BiAffine + ELMo + BERT	94.10	92.03
LR_Parser + BERT	94.14	92.07
LR_Parser + ELMo + BERT	94.23	92.12
RL_Parser + BERT	94.10	91.97
Ensemble_2	94.50	92.46

✓ RL_Parser을 앙상블하여 성능 향상

실험 결과 및 분석

모델	UAS	LAS
LR_Parser + ELMo + BERT	94.23	92.12
Ensemble	94.50	92.46

- ✓ 최종 앙상블 모델로 우수한 성능 달성
- ✓ 단일 모델로는 LR_Parser + ELMo + BERT가 좋은 성능

실험 결과 및 분석

의존 구문 분석	UAS	LAS
나승훈[1]	91.78	88.76
안휘진[2]	92.17	90.08
박성식[3]	92.85	90.65
홍승연[4]	93.12	91.00
박천음[5]	94.06	92.00
제안 모델	94.50	92.46

[1] 나승훈, et al. "Deep Biaffine Attention을 이용한 한국어 의존 파싱", KCC, pp. 584-586, 2017.

[2] 안휘진, et al. "Deep Bi-affine Network와 스택포인트 네트워크를 이용한 한국어 의존 구문 분석 시스템", HCLT, pp. 689-691, 2018.

[3] 박성식, et al. "ELMo와 멀티헤드 어텐션을 이용한 한국어 의존 구문 분석", HCLT, pp. 8-12, 2018.

[4] 홍승연, et al. "BERT와 ELMo 문맥화 단어 임베딩을 이용한 한국어 의존 파싱", 한국정보과학회 학술발표논문집, pp. 491-493, 2019.

[5] 박천음, et al. "BERT를 이용한 한국어 의존 구문 분석", 한국정보과학회 학술발표논문집, pp. 530-532, 2019.

결론

- ✓ 고성능 한국어 의존 구문 분석기 개발
 - ✓ Left-to-Right 포인터 네트워크
 - ✓ Right-to-Left 포인터 네트워크
 - ✓ Deep Biaffine 네트워크
 - ✓ 앙상블 모델
- ✓ Dev 성능 UAS 94.50 / LAS 92.46

감사합니다

