11. RNN

2019년 4월 15일 월요일 오전 9:18

1. ***Definition***

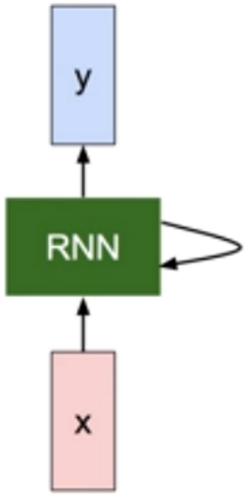
Recurrent Neural Network

We can process a sequence of vectors **x** by applying a recurrence formula at every time step:

$$h_t = f_W(h_{t-1}, x_t)$$

new state — old state — input vector at some time step

some function with parameters W



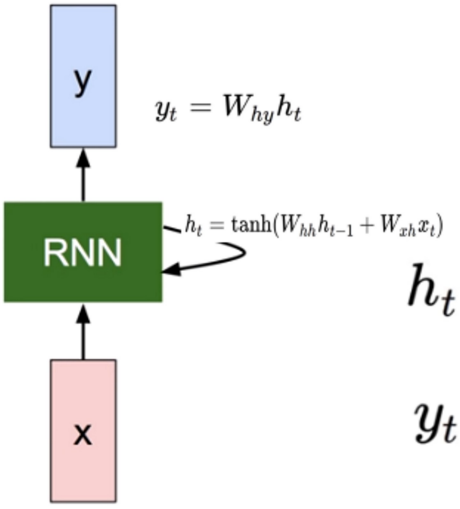
(Vanilla) Recurrent Neural Network

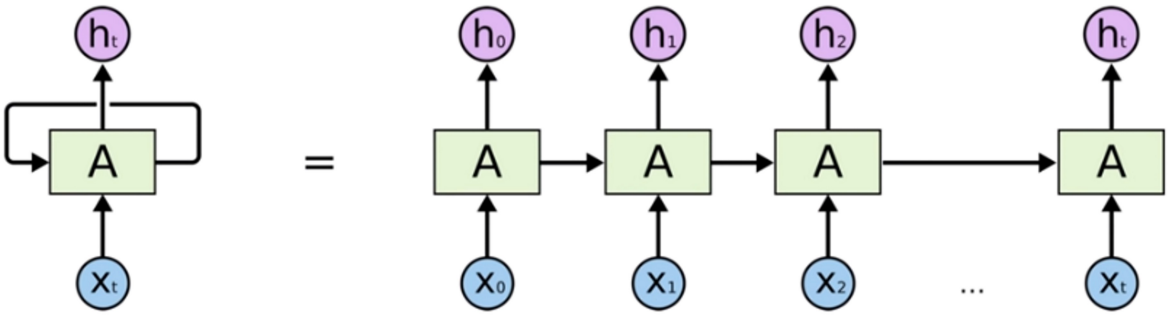
The state consists of a single *"hidden"* vector **h**:

$$h_t = f_W(h_{t-1}, x_t)$$

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$

$$y_t = W_{hy}h_t$$



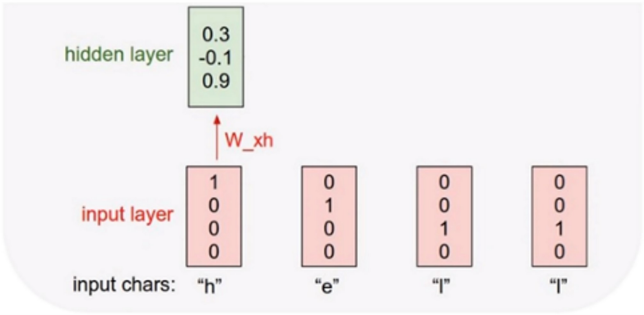


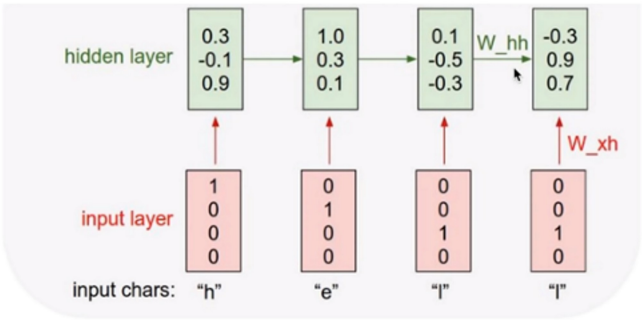
Notice: the same function and the same set of parameters are used at every time step.

텍스트 -> 벡터값 :

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$



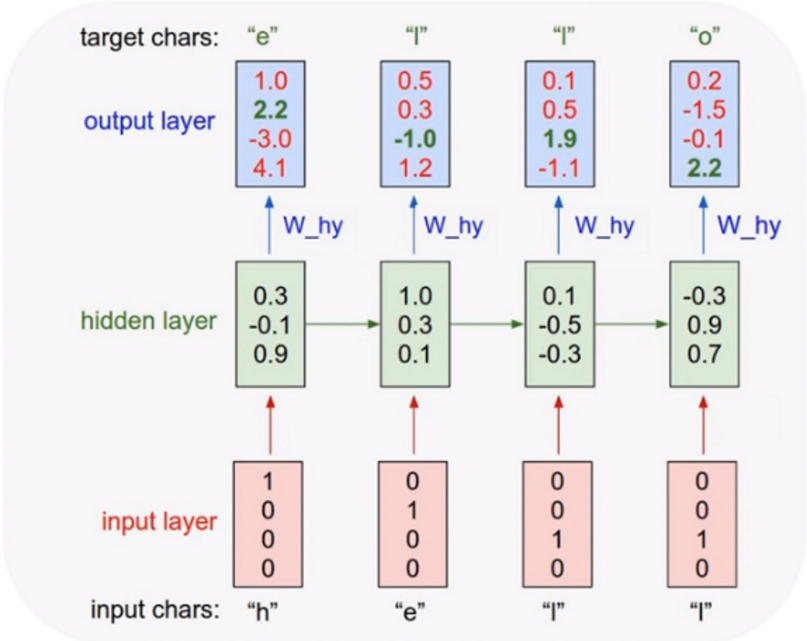


Character-level language model example

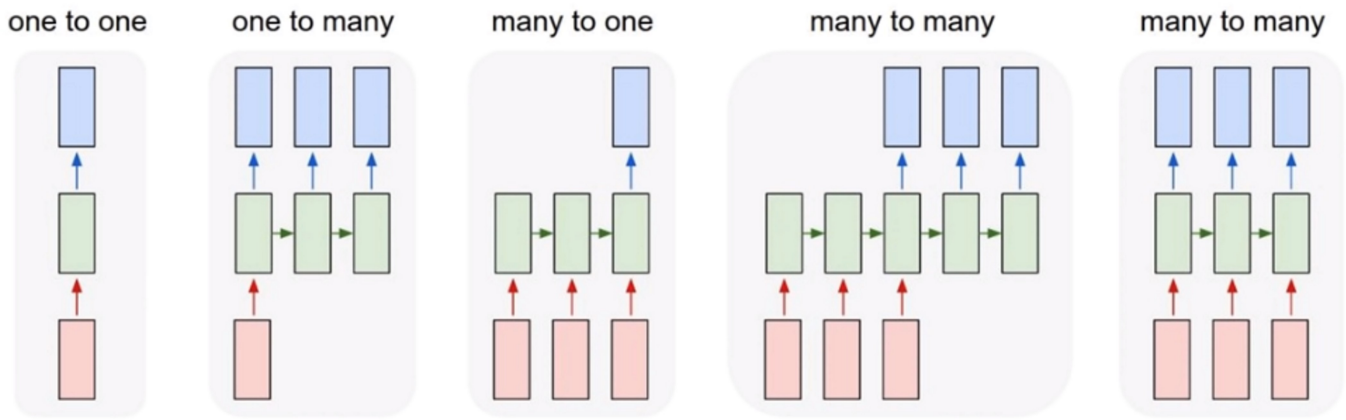
$$y_t = W_{hy}h_t$$

Vocabulary:
[h,e,l,o]

Example training sequence:
"hello"



Recurrent Networks offer a lot of flexibility:



e.g. **Image Captioning**
image -> sequence of words

e.g. **Machine Translation**
seq of words -> seq of words

e.g. **Video classification on frame level**

e.g. **Image Captioning**
image -> sequence of words

LSTM(long-short-term-memory)

GRU(gated recurrent unit)

https://aikorea.org/blog/rnn-tutorial-4/ ***-> LSMT, GRU 튜토리얼***