

Deep Learning-2

ICT Innovation Square

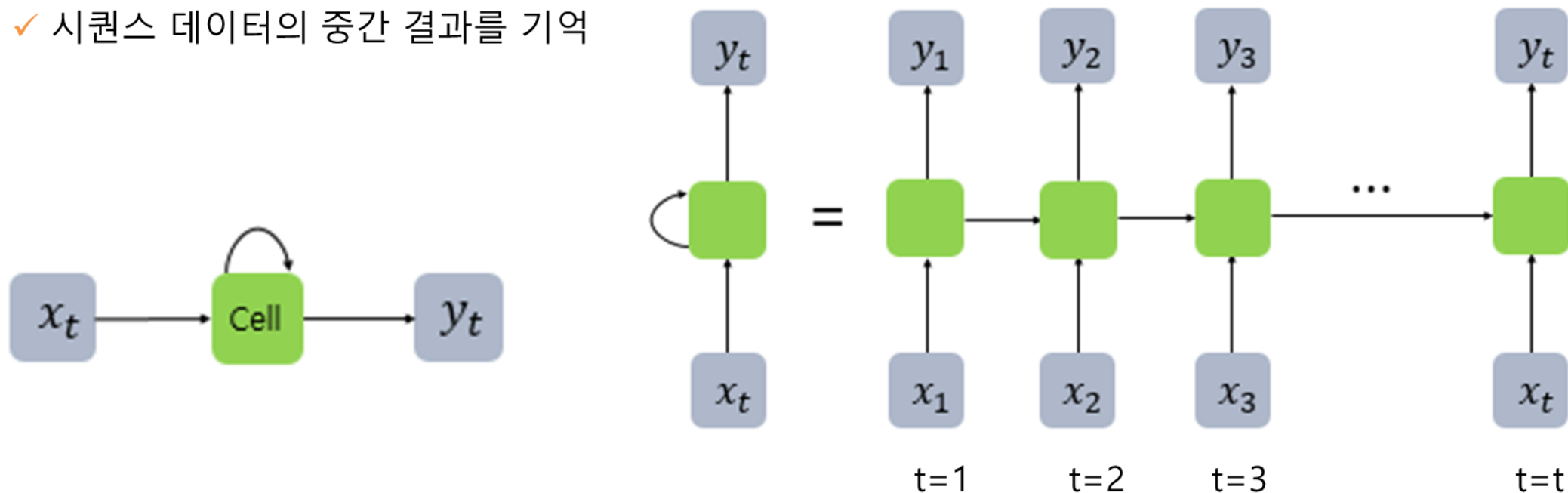
김 보 연

- LeNet
 - http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf
- AlexNet
 - <https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>
- ZFNet
 - <https://arxiv.org/abs/1311.2901>
- VGG-16
 - <https://arxiv.org/pdf/1409.1556.pdf>
- ResNet
 - <https://arxiv.org/pdf/1512.03385.pdf>
- Neural Network 3D Simulation
 - <https://www.youtube.com/watch?v=3JQ3hYko51Y>
- VGG16 Neural Network Visualization
 - <https://www.youtube.com/watch?v=RNnKtNrsmg>

- Object Detection
 - <https://arxiv.org/pdf/1809.02165.pdf>
- R-CNN
 - <https://arxiv.org/pdf/1311.2524.pdf>
- Fast-RCNN
 - <https://arxiv.org/abs/1504.08083>
- Faster-RCNN
 - <https://arxiv.org/pdf/1506.01497.pdf>
- YOLO v5
 - <https://github.com/ultralytics/yolov5>
- FCN
 - https://openaccess.thecvf.com/content_cvpr_2015/papers/Long_Fully_Convolutional_Networks_2015_CVPR_paper.pdf
- U-net
 - https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28

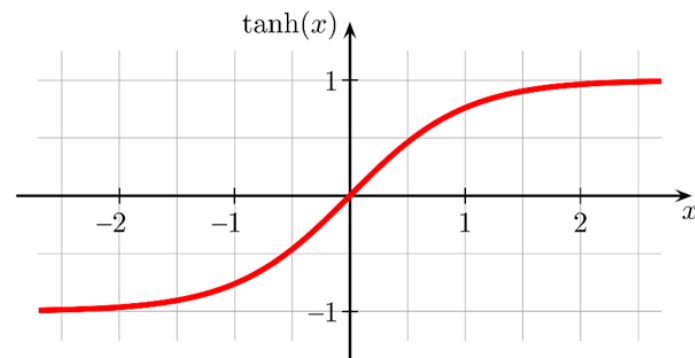
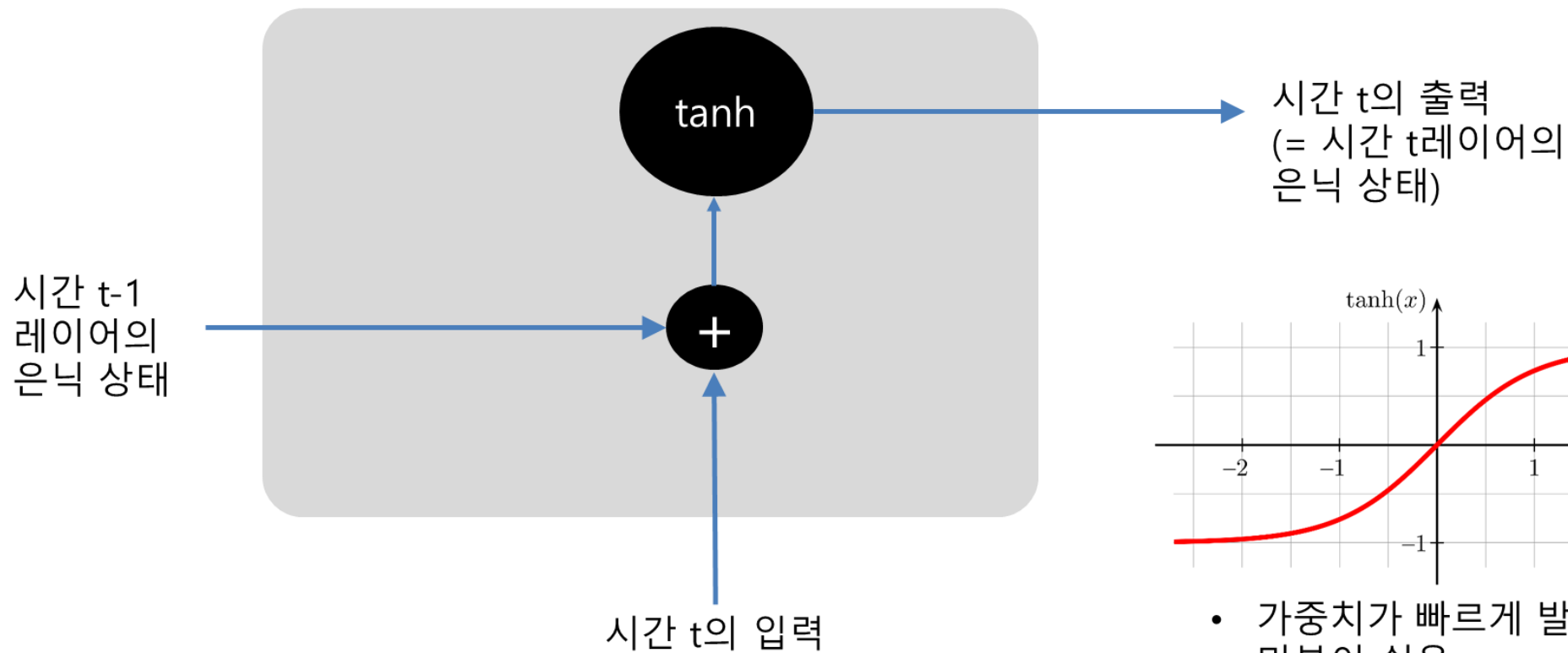
■ 순환 신경망(RNN)

- 주어진 단어 시퀀스를 한 단어씩 처리
- 각 레이어의 출력을 다음 레이어로 전달
- 은닉 상태(hidden state)
 - ✓ 다음 레이어로 이어지는 중간 출력
 - ✓ 시퀀스 데이터의 중간 결과를 기억



<https://www.youtube.com/watch?v=PahF2hZM6cs>

■ 내부 구조



- 가중치가 빠르게 발산하지 않음
- 미분이 쉬움

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$\tanh'(x) = 1 - \tanh^2(x)$$

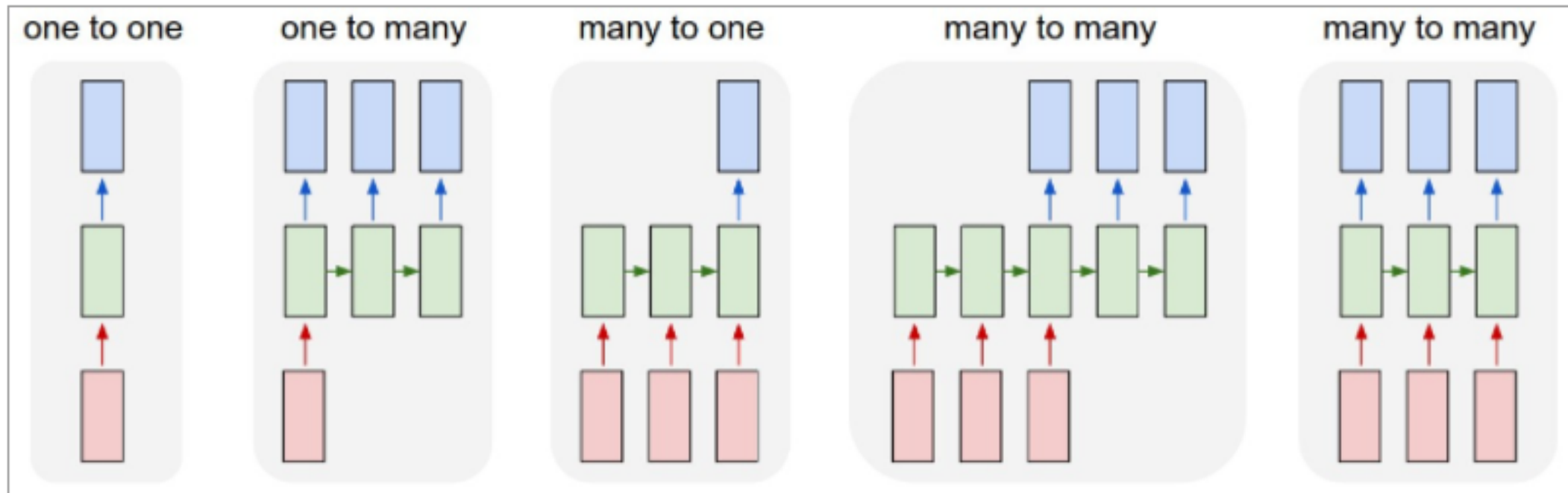
(1)MLP

(2)모델 2

(3)모델 1

(4)모델 3

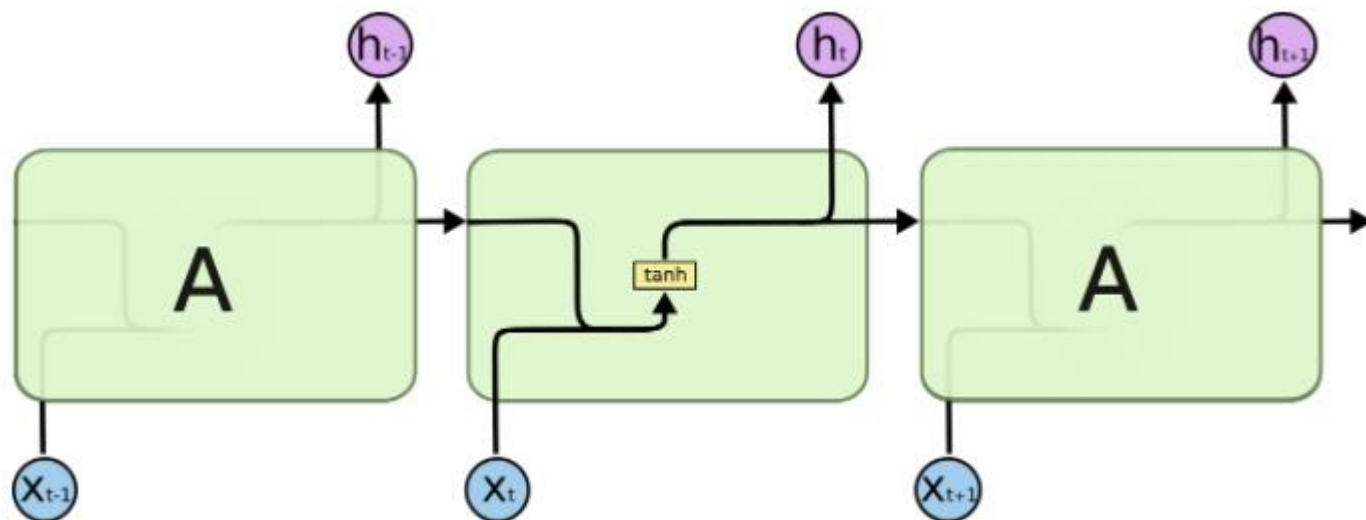
(5) 모델 3



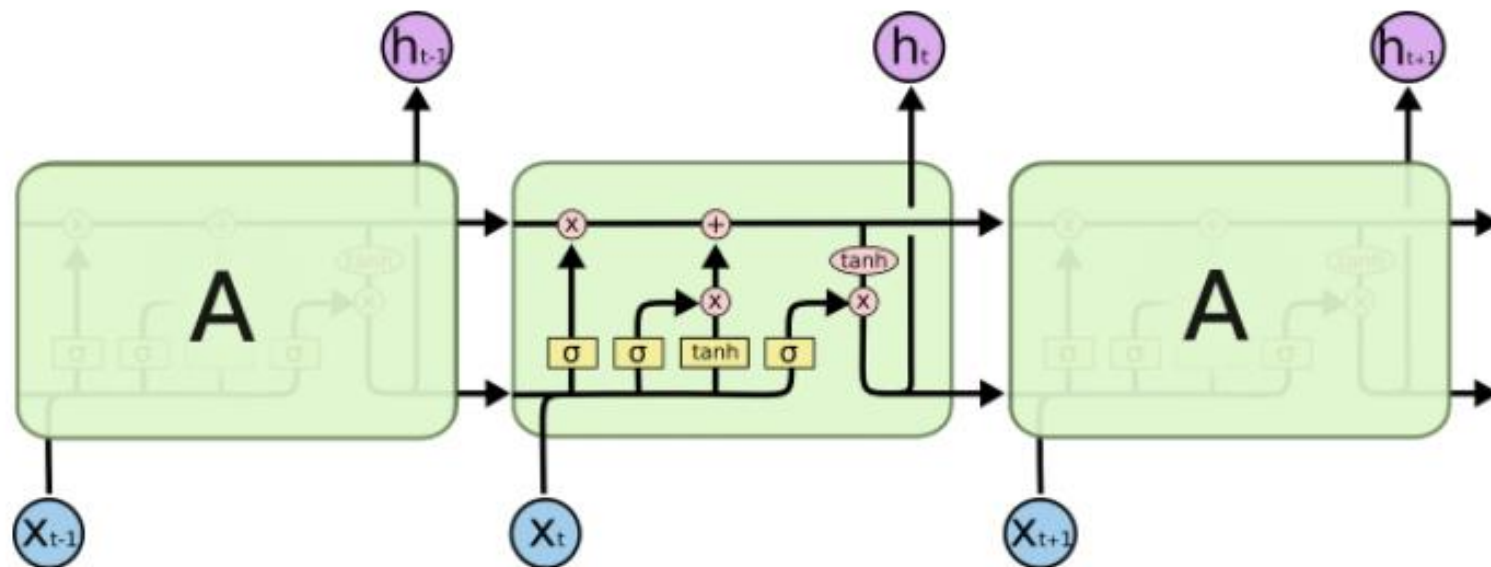
Each rectangle is a vector and arrows represent functions (e.g. matrix multiply). Input vectors are in red, output vectors are in blue and green vectors hold the RNN's state (more on this soon). From left to right: **(1)** Vanilla mode of processing without RNN, from fixed-sized input to fixed-sized output (e.g. image classification). **(2)** Sequence output (e.g. image captioning takes an image and outputs a sentence of words). **(3)** Sequence input (e.g. sentiment analysis where a given sentence is classified as expressing positive or negative sentiment). **(4)** Sequence input and sequence output (e.g. Machine Translation: an RNN reads a sentence in English and then outputs a sentence in French). **(5)** Synced sequence input and output (e.g. video classification where we wish to label each frame of the video). Notice that in every case are no pre-specified constraints on the lengths sequences because the recurrent transformation (green) is fixed and can be applied as many times as we like.

■ RNN


RNN



LSTM




LSTM의 반복 모듈에는 4개의 상호작용하는 layer가 들어있다.



Neural Network
Layer




Pointwise
Operation



Vector
Transfer

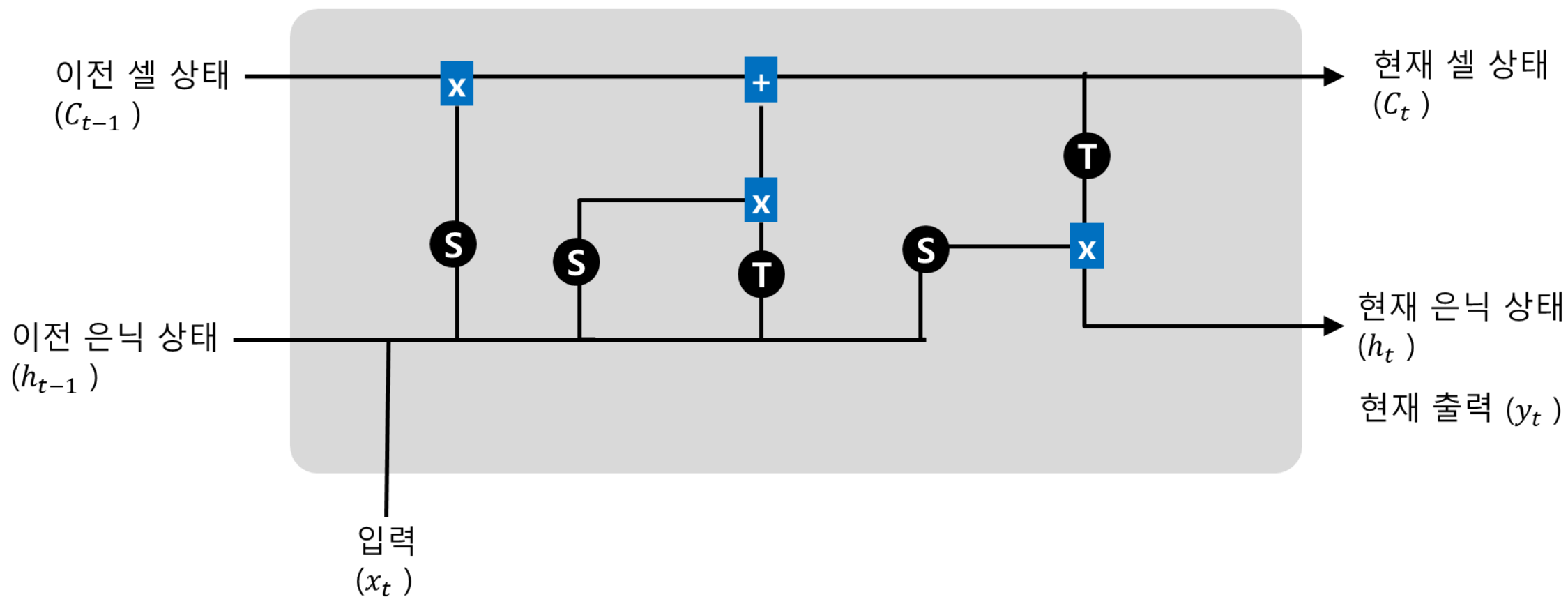


Concatenate



Copy

■ 내부 구조



- 셀 상태: 특정 시점에서 기억해야 할 중요한 정보를 다음 시점으로 전달
- 은닉 상태: 전반에 걸쳐 기억하는 전체 메모리