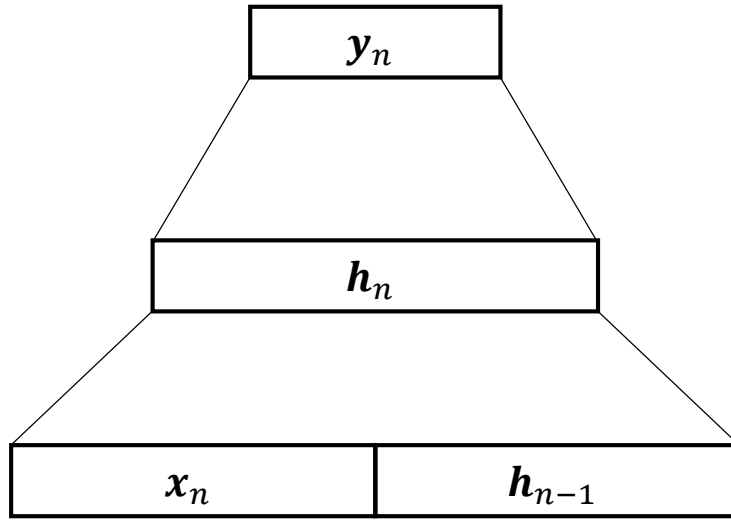
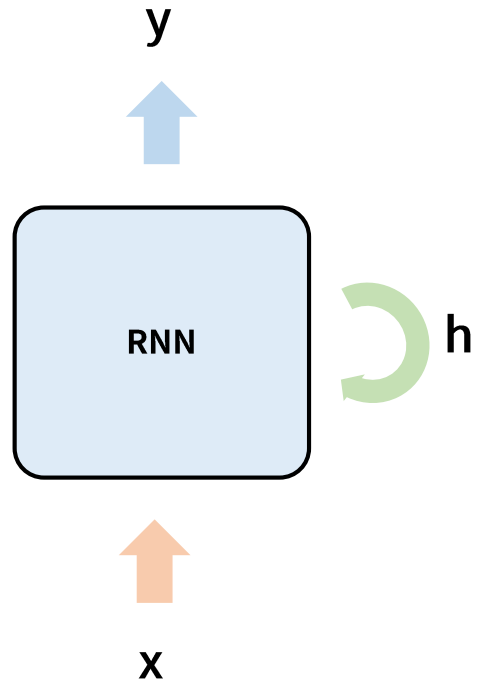


Chapter 06. 순환 신경망(RNN)

STEP1. 심화 순환 신경망

순환 신경망 Recurrent Neural Network

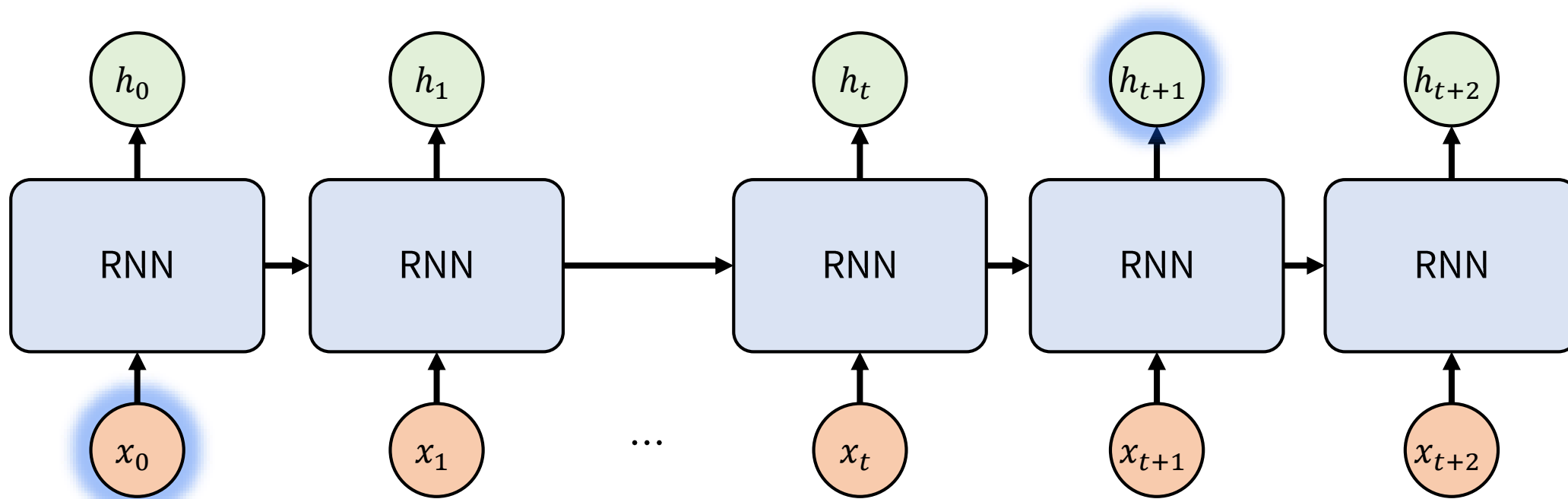


$$y_n = \tanh(W_o h_n + b_o)$$

$$h_n = \tanh(W_x x_n + W_h h_{n-1} + b_{in})$$

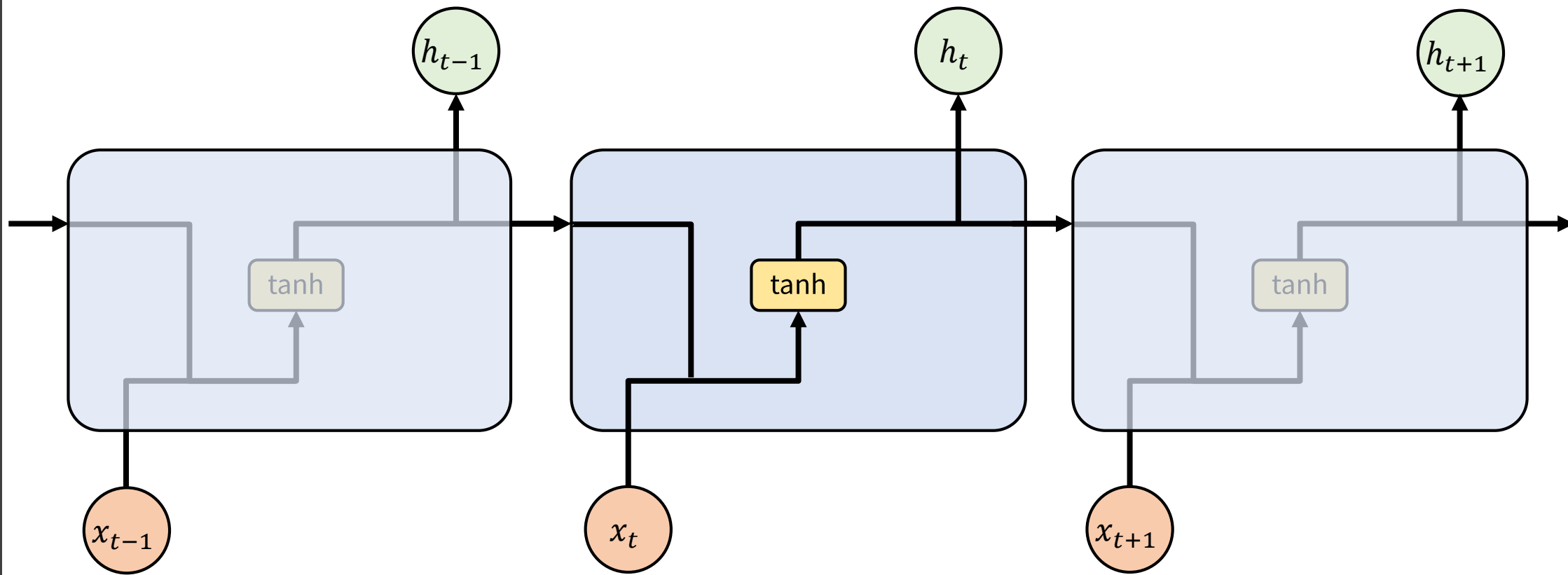
바닐라 순환 신경망에 맛깔 난 토핑을 얹은, 심화 순환 신경망의 구조와 동작을 이해한다.

기울기 소실 문제



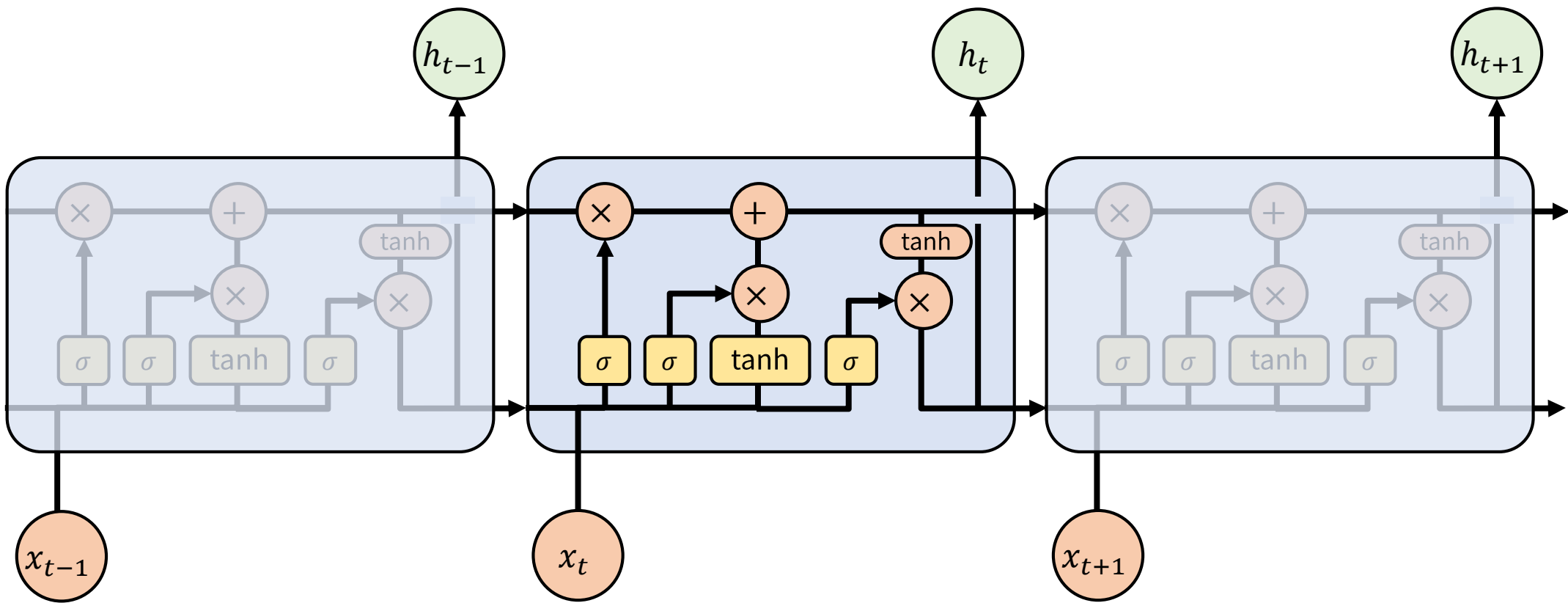
어떤 입력의 정보가 사용되는 **시점이 차이가 많이 날 경우**, 학습 능력이 저하된다.

Vanilla RNN



바닐라 RNN을 다른 형태로 표현하였다. 이 표현법을 기준으로 변형 알고리즘을 알아보자.

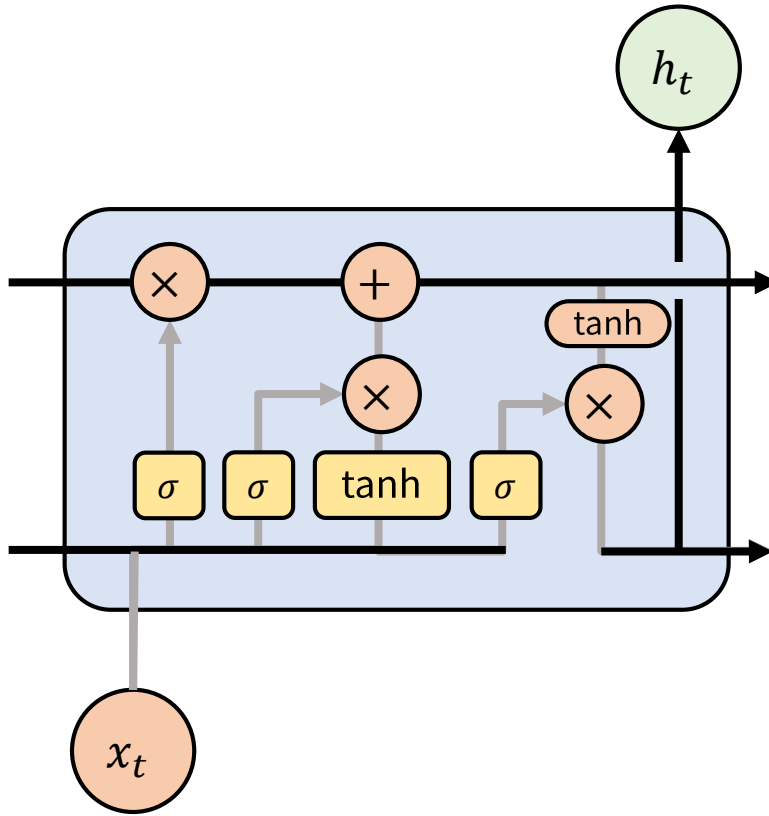
LSTM Long Short-Term Memory



Vanilla RNN을 개선한 LSTM 구조. 조금 복잡해 보이지만 하나씩 천천히 알아보자.

기억할 것은 오래 기억하고, 잊을 것은 빨리 잊어버리는 능력이 있다.

Cell State, Hidden State

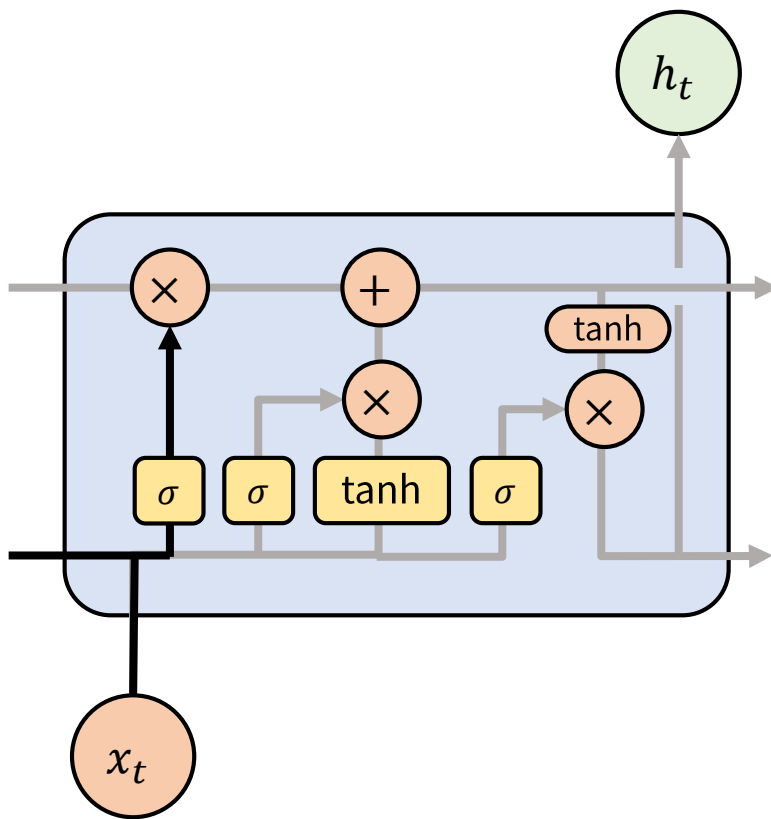


Cell State – 기억을 오랫동안 유지할 수 있는 구조
새로운 특징을 덧셈으로 받는 구조 (Residual Network)

Hidden State – 계층의 출력/다음 타임 스텝으로 넘기는 정보

RNN과 달리, Cell State가 있어서 ‘기억’에 관한 부분을 전담한다.

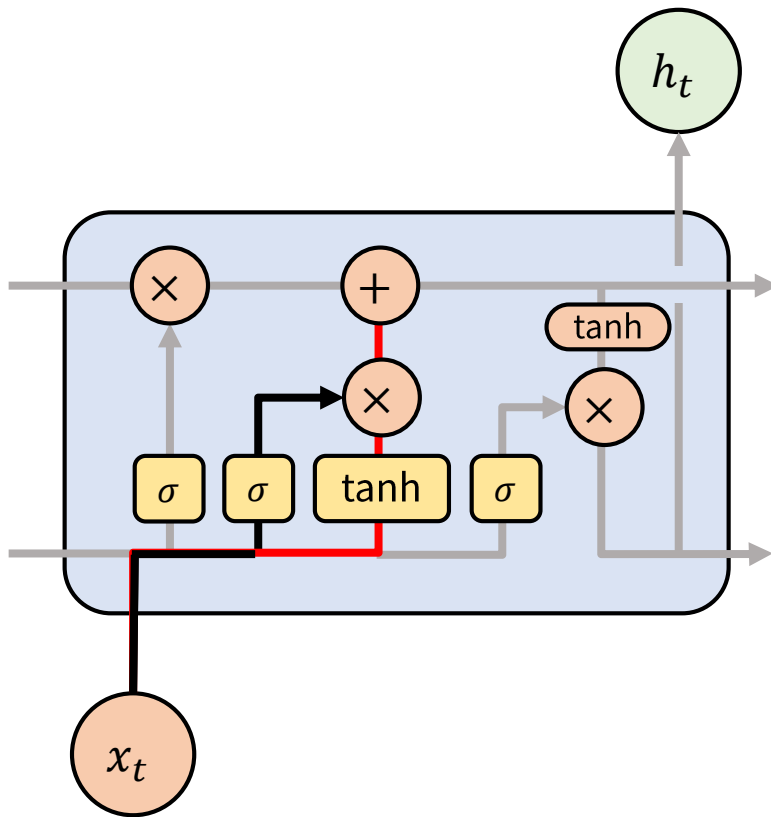
Forget Gate



Forget Gate

- Sigmoid 활성화 함수로, 0~1의 출력 값을 가짐
- Cell State에 이를 곱해 주어서 ‘얼만큼 잊을지’를 결정

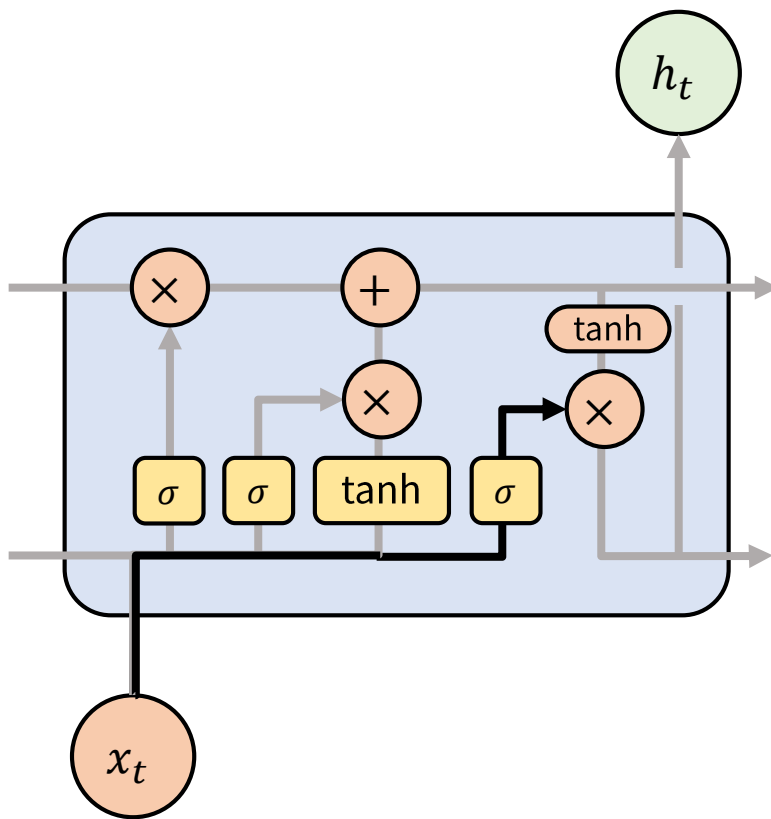
Input Gate



Input Gate

- Sigmoid 활성화 함수로, 0~1의 출력 값을 가짐
- 새롭게 추출한 특징을 얼마나 사용할 지 결정

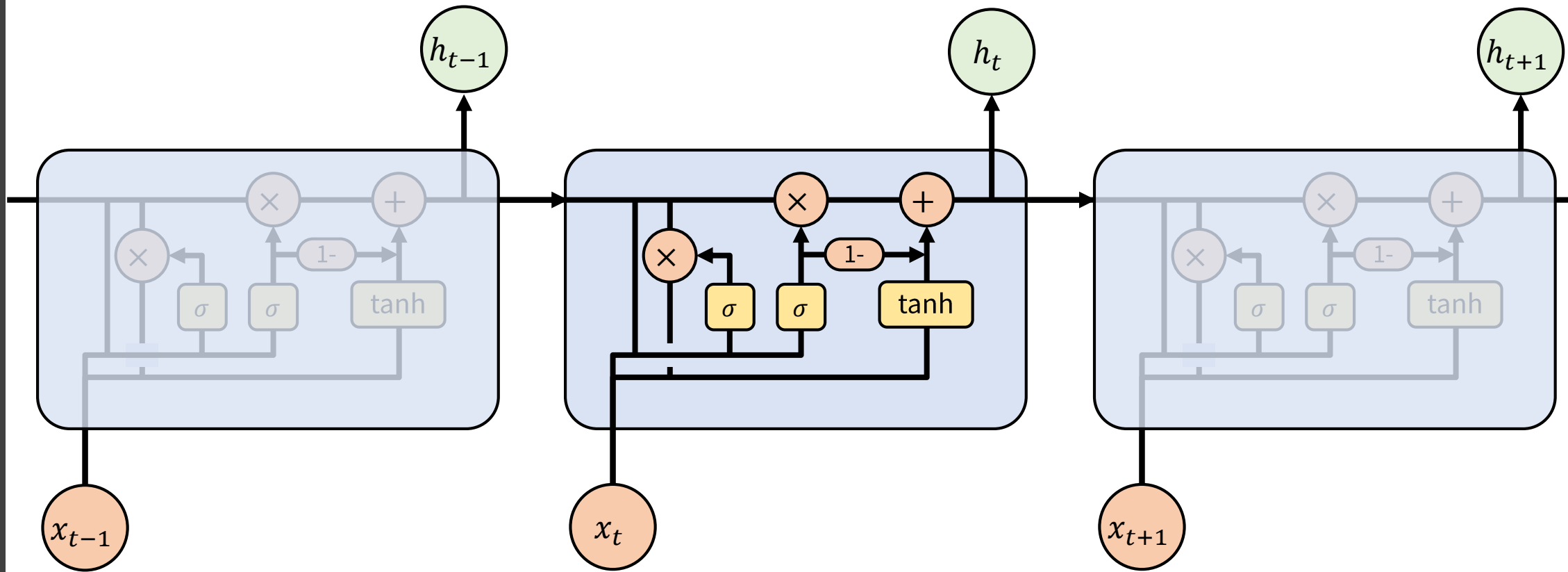
Output Gate



Output Gate

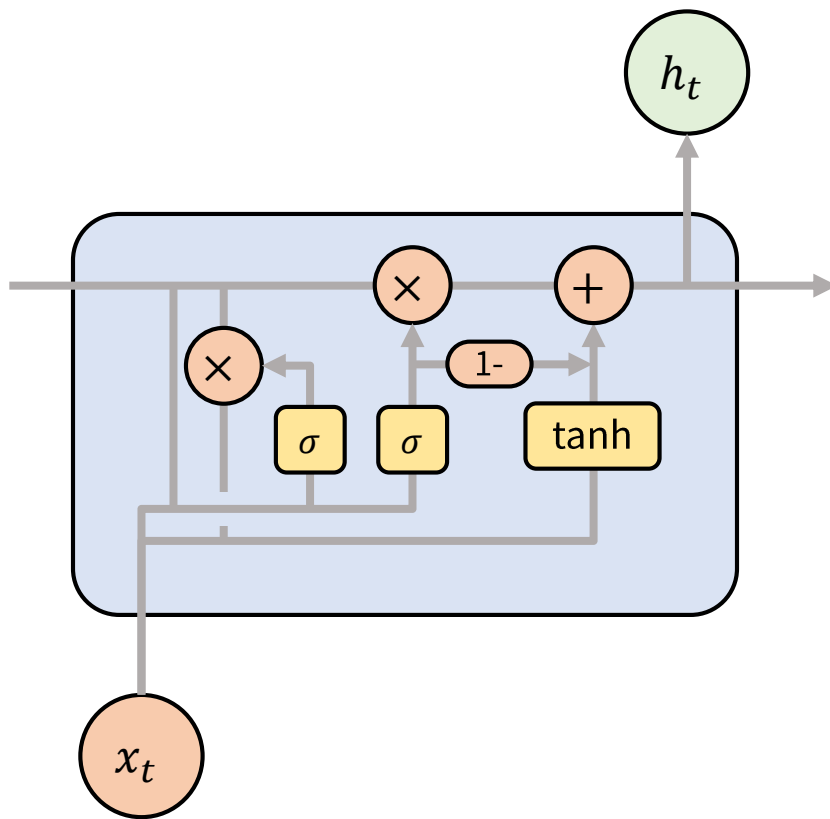
- Sigmoid 활성 함수로, 0~1의 출력 값을 가짐
- Cell로부터 **출력을 얼마나 내보낼지 결정**하는 역할

GRU Gated Recurrent Unit



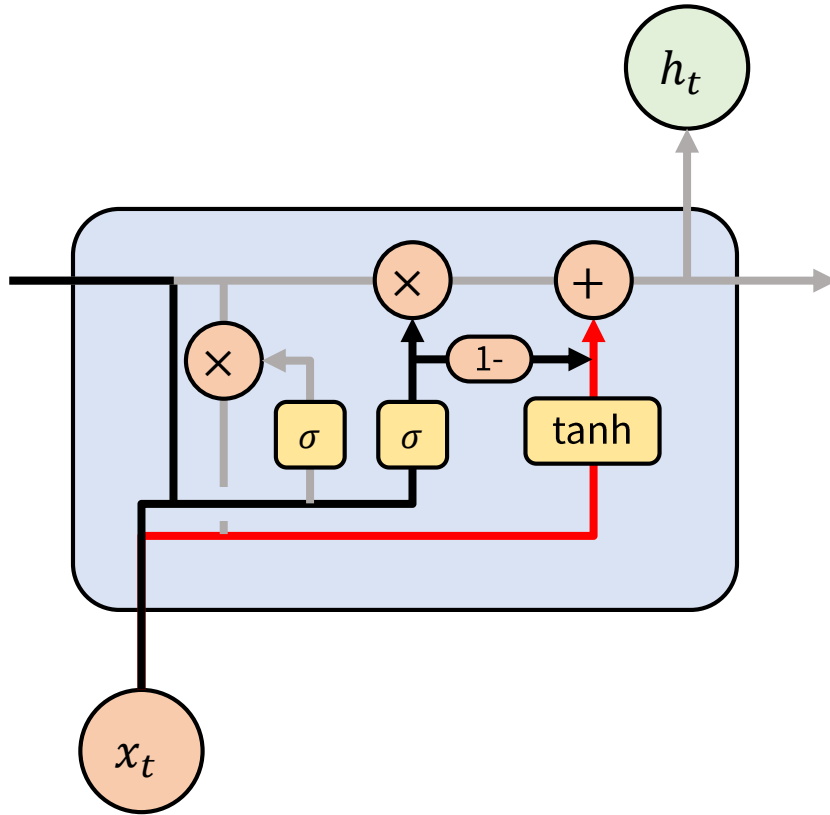
LSTM을 간소화한 버전이라고 할 수 있는 GRU. 마찬가지로 차근차근 보면 쉽게 의미를 알 수 있다.

GRU의 구조



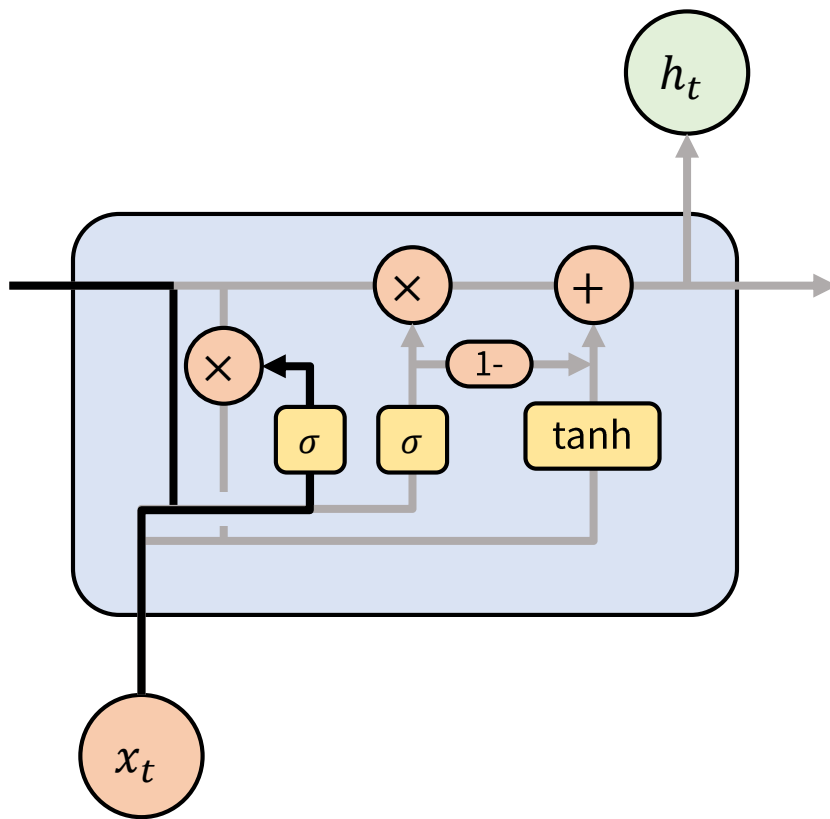
- Cell State가 없고, Hidden State만 존재한다.
- Forget Gate와 Input Gate를 결합하였다.
- Reset Gate를 추가하였다.

Forget Gate & Input Gate



- LSTM과 동일한 Forget Gate를 사용한다.
- Forget Gate를 1에서 빼서 Input Gate로 사용하였다.

Reset Gate



Reset Gate

- Sigmoid 활성화 함수로, 0~1의 값을 가진다.
- 이전 Hidden state를 얼마나 사용할지 정하는 역할
- 0에 가까운 값이 되면 'Reset'이 된다. (ex. 새 문장의 시작)