

텍스트마이닝 세미나

ToBig's 9기 신용재

Long Short-Term Memory models

a.k.a LSTM

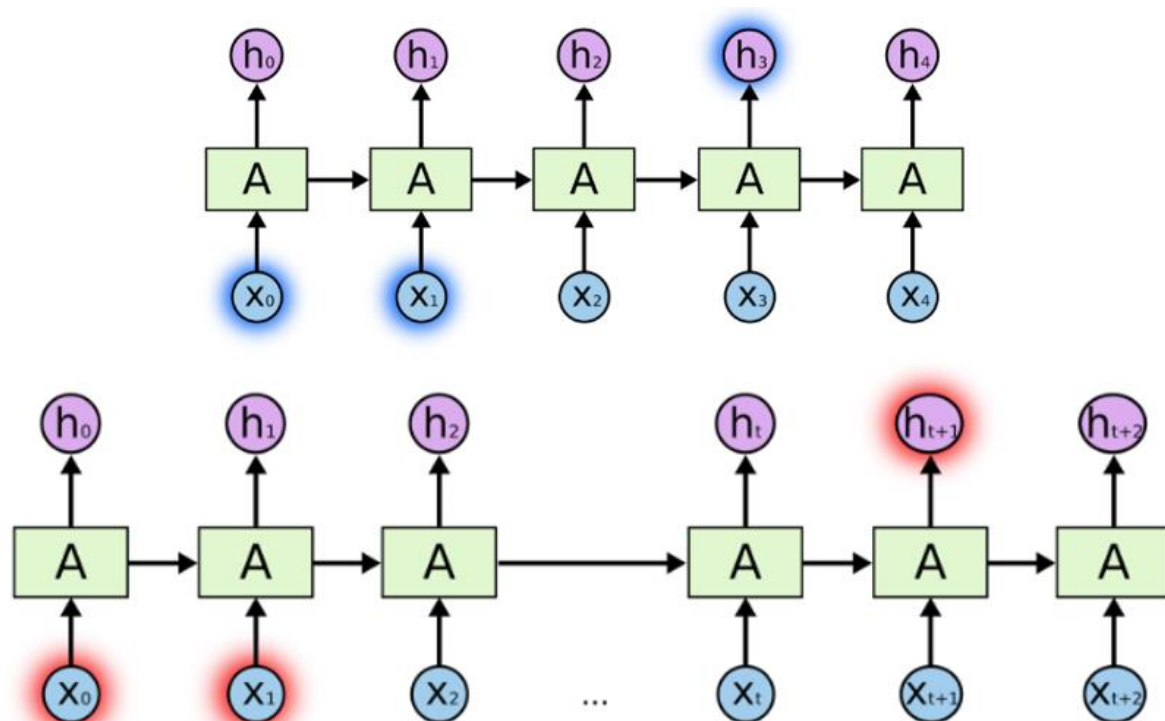
Contents

Unit 01 | LSTM이란?

Unit 02 | LSTM의 구조

Unit 03 | RNN과 LSTM의 예제

Unit 01 | LSTM이란?



<관련 정보와 그 정보를 사용하는 지점 사이 거리가 멀 경우 RNN 학습능력 저하>

*RNN의 관련정보와 그 정보를 사용하는 지점 사이 거리가 멀 경우 Vanishing gradient problem 발생

*LSTM이란.

오랜 기간이 지난 정보도 기억할 수 있는 RNN의 종류.
RNN의 hidden-state에 cell-state를 추가한 구조

the clouds are in the *sky*,

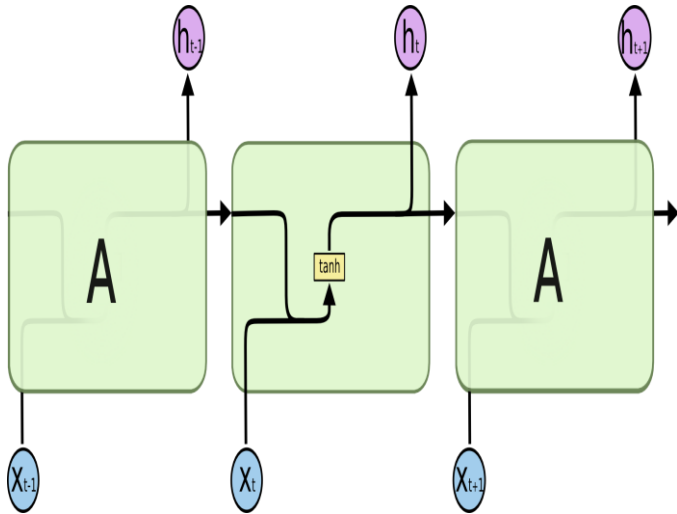
I grew up in France... I speak fluent *French*.

Cell State란.

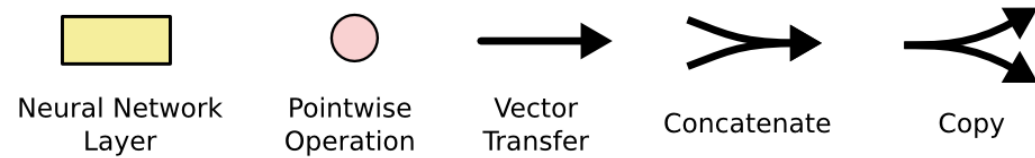
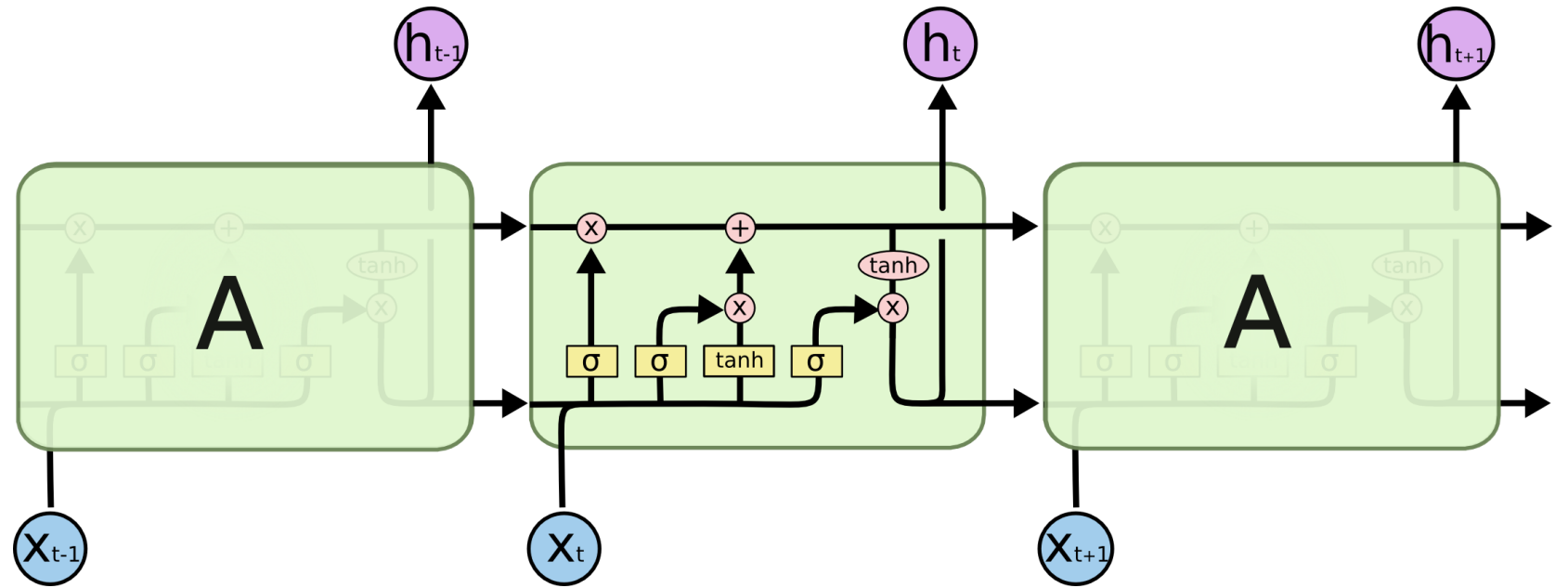
정보가 큰 변함없이 계속적으로 다음 단계에 전달되
게끔 하는 cell.

Unit 02 | LSTM의 구조

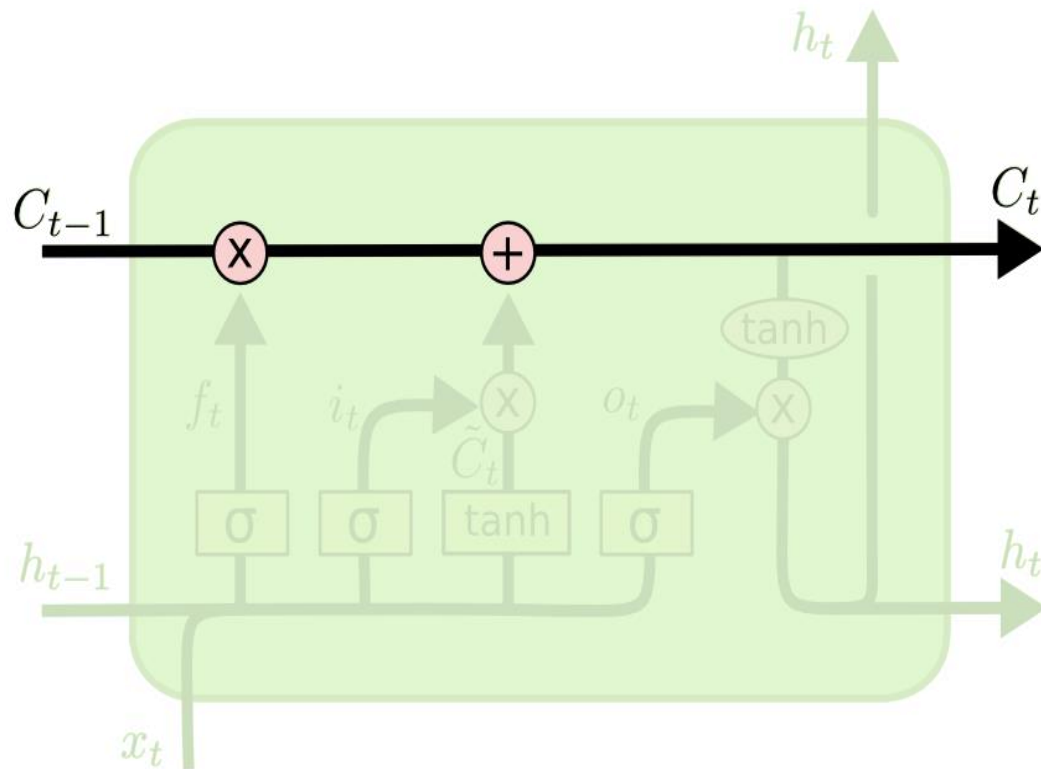
RNN



LSTM

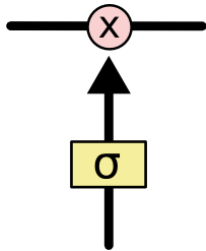


Unit 02 | LSTM의 구조



Cell State: 컨베이어 벨트처럼 전체 체인을 통과한다.

Gate를 통해 cell state에 정보를 더할지 뺄지 결정한다. Gate는 시그모이드와 점단위 곱 연산으로 이루어진다. 총 3개!



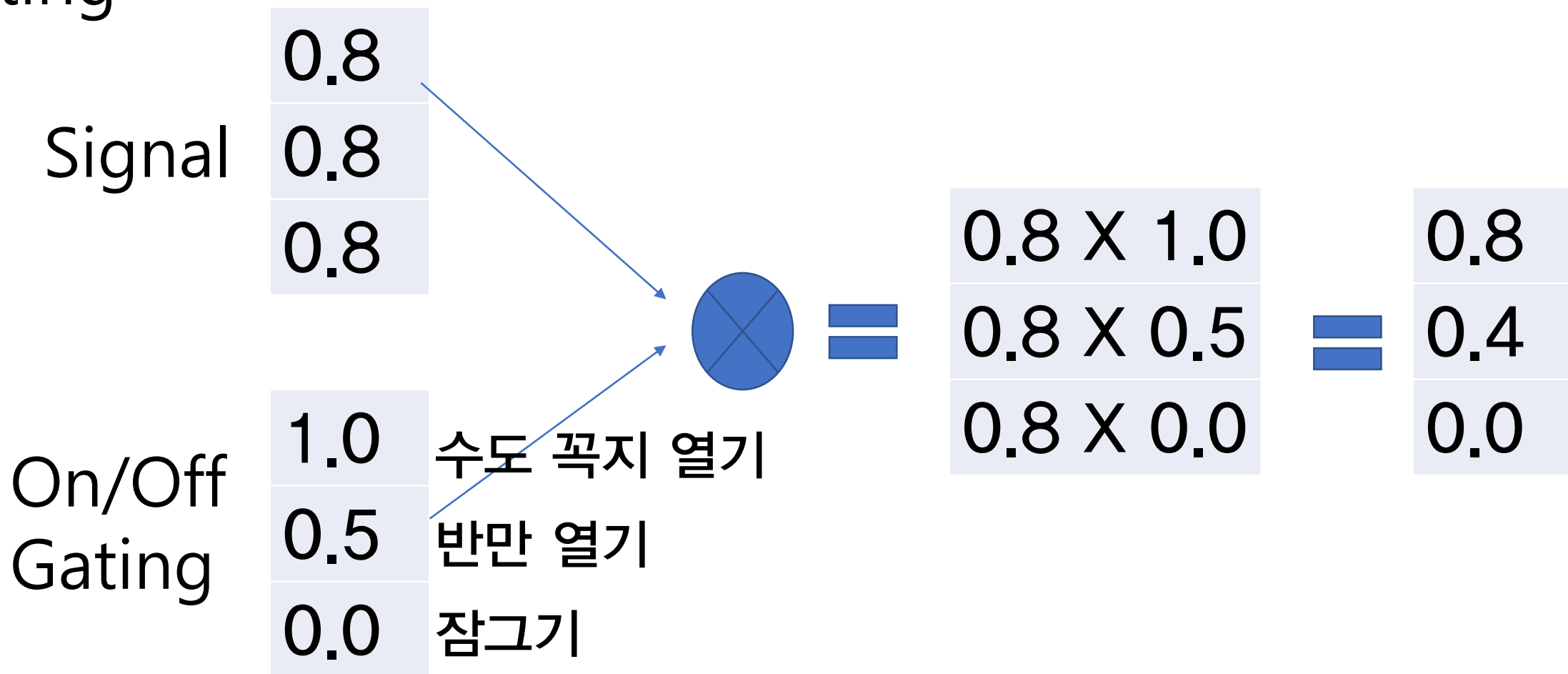
시그모이드 레이어의 값은 0~1.

0: 다 막기

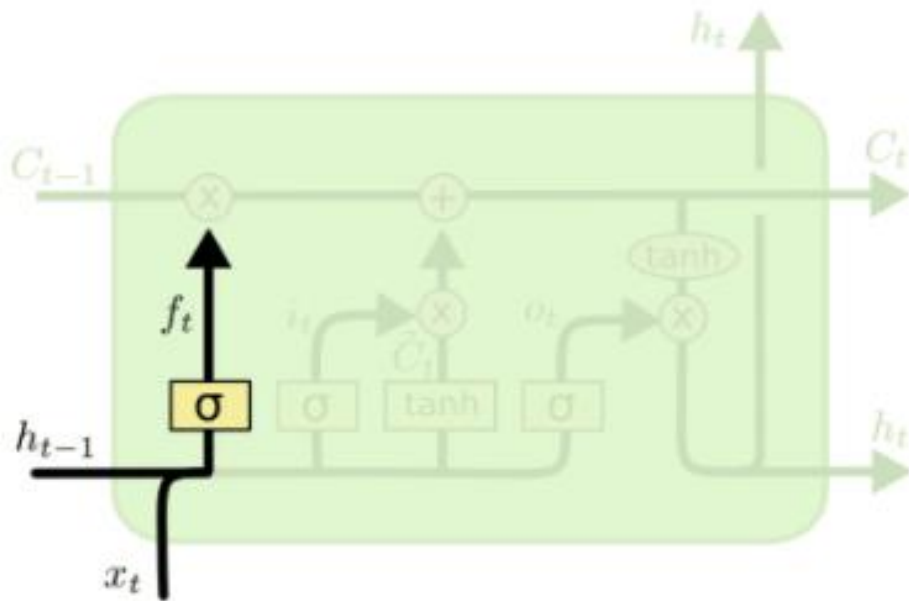
1: 다 통과시키기

Unit 02 | LSTM의 구조

Gating



Unit 02 | LSTM의 구조



$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f)$$

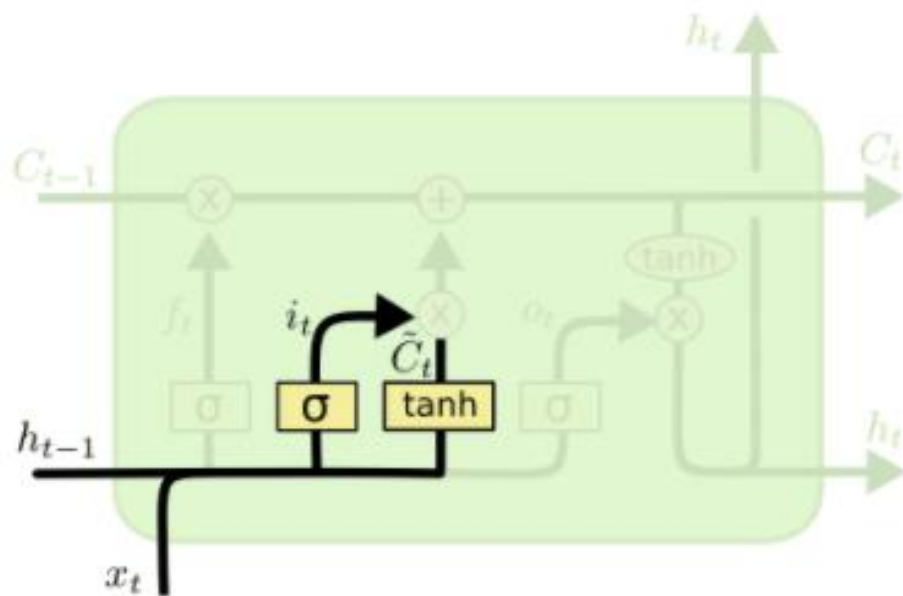
1. Forget gate (과거정보 잊기!)

Cell state 정보를 버릴지 말지 결정. (0:잊기~1:기억)

Ex) 현재 주어의 성별을 기억하고 있다.

새로운 주어가 등장했을 때, 지난 주어의 성별을 잊을지 말지

Unit 02 | LSTM의 구조



2. 새로운 정보를 기억(저장)할지 결정(현재 정보 기억하기!!)
무슨 새로운 단어를 cell state에 담을지.

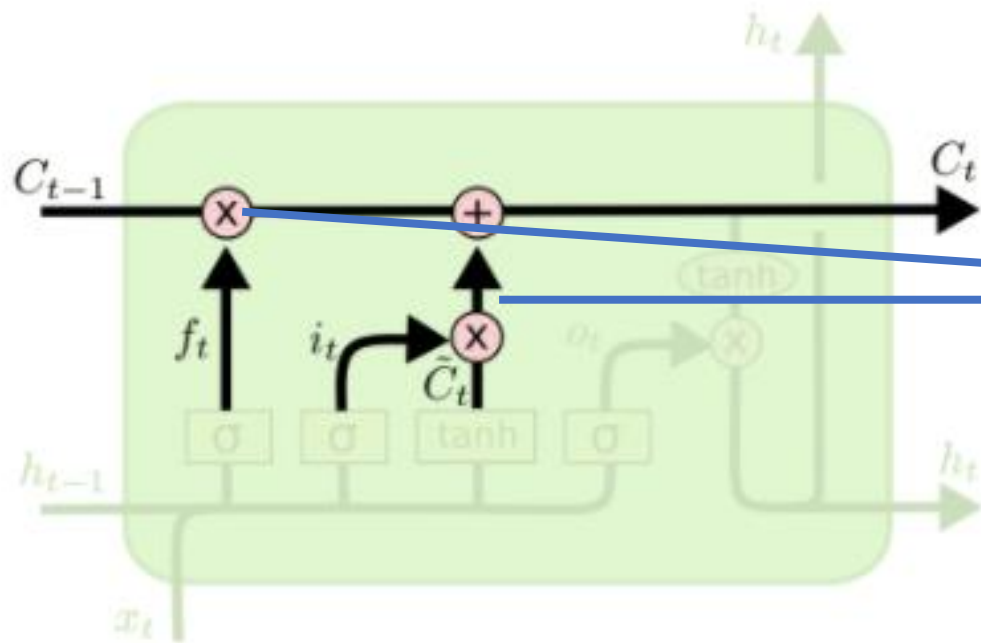
-sigmoid layer(input gate): 어떤 값을 업데이트할지 결정
-tanh layer: cell state에 더해질만한 후보 벡터를 만든다.

Ex) 새로운 주어에 해당하는 성별을 추가하기 원한다.
그 전에 있었던 주어의 성별을 잊고,

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Unit 02 | LSTM의 구조



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

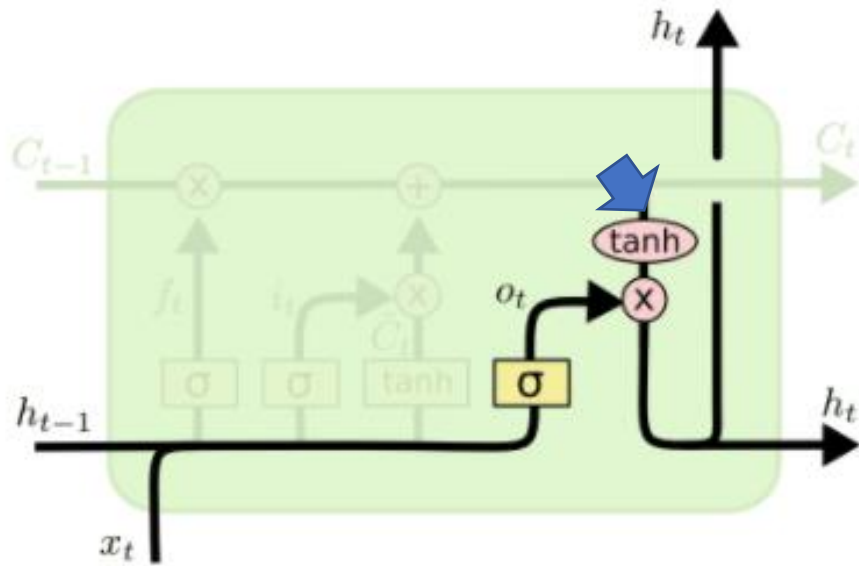
3. 이전의 셀 스테이트를 업데이트 시킨다.
1,2 단계에서 무엇을 할지 정했으니, 계산만 하면 된다.

→ 잊기로 한 것(게이트 통과)과 기억하기로 한 것(게이트 통과)의 합

새로운 후보 값이 기존 값에 영향을 주는 연산을 한다.

Ex) 1,2 단계에서 이루어진 결정을 이행한다.
과거 주어의 성별을 잊고,
새로운 주어에 해당하는 성별을 추가한다.

Unit 02 | LSTM의 구조



$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

4. Output으로 무엇을 낼지 결정한다.

Cell state에 의해 결정되지만, filter된 버전으로 결정된다.

-Sigmoid layer: cell state의 어떤 부분을 출력할지 결정.

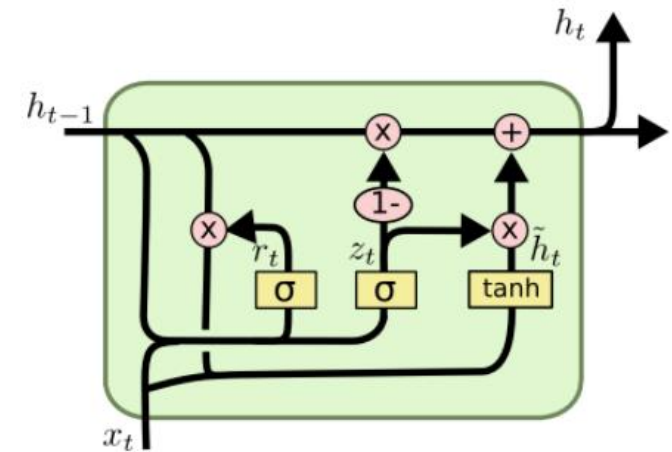
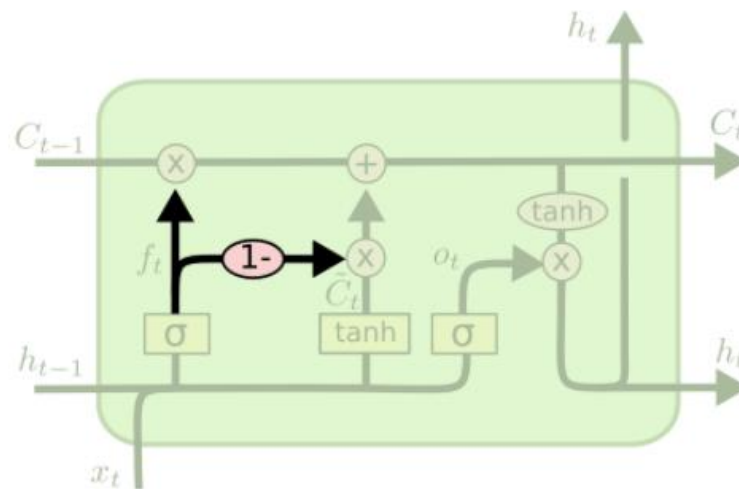
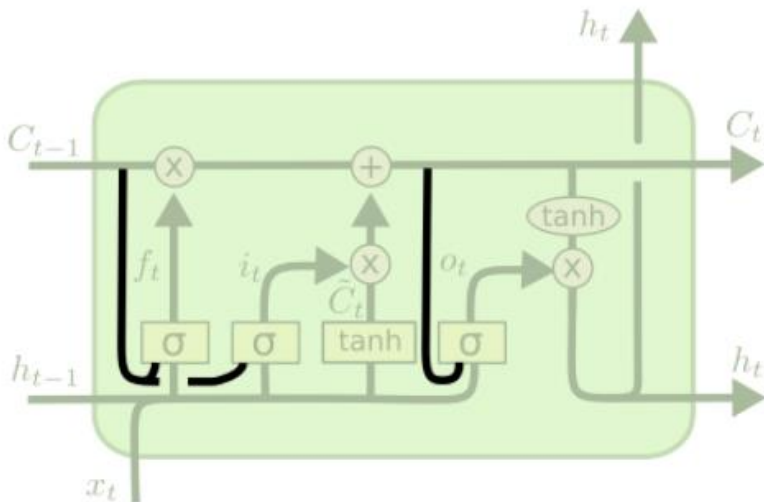
-tanh(-1~1)에 연산된 cell state를 통과시킨다.

-sigmoid layer를 통과한 값과 tanh를 통과한 cell state를 곱한다.

Ex)방금 전에 주어를 봤으니, 그에 맞는 동사를 출력으로 내놓고 싶다. 예를 들어, 주어가 복수인지 단수인지에 따라 어떤 동사를 취할지 알 수 있게 한다.

Unit 02 | LSTM의 구조

LSTM의 다른 형태



Gated Recurrent Unit

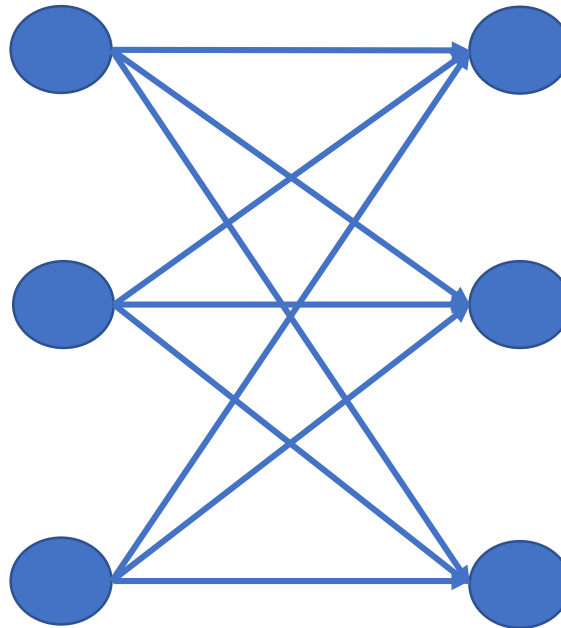
Unit 03 | RNN과 LSTM의 예제

오늘의 저녁 메뉴는?

요일

요일

야근



Unit 03 | RNN과 LSTM의 예제

RNN

어제 피자로 예측

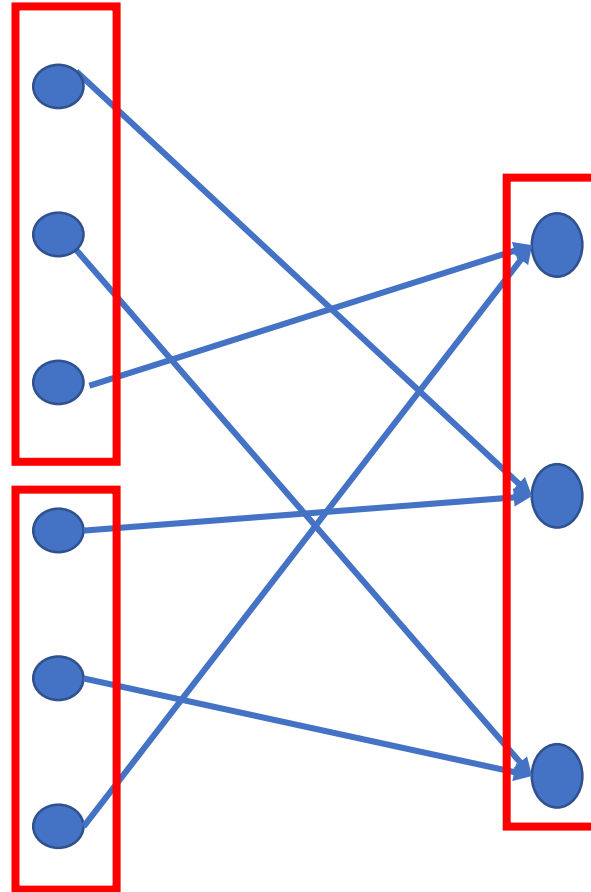
어제 스시로 예측

어제 와플로 예측

어제 피자

어제 스시

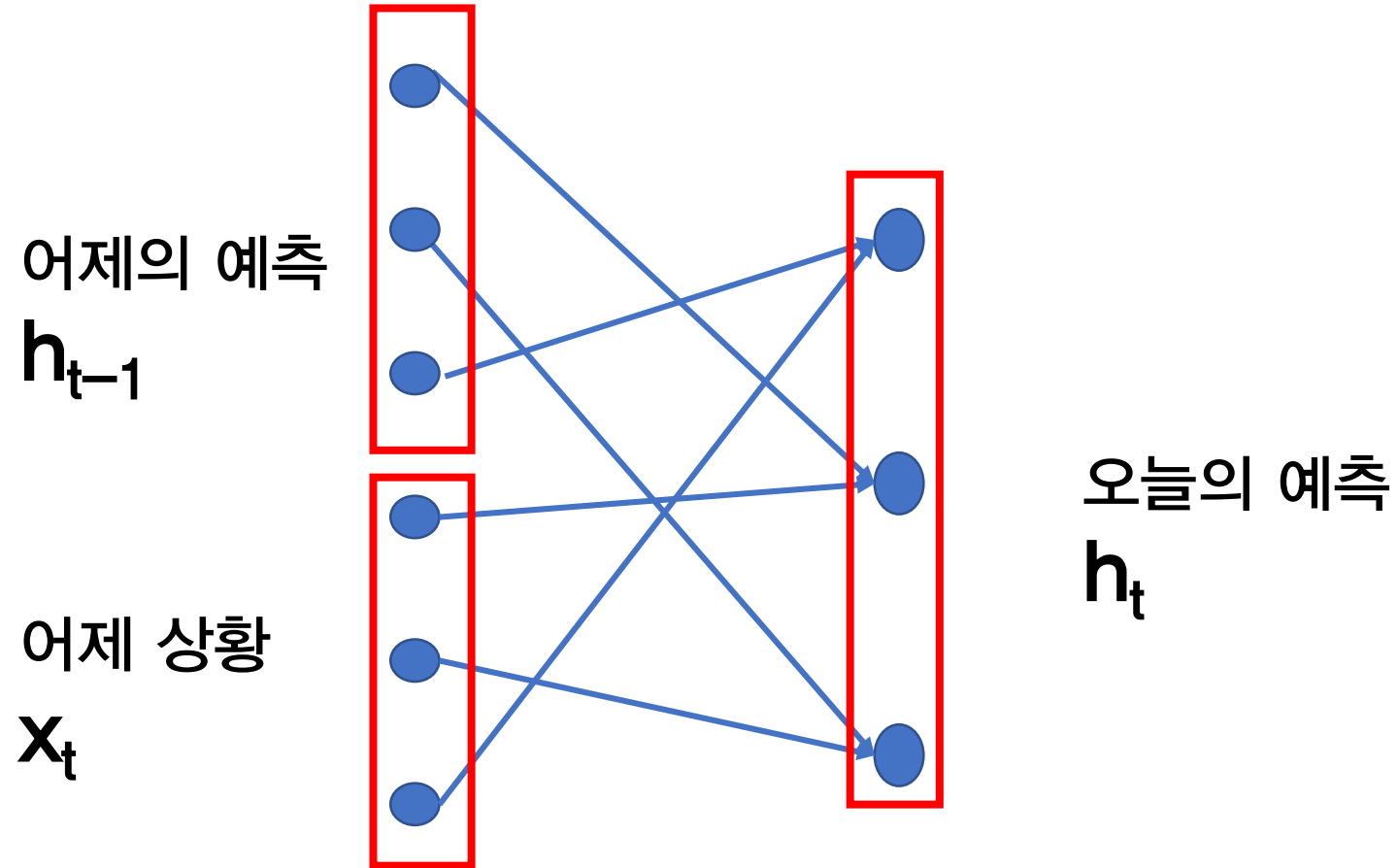
어제 와플



피자→스시
스시→와플
와플→피자

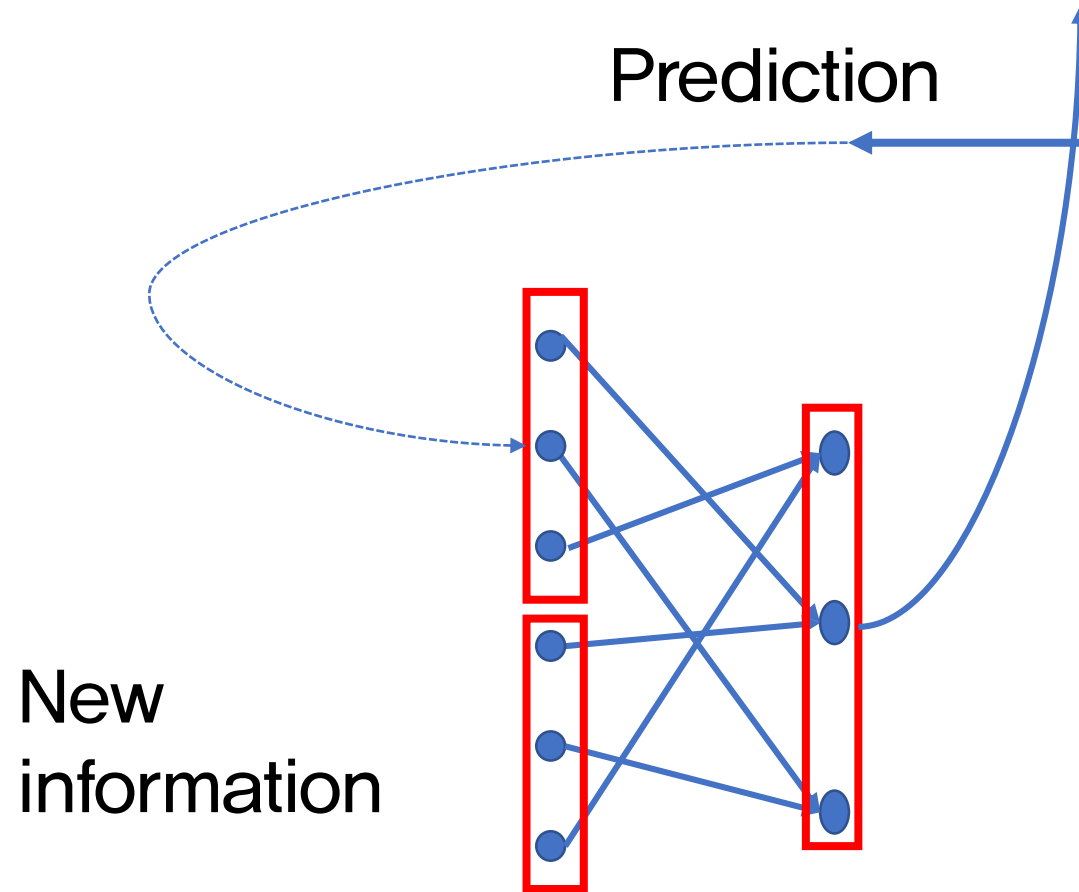
Unit 03 | RNN과 LSTM의 예제

RNN



Unit 03 | RNN과 LSTM의 예제

RNN



Unit 03 | RNN과 LSTM의 예제

간단한 예시

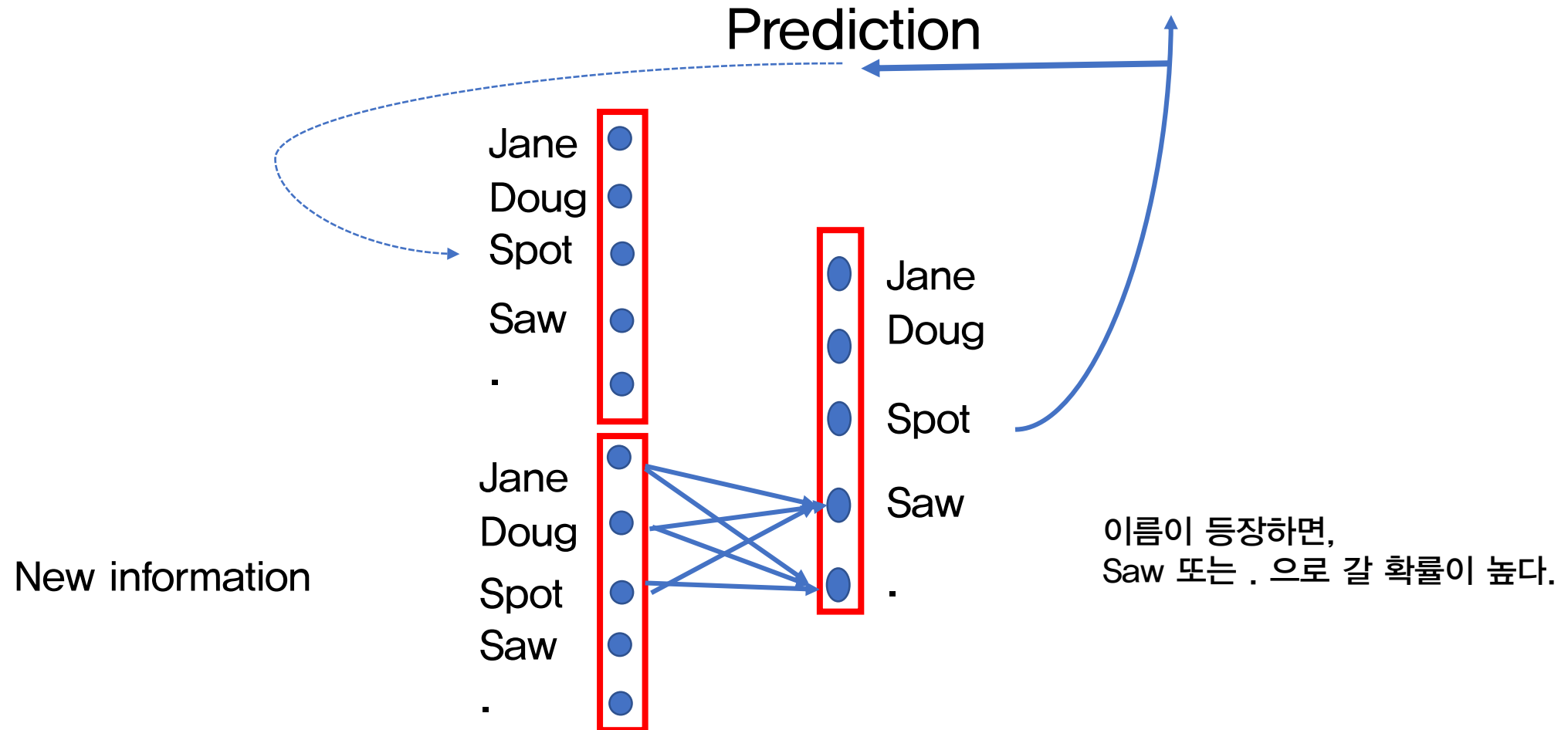
Doug saw Jane
Jane saw Spot.
Spot saw Doug

Dictionary:
{Doug, Jane, Spot, saw, .}

문제:
Dictionary에 있는 단어들을 사용해
문법적으로 올바른 순서로 나열

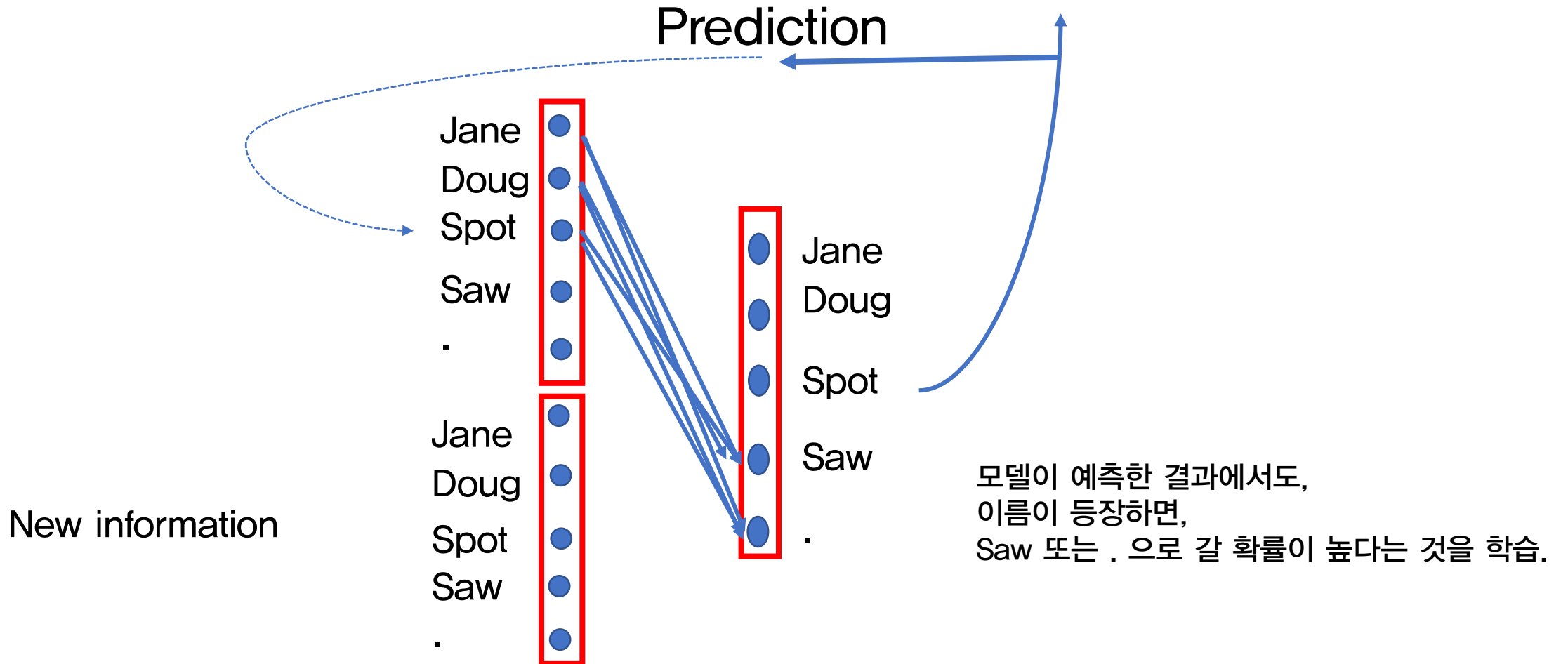
Unit 03 | RNN과 LSTM의 예제

RNN



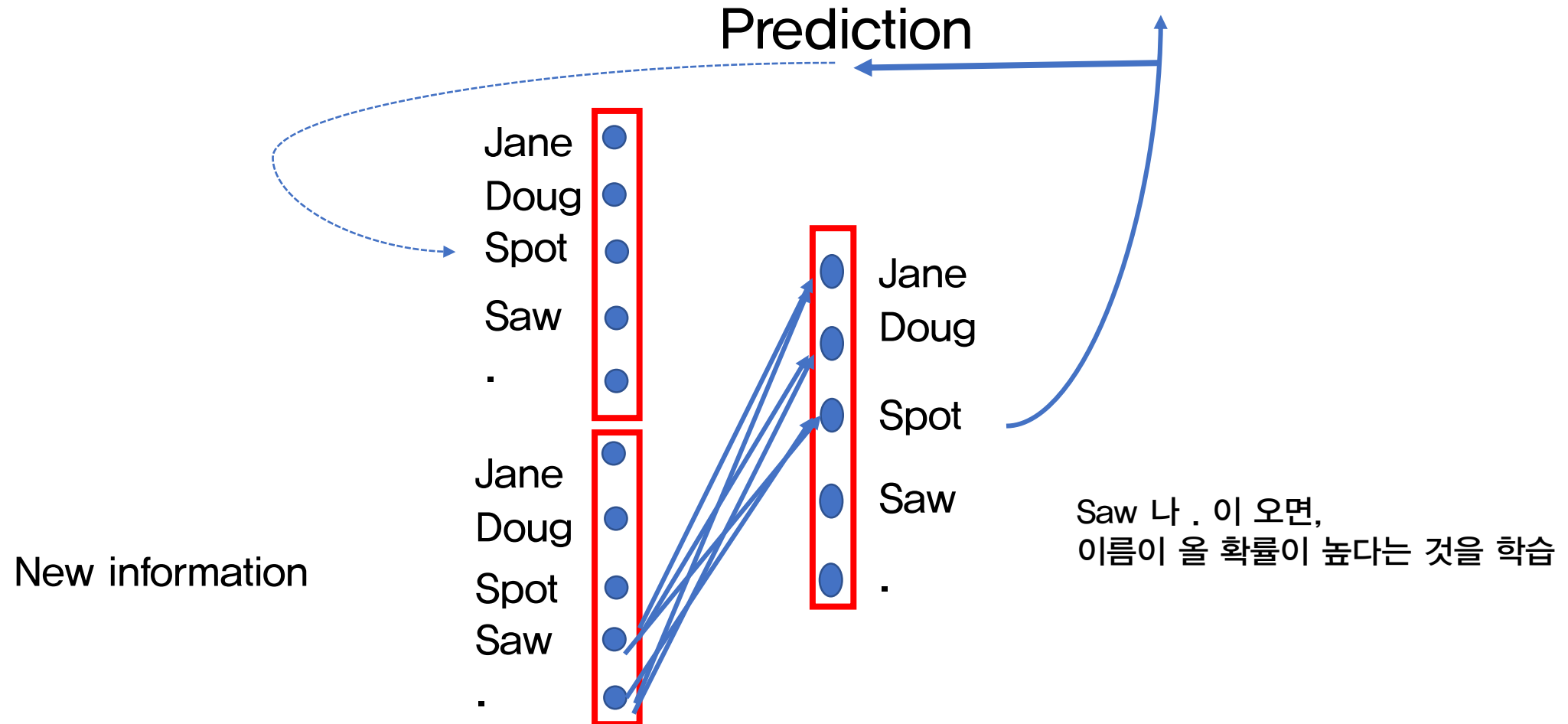
Unit 03 | RNN과 LSTM의 예제

RNN



Unit 03 | RNN과 LSTM의 예제

RNN



Unit 03 | RNN과 LSTM의 예제

RNN의 한계 Short-term memory만 활용한다면,,,

Doug saw Doug.

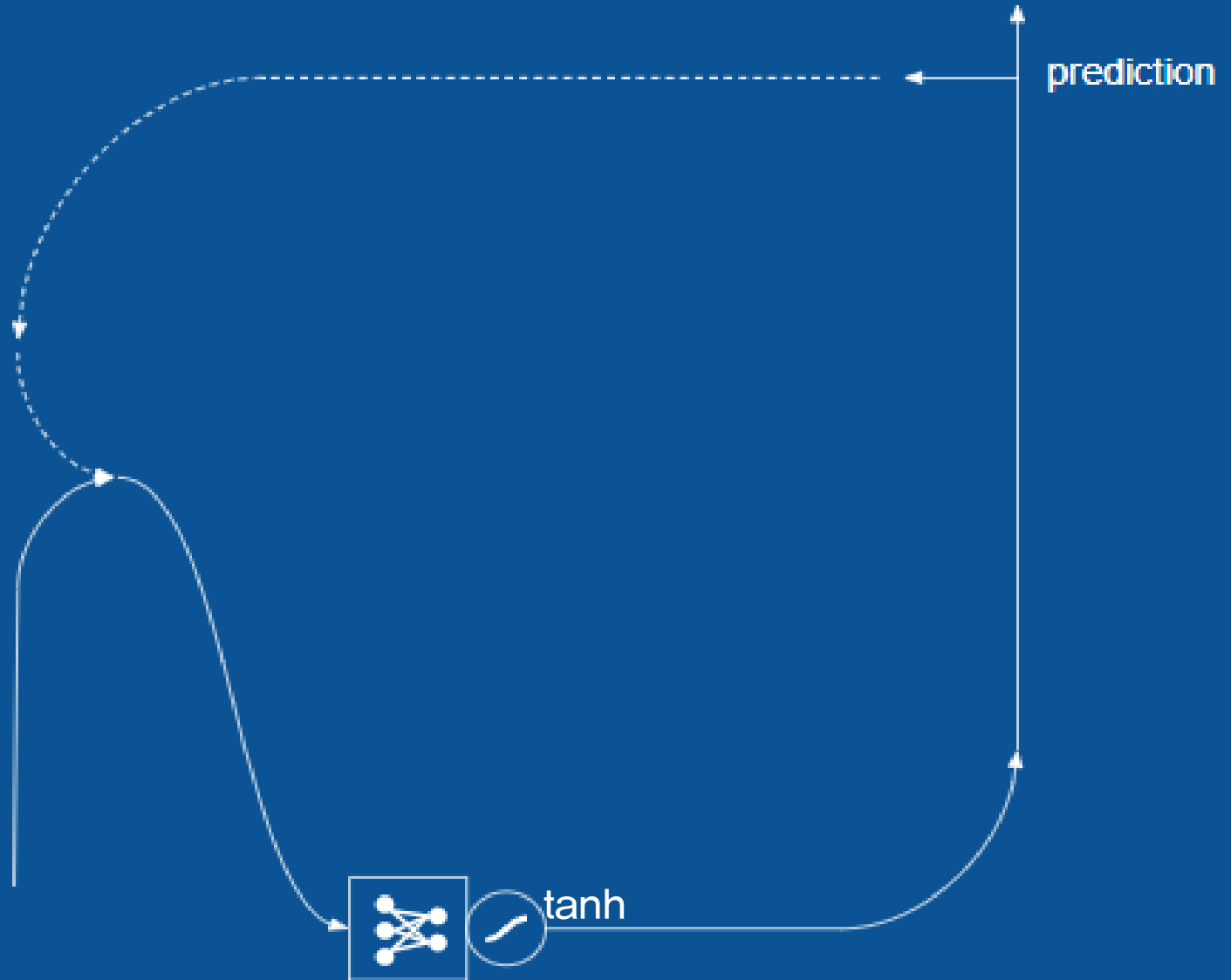
Jane saw Spot saw Doug saw ...

Spot. Doug. Jane.

Unit 03 | RNN과 LSTM의 예제

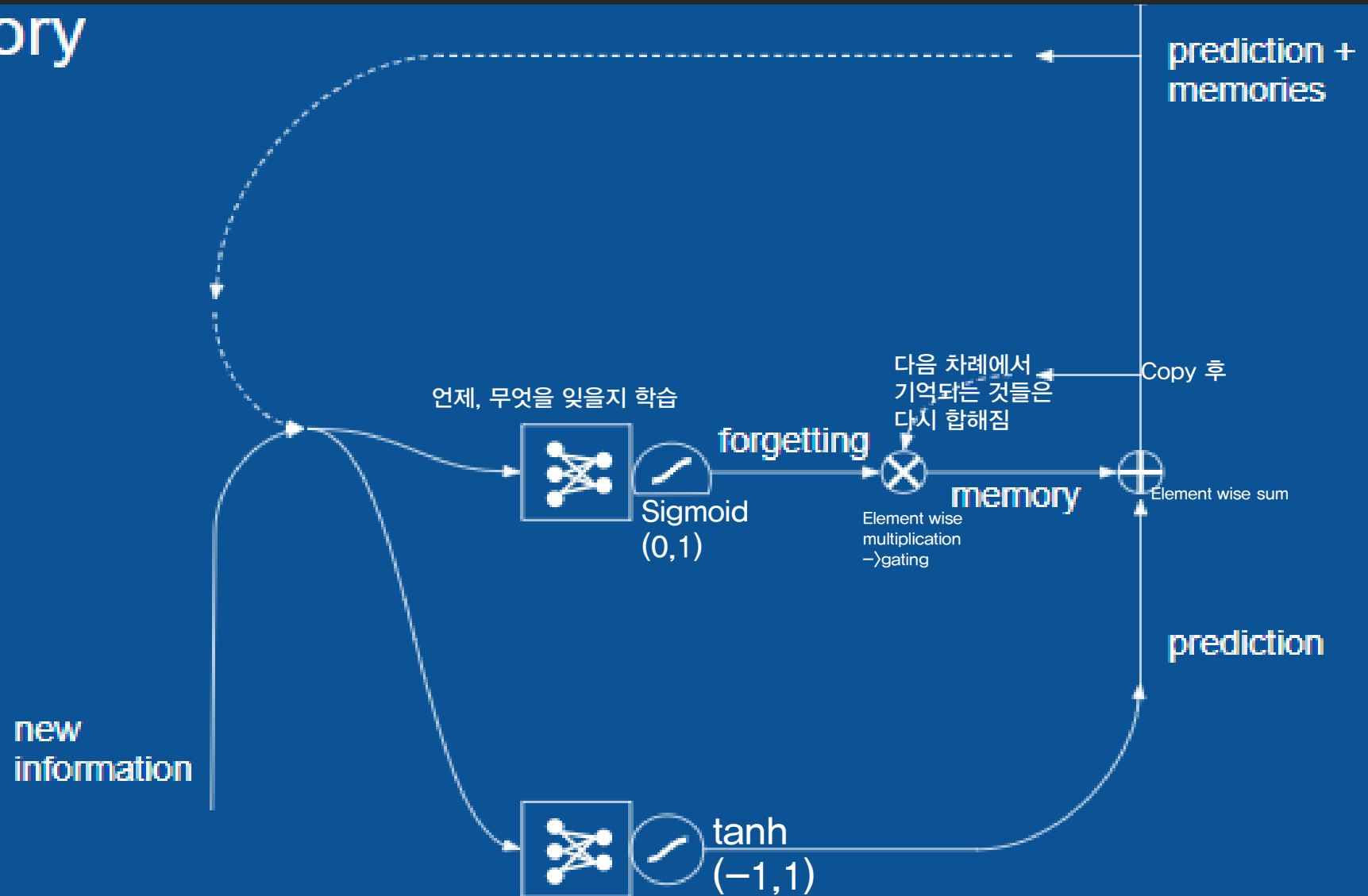
recurrent
neural
network

new
information



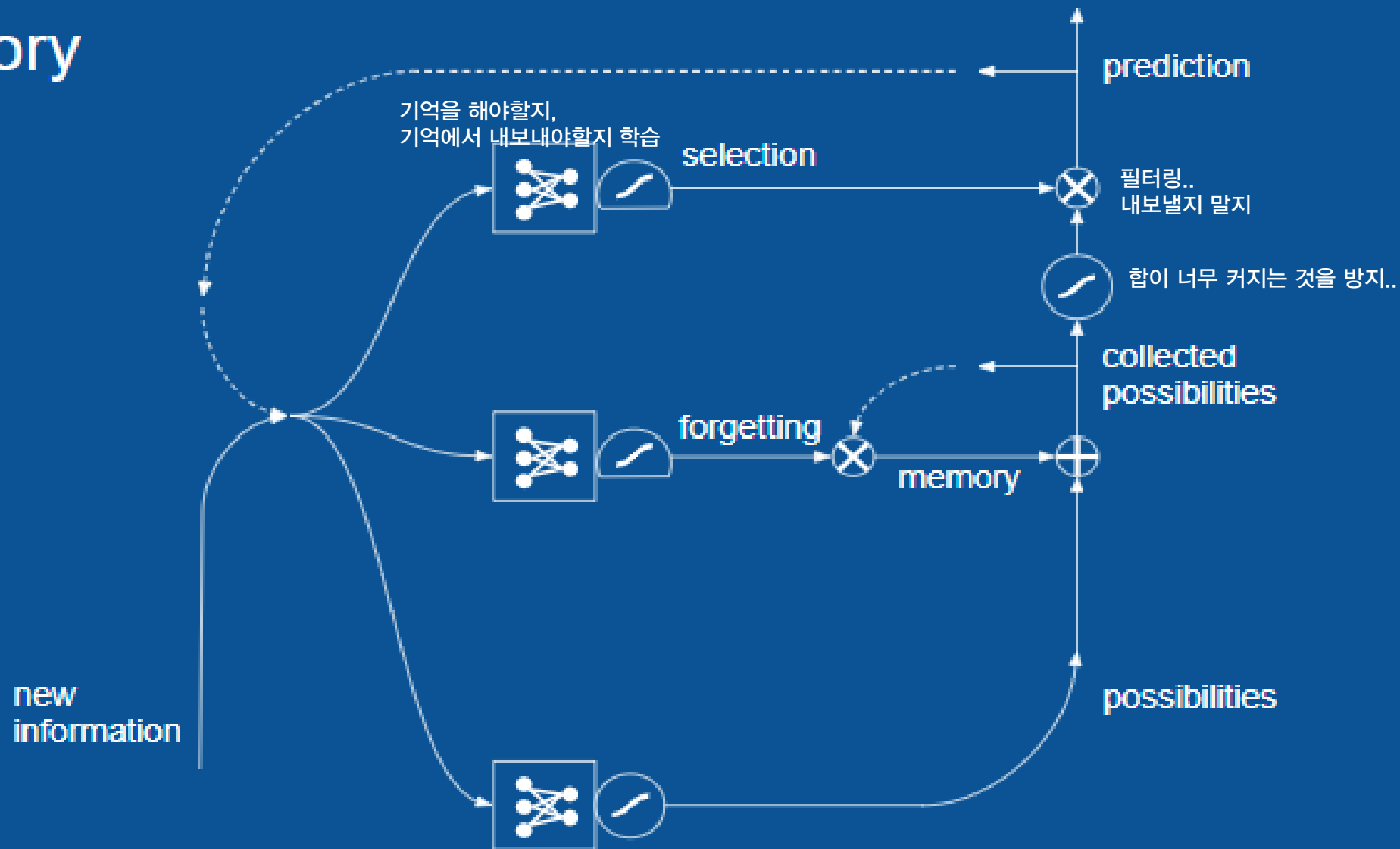
Unit 03 | RNN과 LSTM의 예제

memory



Unit 03 | RNN과 LSTM의 예제

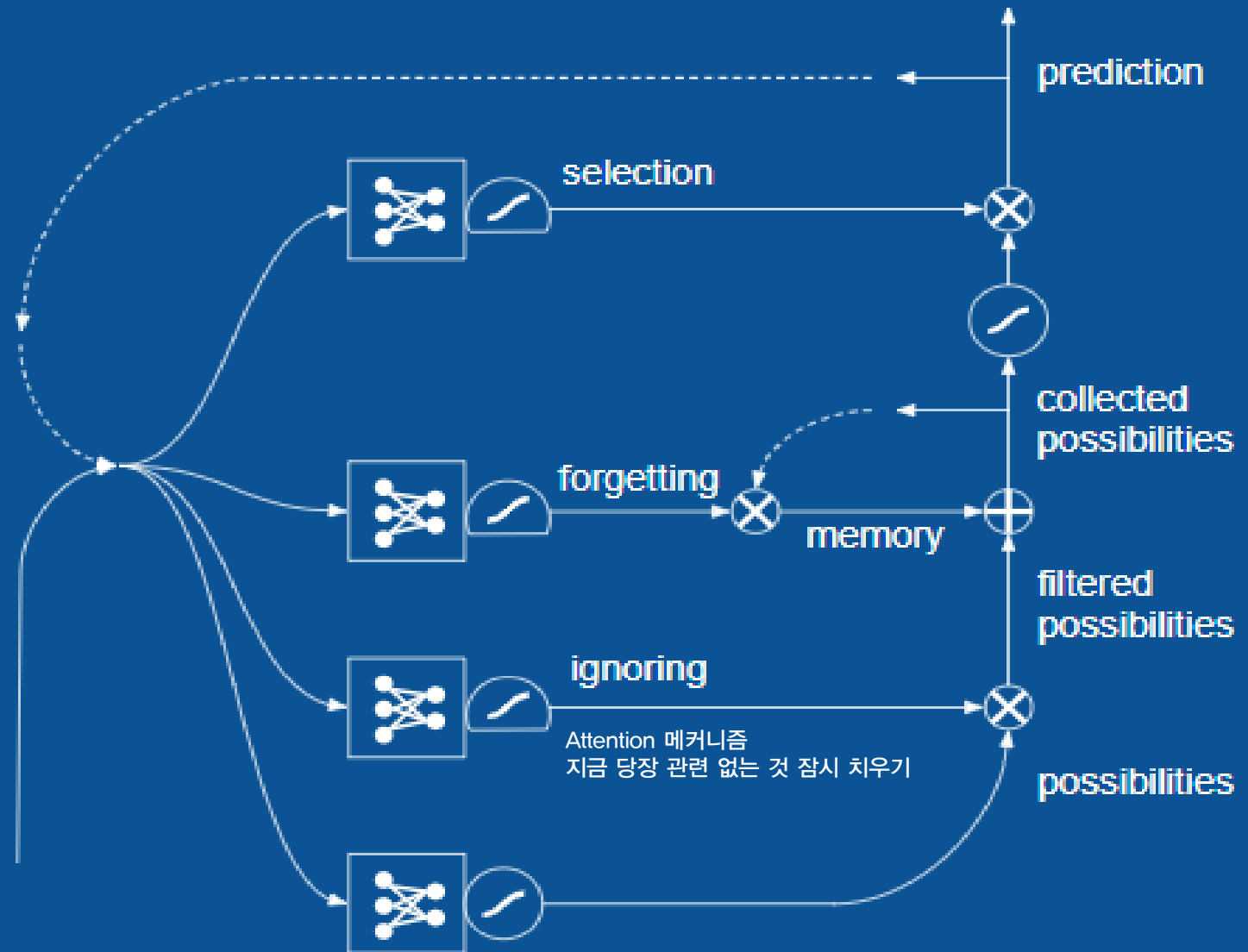
memory



Unit 03 | RNN과 LSTM의 예제

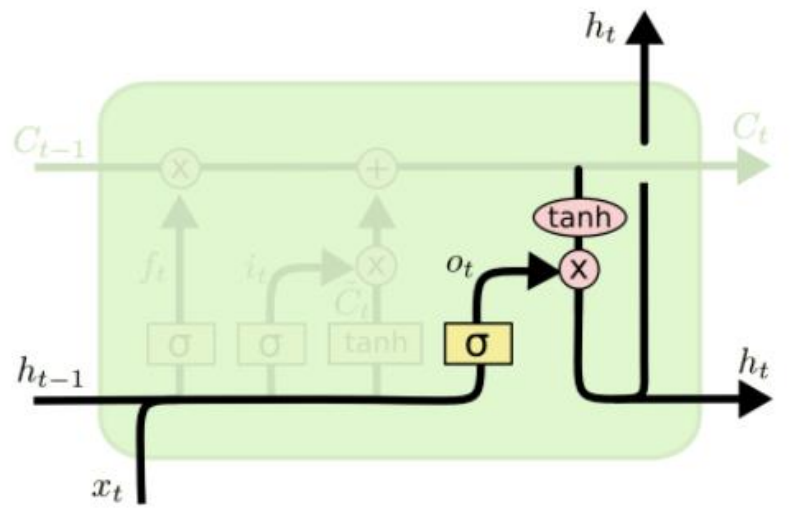
long
short-term
memory

new
information

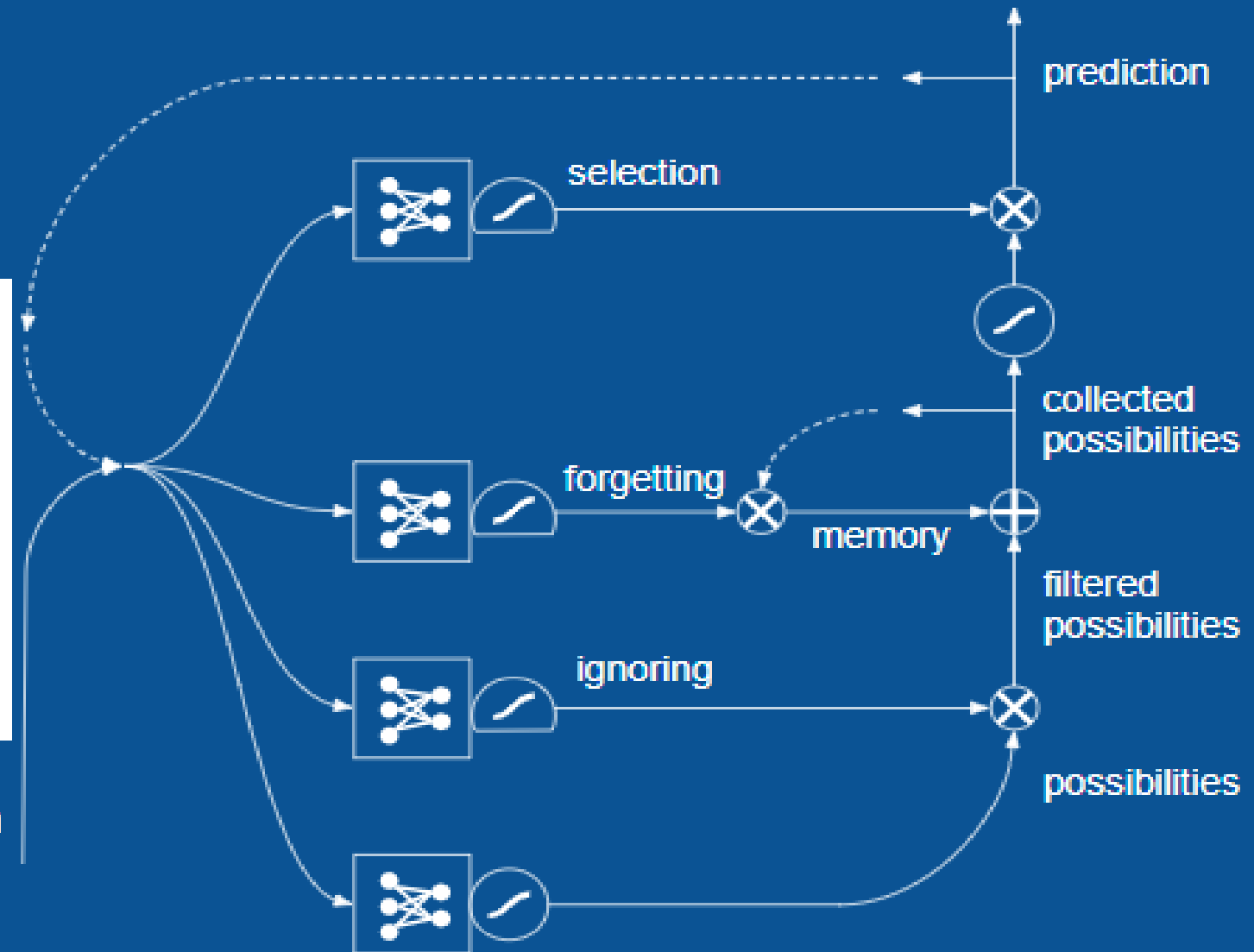


Unit 03 | RNN과 LSTM의 예제

long
short-term
memory



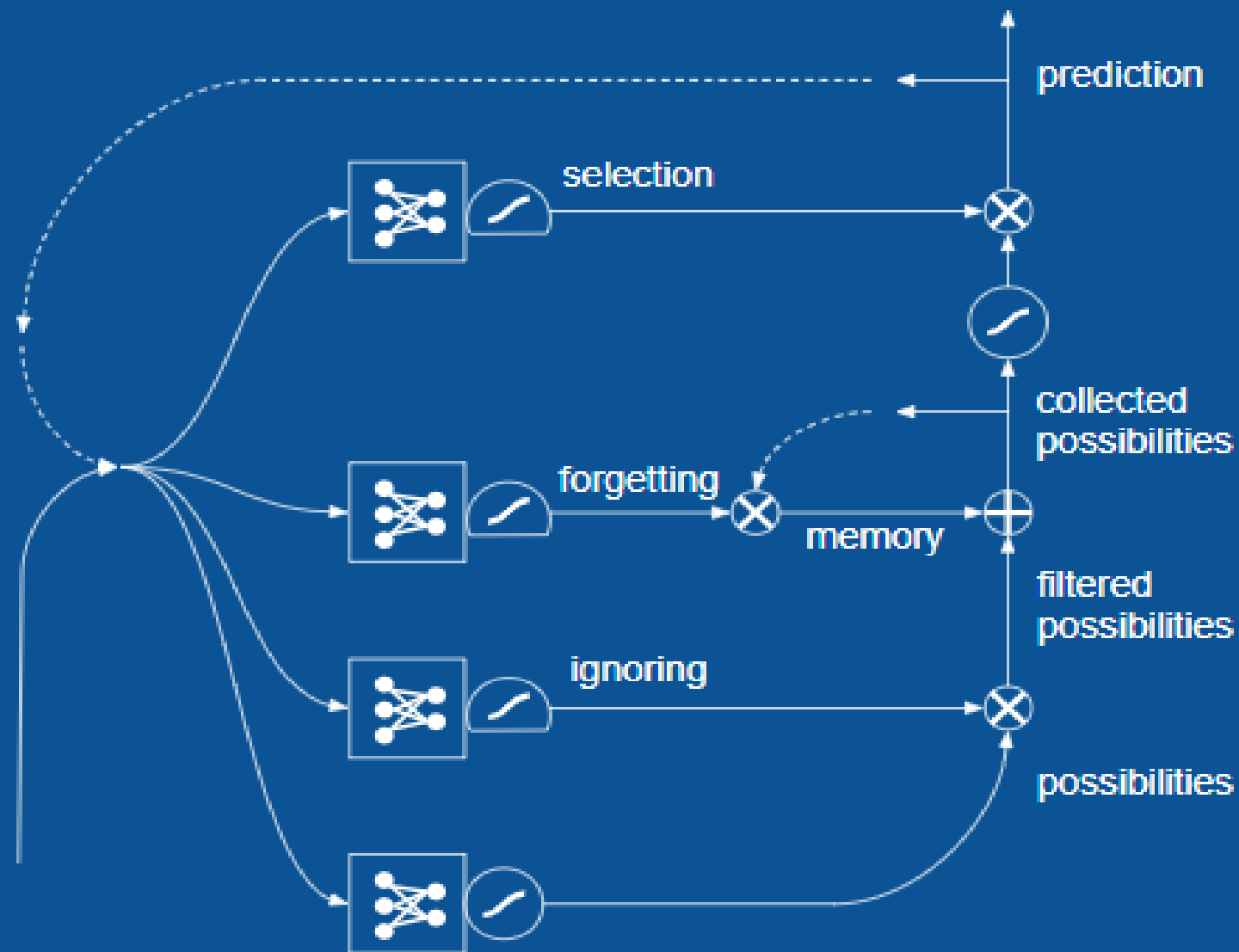
new
information



Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Jane saw Spot.
Doug ...

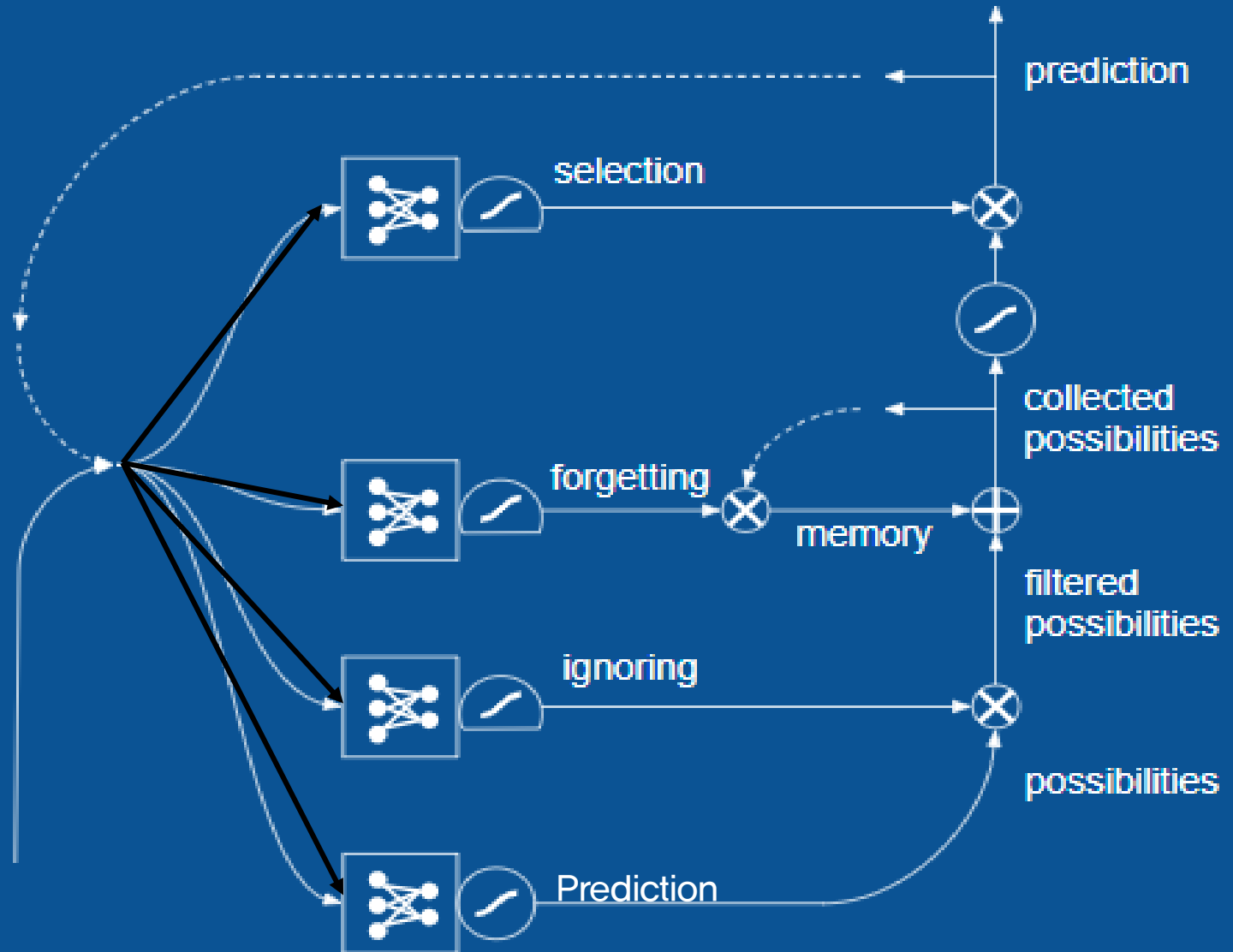


Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Doug,
Jane,
Spot

Jane saw Spot.
Doug ...

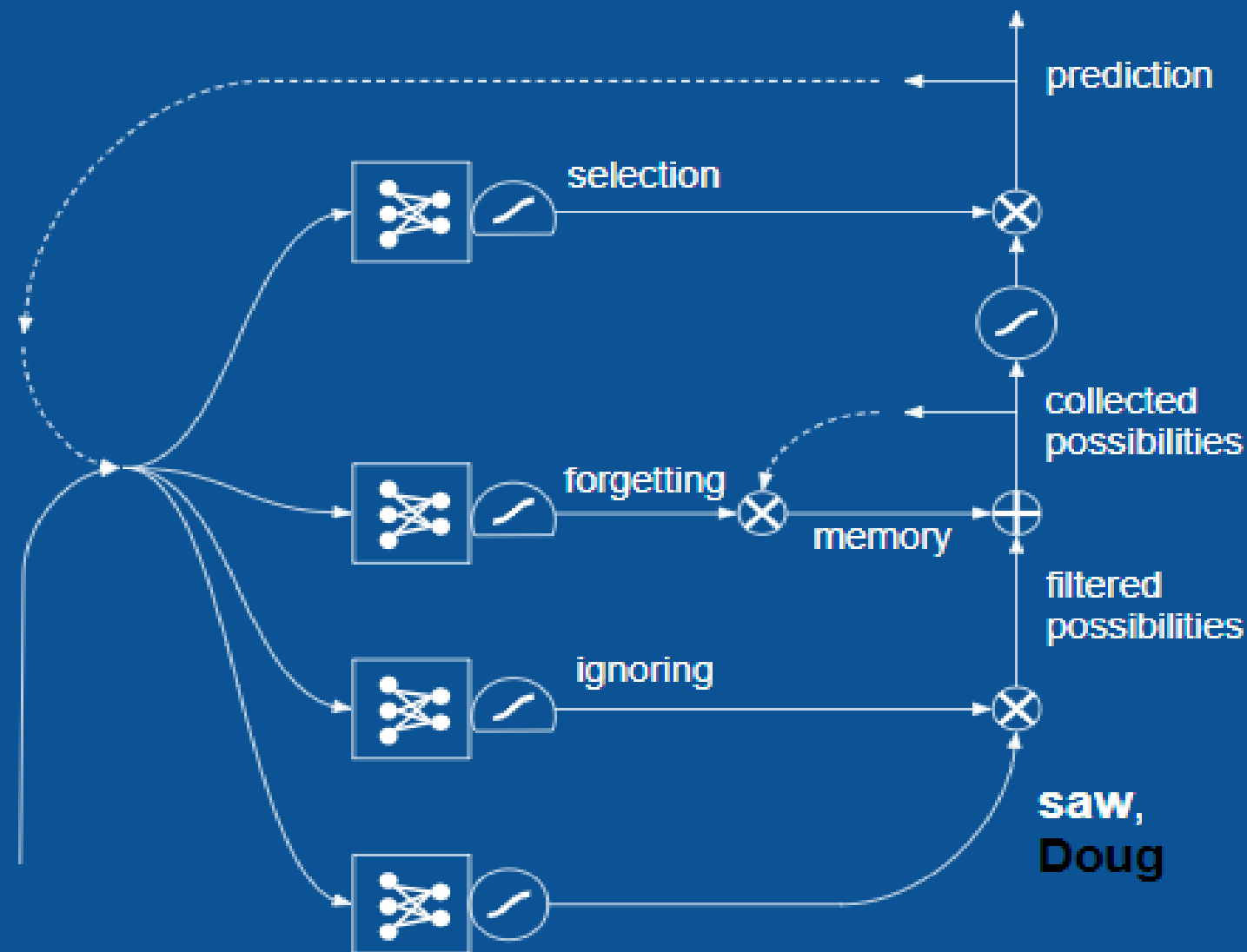


Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Doug,
Jane,
Spot

Jane saw Spot.
Doug ...



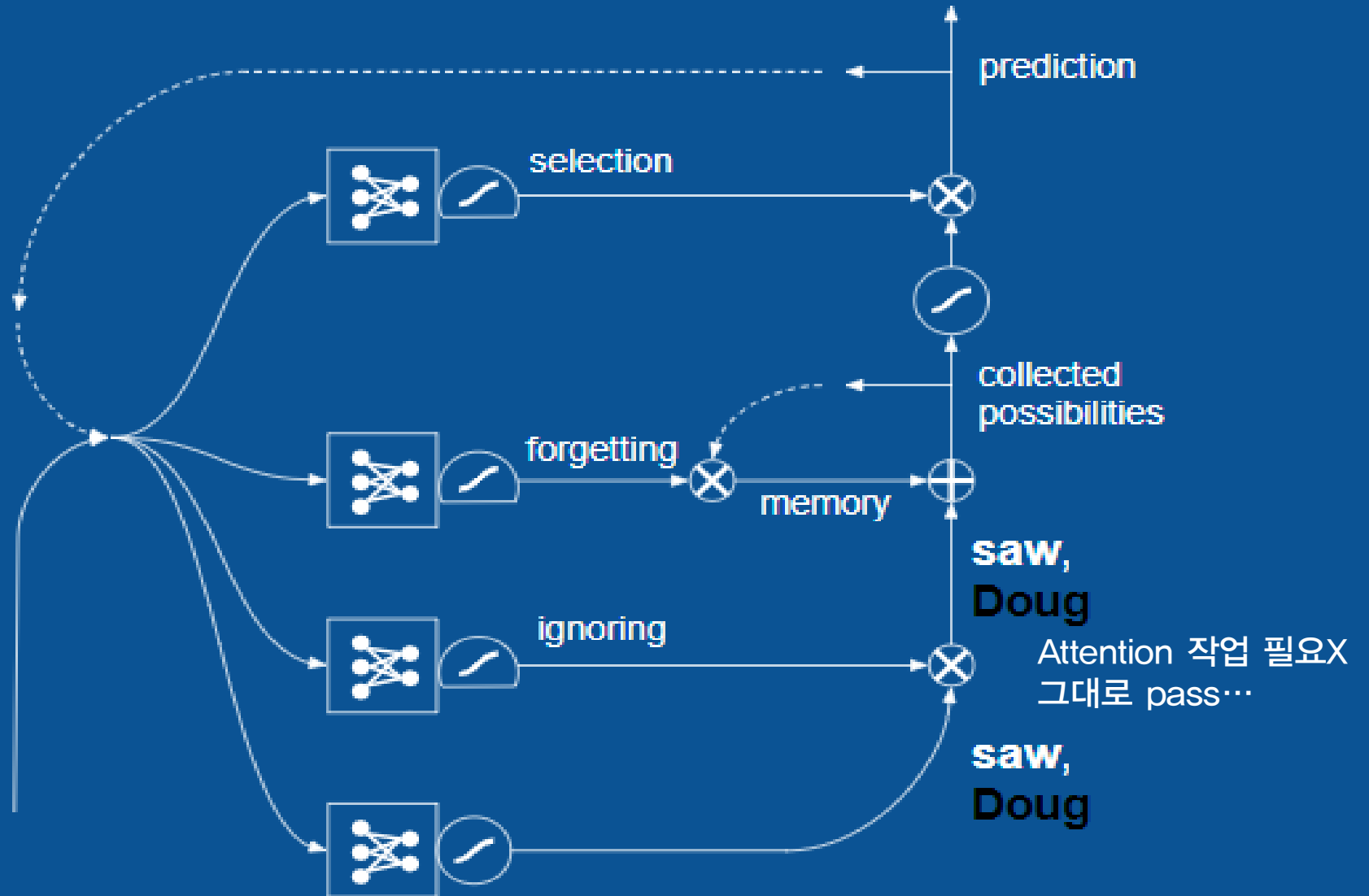
saw,
Doug

Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Doug,
Jane,
Spot

Jane saw Spot.
Doug ...

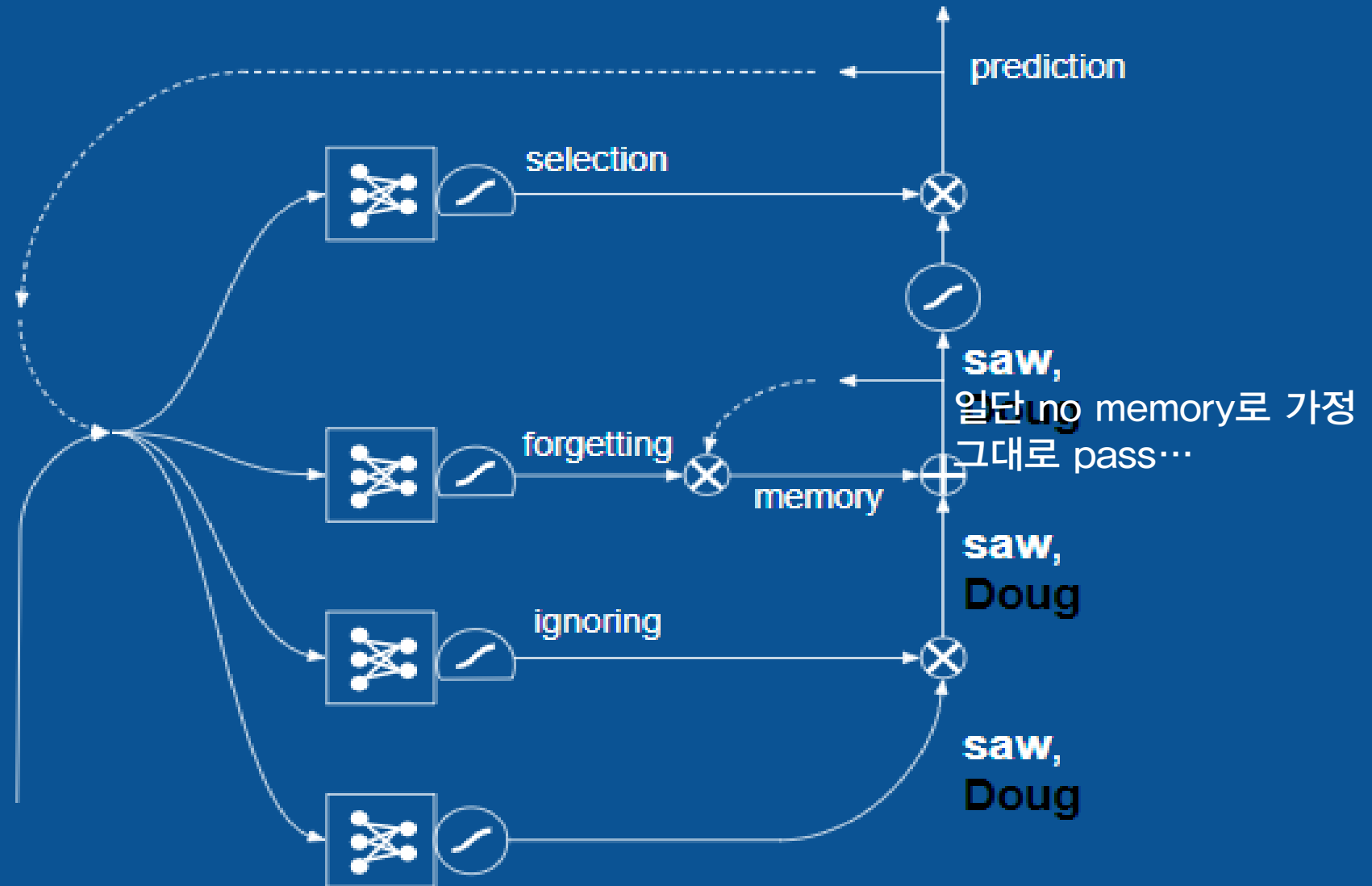


Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Doug,
Jane,
Spot

Jane saw Spot.
Doug ...

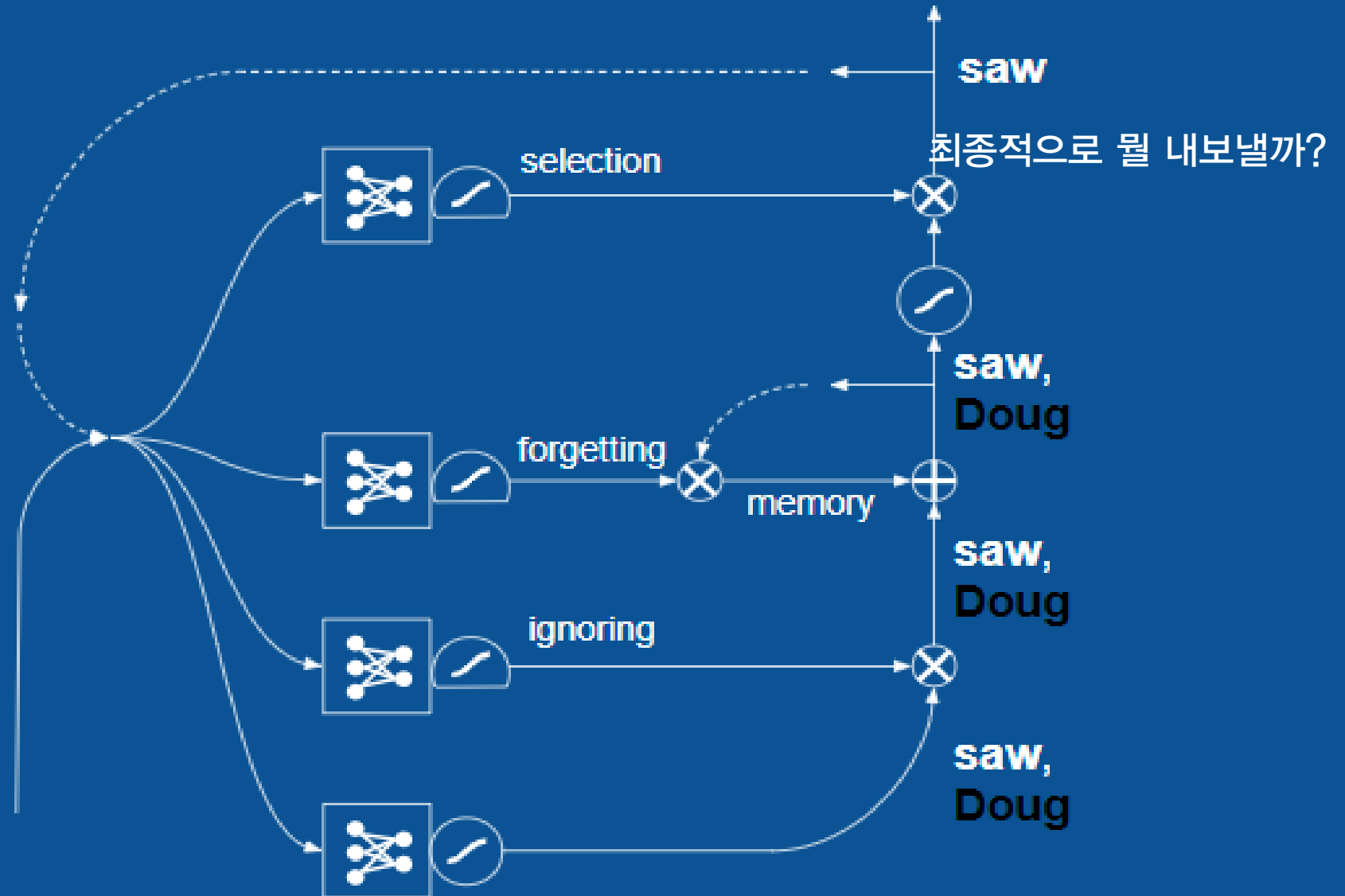


Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

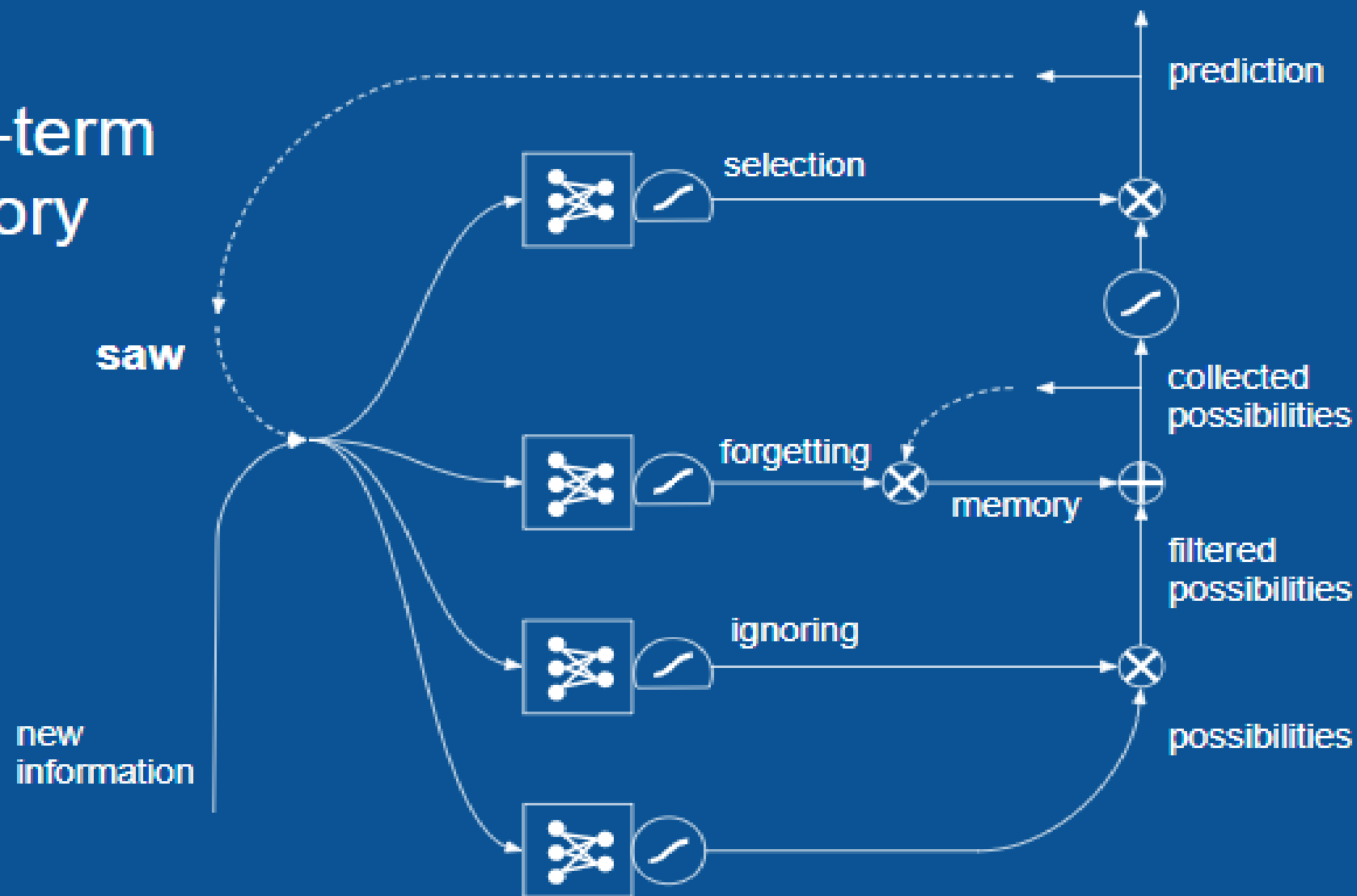
Doug,
Jane,
Spot

Jane saw Spot.
Doug ...



Unit 03 | RNN과 LSTM의 예제

long
short-term
memory



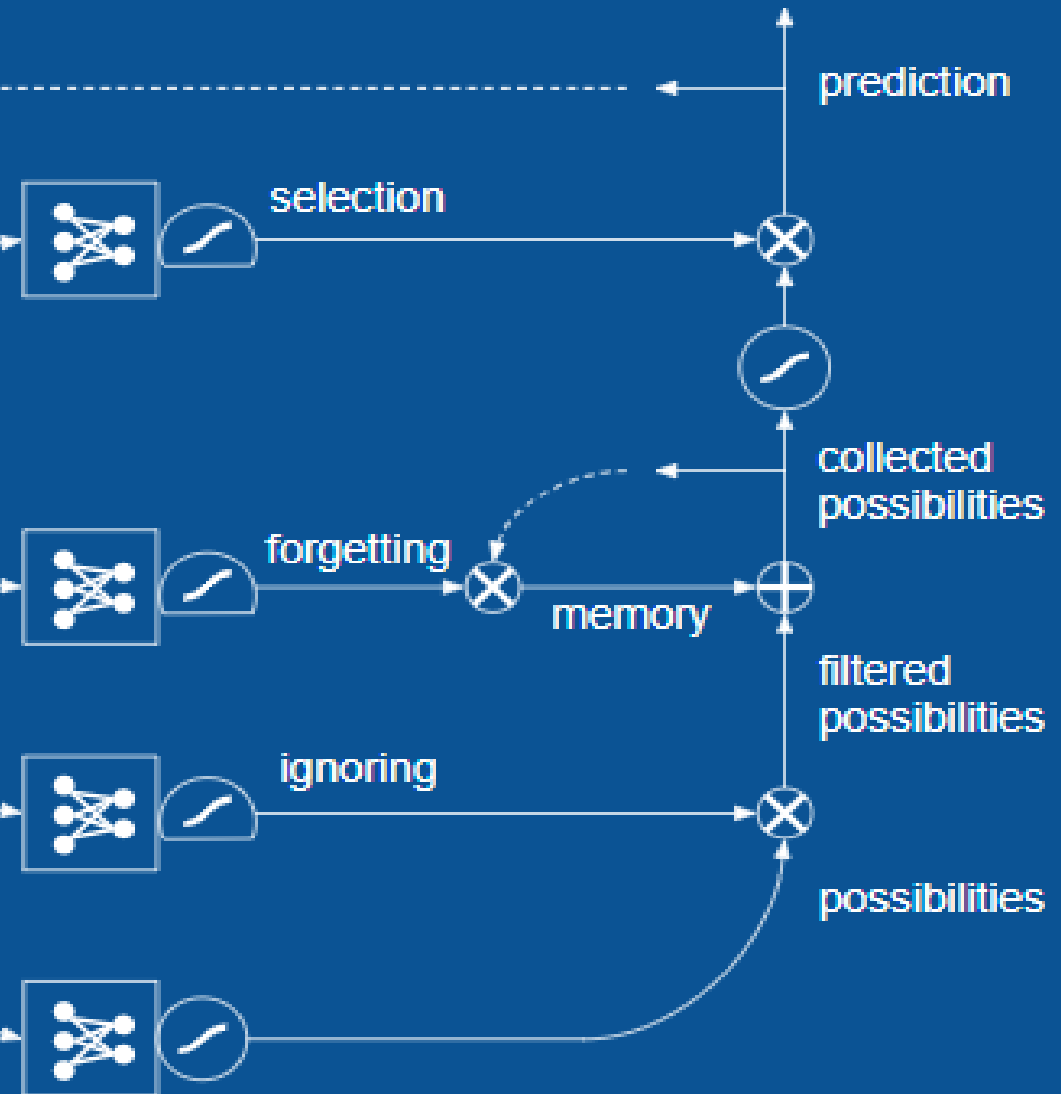
Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

saw

Jane saw Spot.
Doug **saw** ...

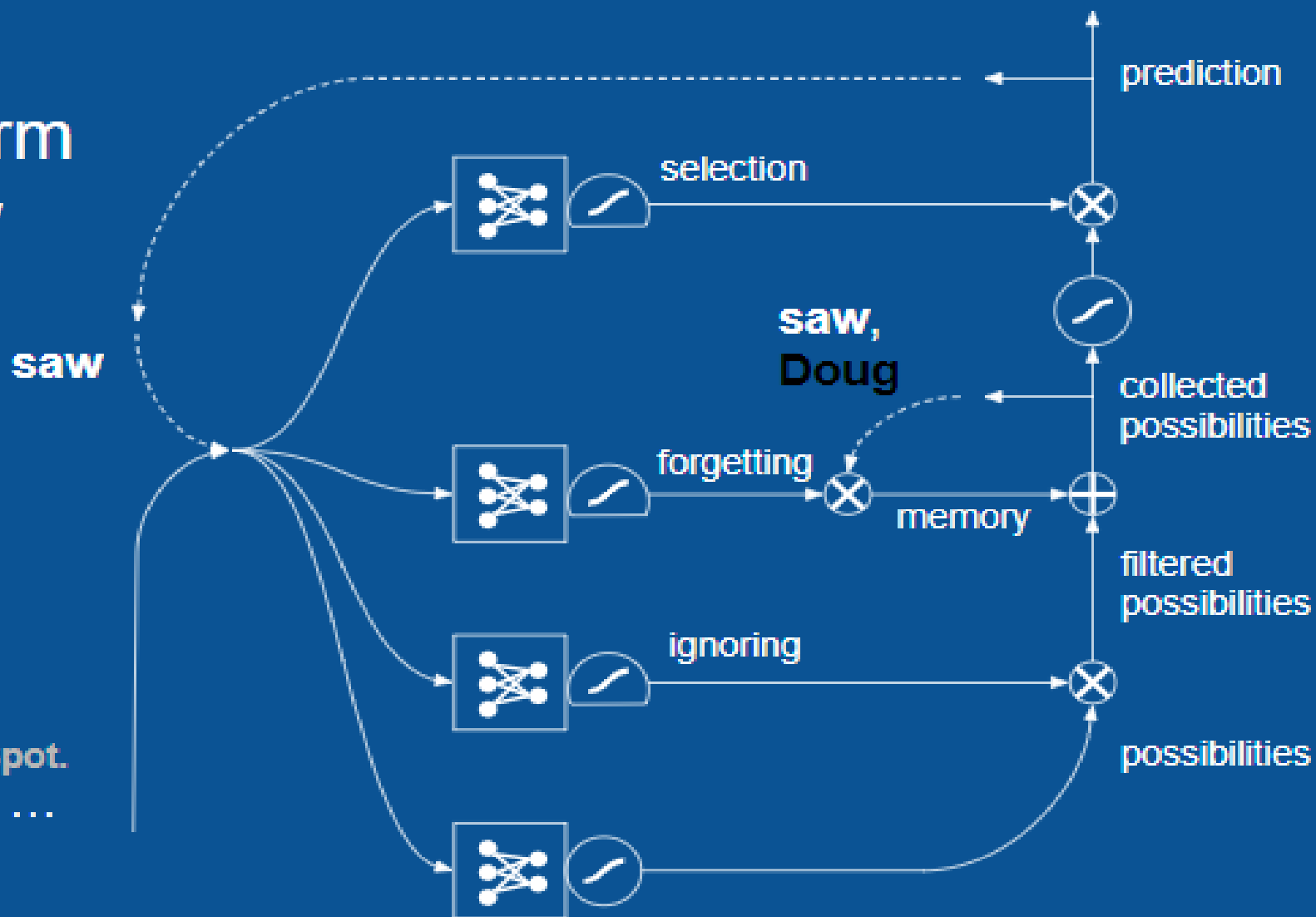
Doug, Jane, Spot 이 나오겠군..



Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

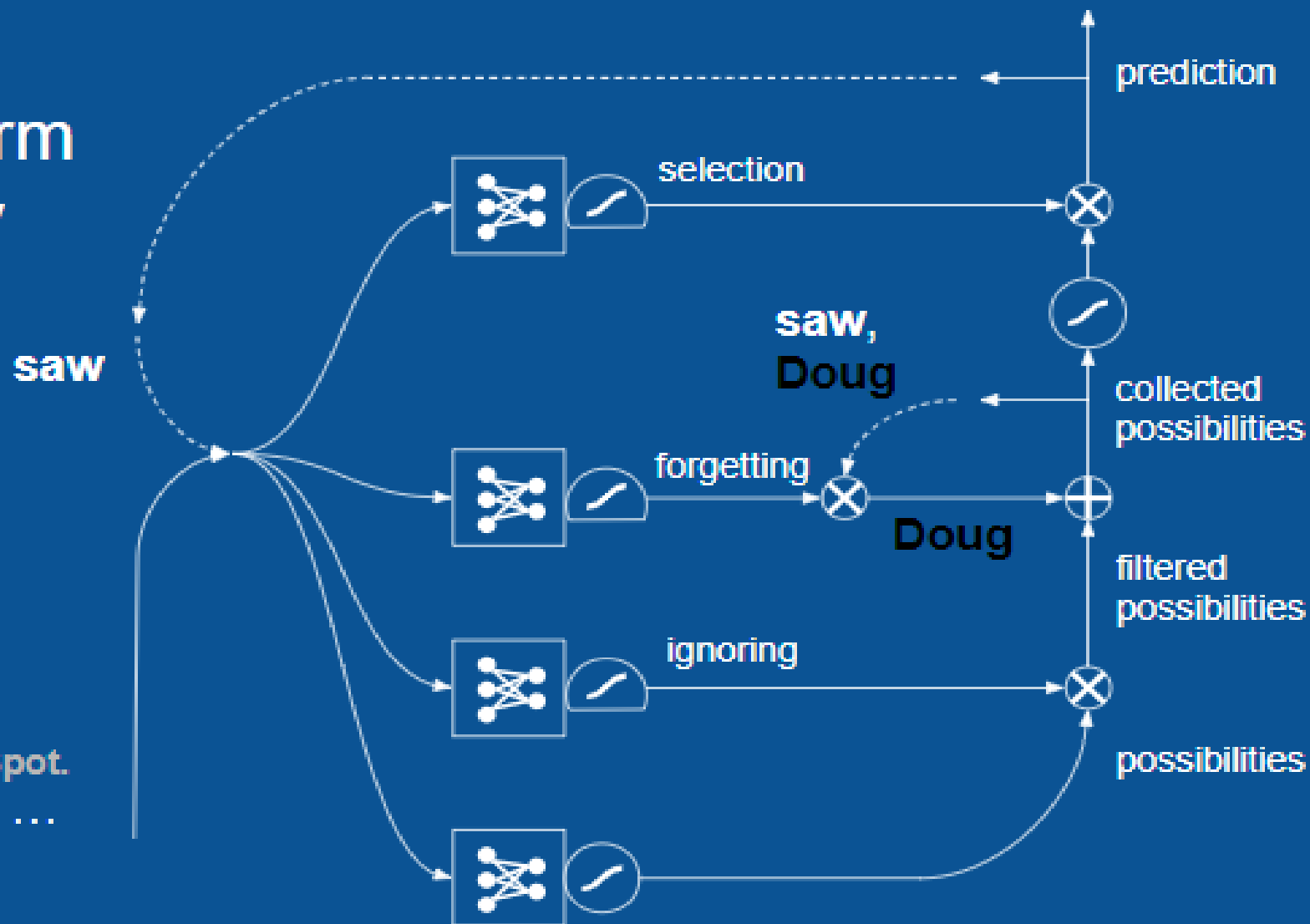
Jane saw Spot.
Doug **saw** ...



Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Jane saw Spot.
Doug **saw** ...

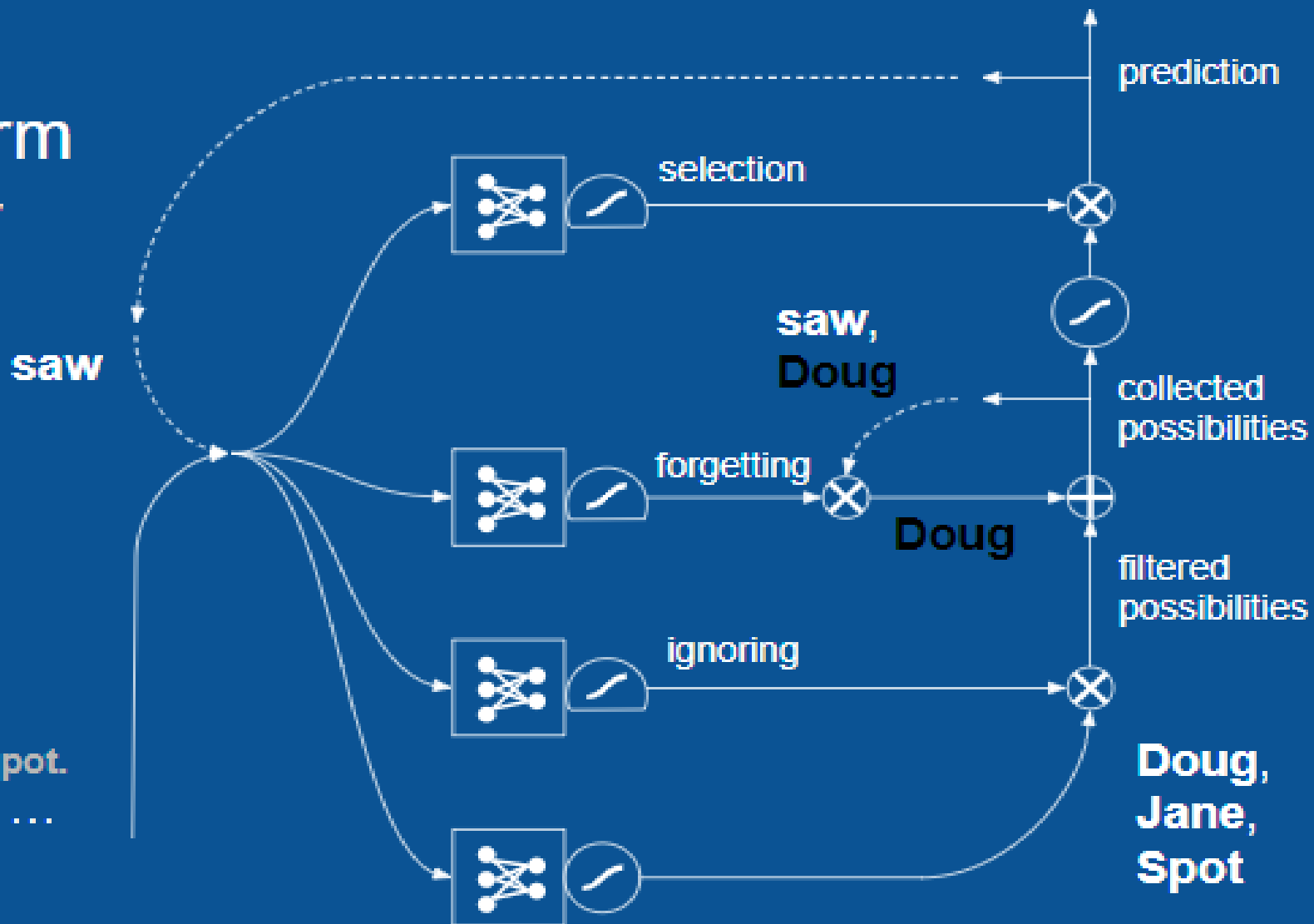


Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Jane saw Spot.
Doug **saw** ...

saw



Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Jane saw Spot.
Doug **saw** ...

saw

방금 saw를 내보냈음.
이름과 관련된 것을 기억할래

selection

forgetting

ignoring Attention 작업 필요
그대로 pass...

saw,
Doug

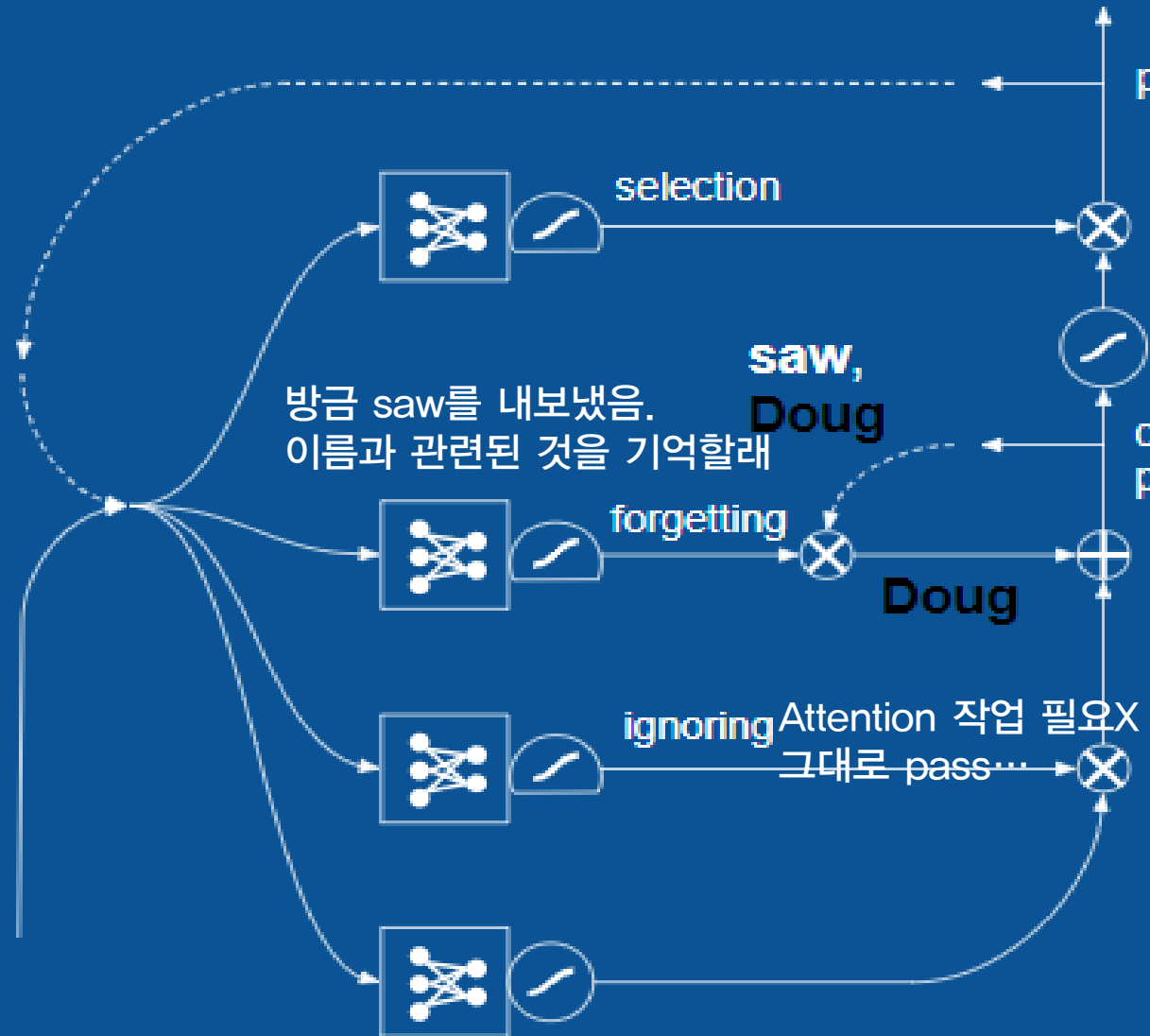
Doug

prediction

collected
possibilities

Doug,
Jane,
Spot

Doug,
Jane,
Spot

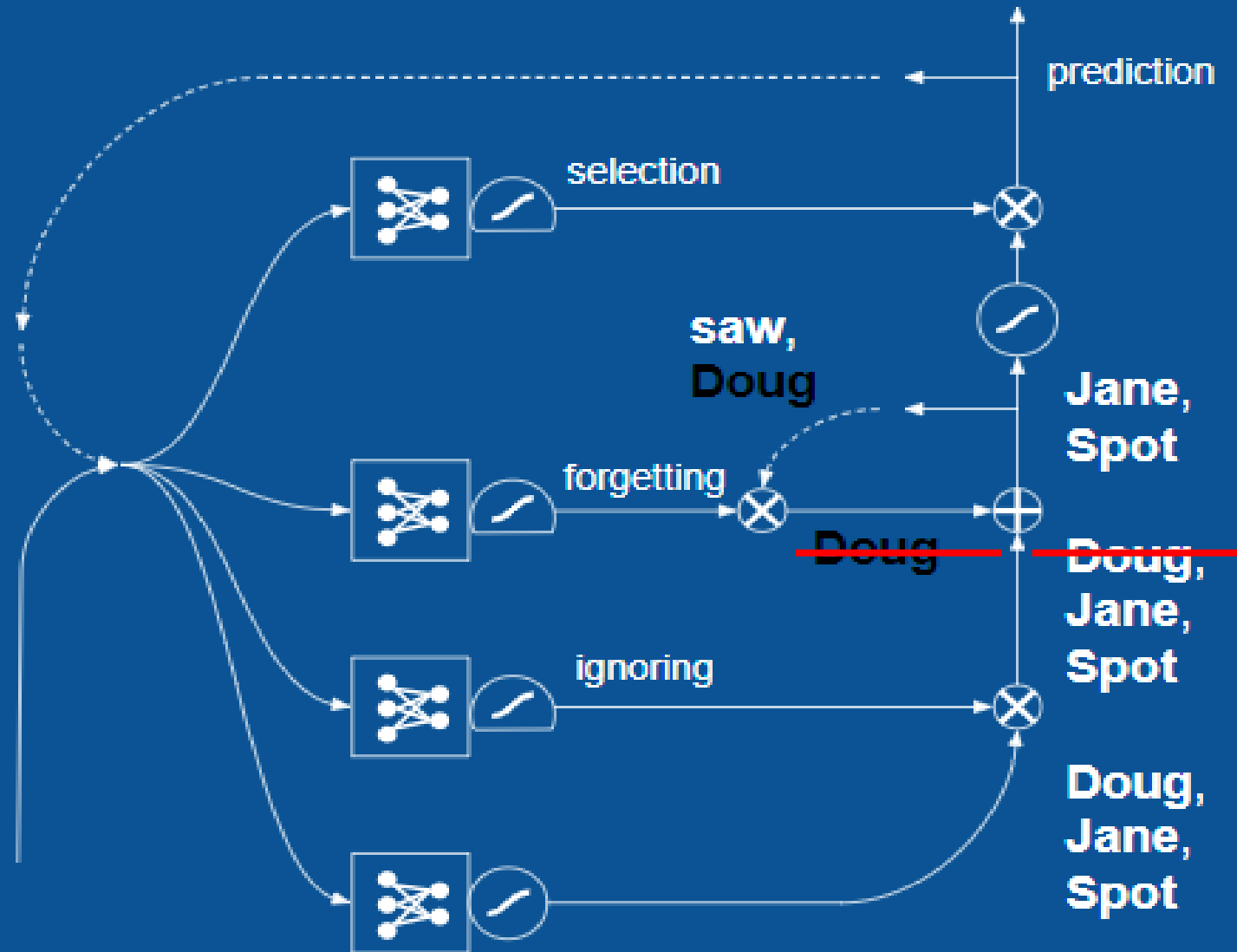


Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Jane saw Spot.
Doug **saw** ...

saw



Unit 03 | RNN과 LSTM의 예제

long
short-term
memory

Doug saw Doug..
방지..
물론, RNN도 가능

Jane saw Spot.
Doug **saw** ...

saw

selection

forgetting

ignoring

saw,
Doug

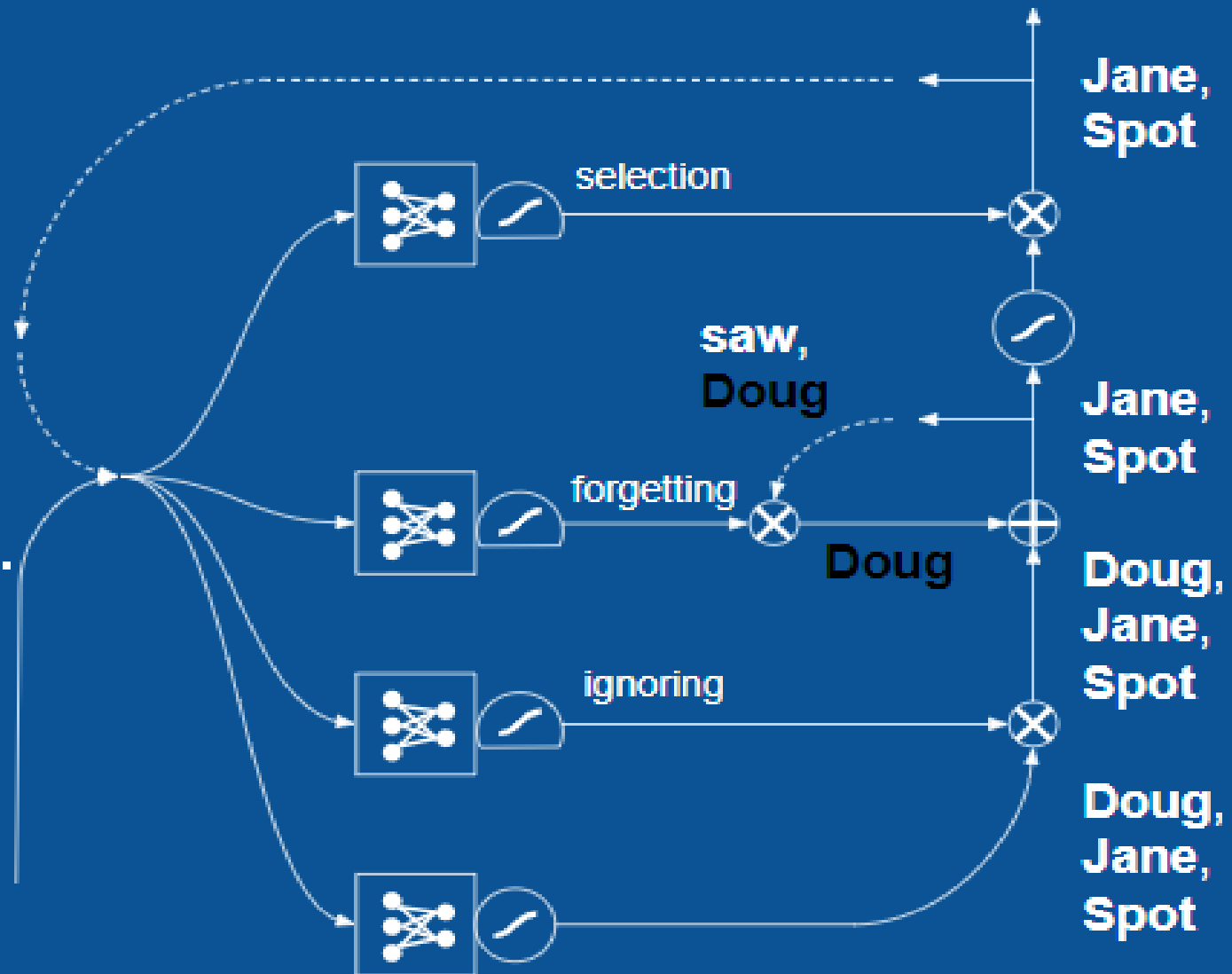
Doug

Jane,
Spot

Jane,
Spot

Doug,
Jane,
Spot

Doug,
Jane,
Spot





Q & A

들어주셔서 감사합니다.