Recurrent Neural Networks

Notation

Motivating example

NLP

This a problem called named-entity recognition and this is used by search engines for example to index all of, say the last 24 hours news of all the people mentioned in the news articles, so that they can index them appropriately.

x: Harry Potter and Hermione Granger invented a new spell.

$\rightarrow$ $x^{\langle 1 \rangle}$ $x^{\langle 2 \rangle}$ $x^{\langle 3 \rangle}$ — — . . $x^{\langle t \rangle}$ — — . . $x^{\langle 9 \rangle}$

$T_x = 9$

$\rightarrow$ y: 1   1   0   1   1   0   0   0   0

$y^{\langle 1 \rangle}$ $y^{\langle 2 \rangle}$ $y^{\langle 3 \rangle}$ — . . . $y^{\langle 9 \rangle}$

$T_y = 9$

$x^{(i)\langle t \rangle}$ $T_x^{(i)} = 9$ 15

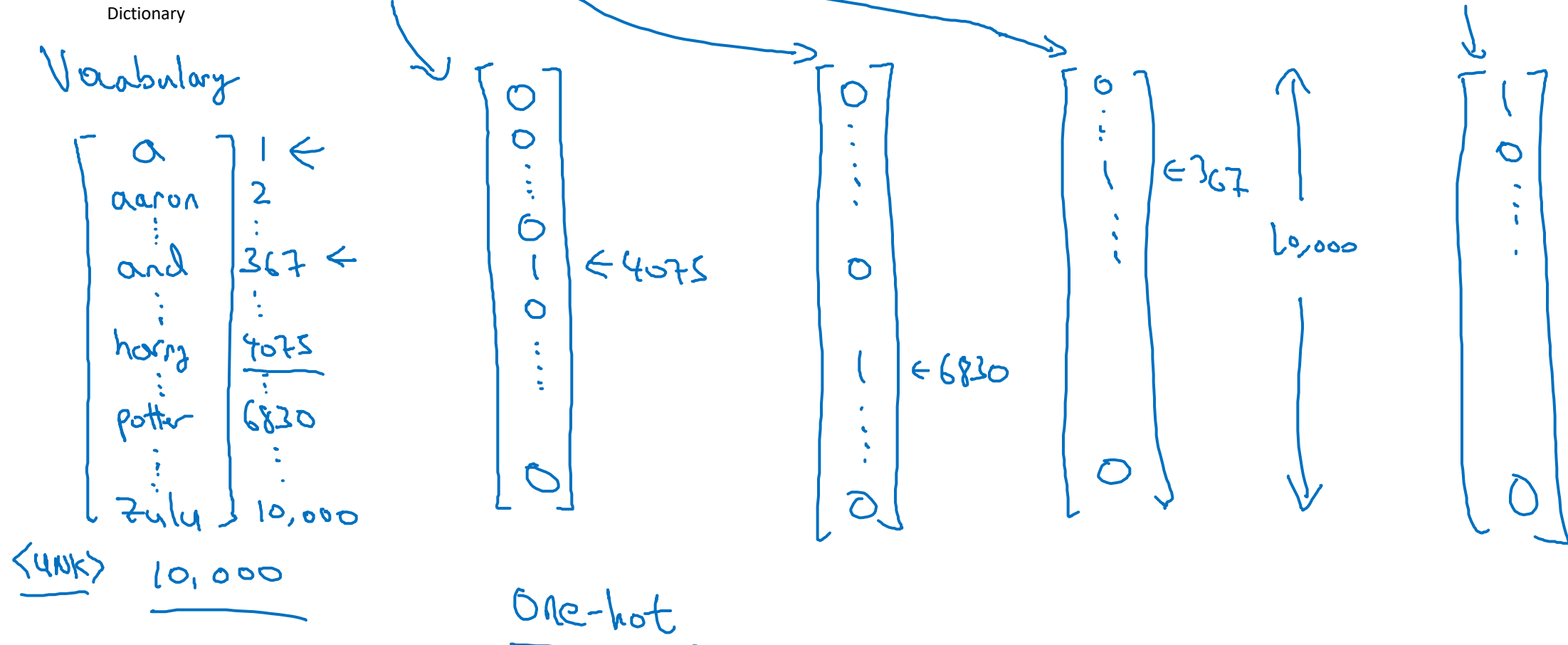$y^{(i)\langle t \rangle}$ $T_y^{(i)}$

Andrew Ng

Representing words

$x^{<t>}$ x_t for each of the values of t in a sentence will be a one-hot vector. (x, y)

$x \longrightarrow y$

x: Harry Potter and Hermione Granger invented a new spell.

$x^{<1>}$ $x^{<2>}$ $x^{<3>}$ to index the different position.   … $x^{<7>}$ $x^{<9>}$

Dictionary

Vocabulary

$$\begin{bmatrix} a \\ aaron \\ \vdots \\ and \\ \vdots \\ harry \\ \vdots \\ potter \\ \vdots \\ zulu \end{bmatrix} \begin{matrix} 1 \leftarrow \\ 2 \\ \vdots \\ 367 \leftarrow \\ \vdots \\ 4075 \\ \\ 6830 \\ \vdots \\ 10,000 \end{matrix}$$

$\langle UNK \rangle$ 10,000

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow 4075$$

One-hot

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow 6830$$

$$\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow 367$$

10,000

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Andrew Ng

Representing words

x: Harry Potter and Hermione Granger invented a new spell.

$x^{<1>}$ $x^{<2>}$ $x^{<3>}$ ... $x^{<9>}$

And = 367
Invented = 4700
A = 1
New = 5976
Spell = 8376
Harry = 4075
Potter = 6830
Hermione = 4200
Gran… = 4000