Sequence

is a collection of ordered items.

Sequences entail 2 kinds of information:
1. what are the individual items;
2. how those items are arranged.

Time flies like an arrow 

An arrow flies like time 

An arrow time like flies 

Modeling a sequence requires to account for the order of items to preserve the meaning.

.

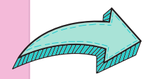
How to model a sequence ?

word embeddings
! are INDEPENDENT
Know nothing about words before or after

Design a **modeling strategy** to capture information about the dependency between items.

Modeling straregy

1. encode items sequentially
2. maintain a **memory** of items seen so far.

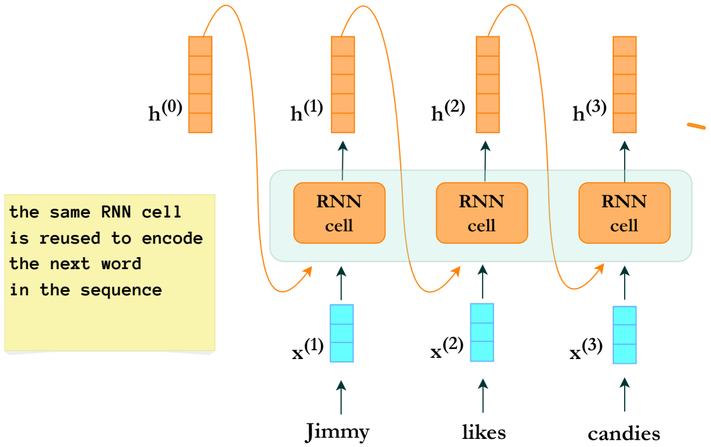


HIDDEN STATE

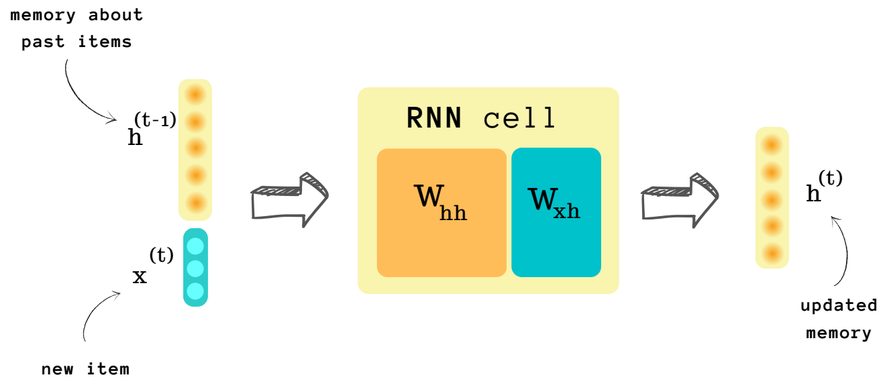
RNN

Goal of RNN:
to bake a sequence representation

Recurrent Neural Network

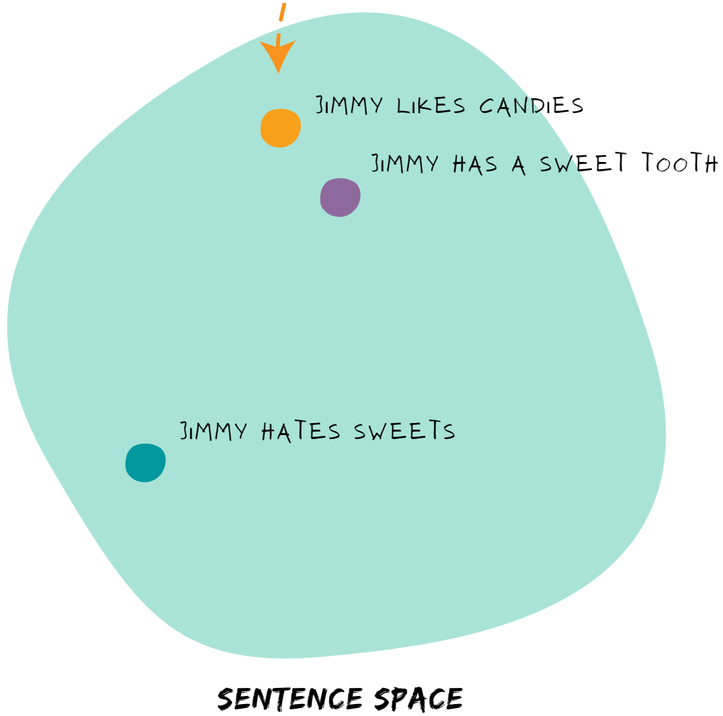


the same RNN cell is reused to encode the next word in the sequence



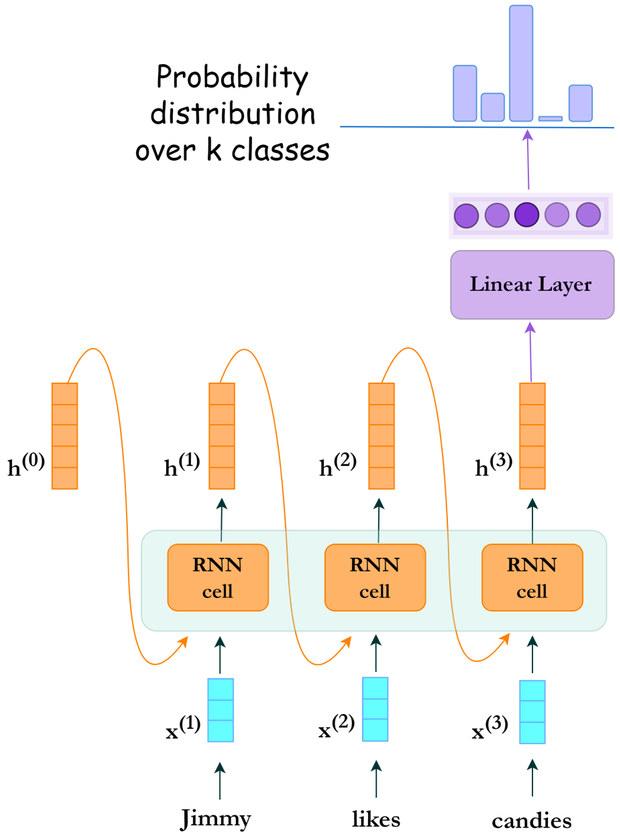
$$h^{(t)} = \sigma\left([W_{hh}; W_{xh}] @ \begin{bmatrix} h^{(t-1)} \\ x^{(t)} \end{bmatrix} + \text{bias} \right)$$

$$h^{(t)} = \sigma\left(W_{hh} h^{(t-1)} + W_{xh} x^{(t)} + \text{bias} \right)$$

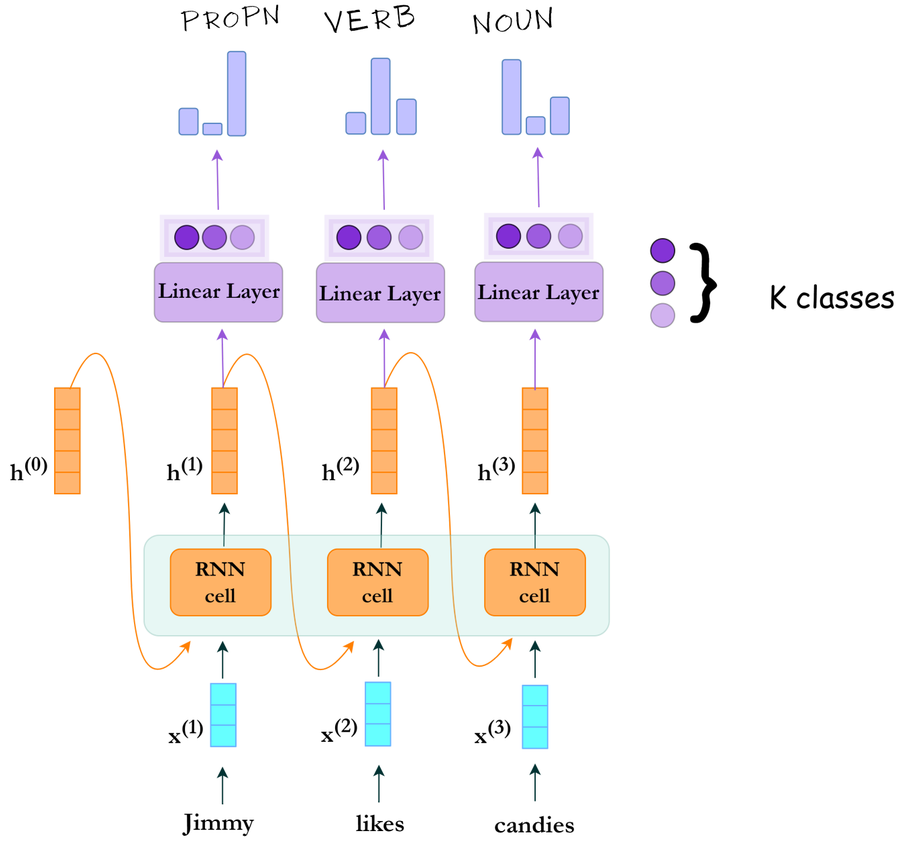


WHEN DO WE NEED RNNS?

Sequence Classification



Sequence Prediction



TEXT GENERATION with RNN

Output distribution

$$y^{(t)} = \text{Softmax}\left(U\, h^{(t)} + \text{bias}\right)$$

Hidden states

$$h^{(t)} = \sigma\left(W_{hh}\, h^{(t-1)} + W_{xh}\, x^{(t)} + \text{bias}\right)$$

Word embeddings

$$x^{(t)} = E\left(\text{word}^{(t)}\right)$$