كلية الهندسة
قسم هندسة الروبوت و الأنظمة الذكية

الشبكات العصبية

محاضرة 4

NLP and Word Embeddings

د. فادي متوج

Word representation

Word representation

V = [a, aaron, ..., zulu, <UNK>] $|V| = 10{,}000$

1-hot representation

$$
\begin{array}{cccccc}
\text{Man} & \text{Woman} & \text{King} & \text{Queen} & \text{Apple} & \text{Orange} \\
(5391) & (9853) & (4914) & (7157) & (456) & (6257) \\
\begin{bmatrix}0\\0\\0\\0\\\vdots\\1\\\vdots\\0\\0\end{bmatrix} &
\begin{bmatrix}0\\0\\0\\0\\0\\\vdots\\1\\\vdots\\0\end{bmatrix} &
\begin{bmatrix}0\\0\\0\\\vdots\\1\\\vdots\\0\\0\\0\end{bmatrix} &
\begin{bmatrix}0\\0\\0\\0\\0\\\vdots\\1\\\vdots\\0\end{bmatrix} &
\begin{bmatrix}0\\\vdots\\1\\\vdots\\0\\0\\0\\0\\0\end{bmatrix} &
\begin{bmatrix}0\\0\\0\\0\\0\\\vdots\\1\\\vdots\\0\end{bmatrix} \\
O_{5391} & O_{9853} & \uparrow & \uparrow & \uparrow & \uparrow
\end{array}
$$

I want a glass of $\boxed{\text{orange juice}}$

I want a glass of $\boxed{\text{apple \_\_\_\_\_ ?.}}$

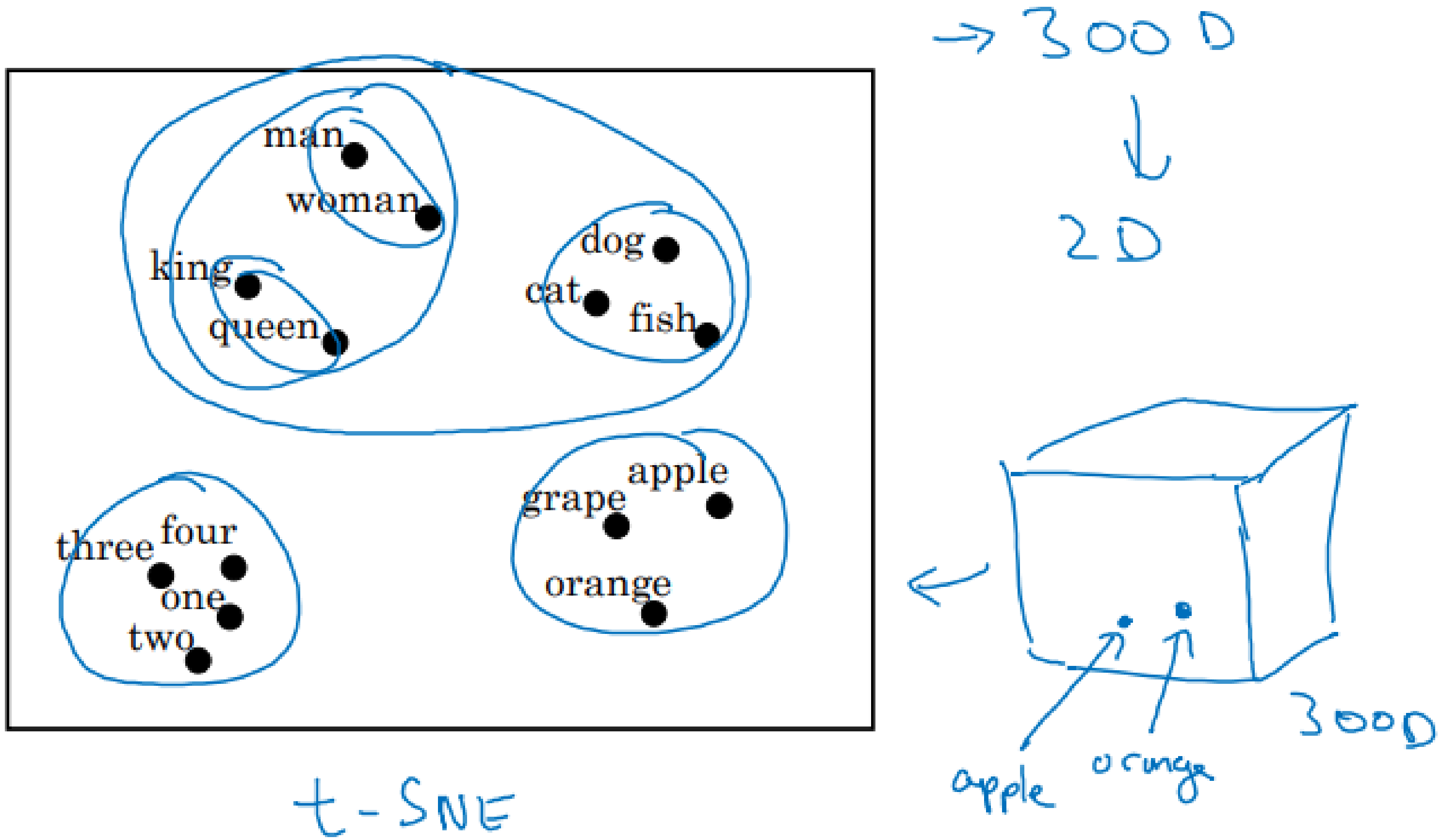
Featurized representation: word embedding

300	Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)
Gender	-1	1	-0.95	0.97	0.00	0.01
Royal	0.01	0.02	0.93	0.95	-0.01	0.00
Age	0.03	0.02	0.7	0.69	0.03	-0.02
Food	0.04	0.01	0.02	0.01	0.95	0.97
⋮	⋮	⋮				
size						
cost						
alive						
verb						
	e_{5391}	e_{9853}				

I want a glass of orange juice

I want a glass of apple juice

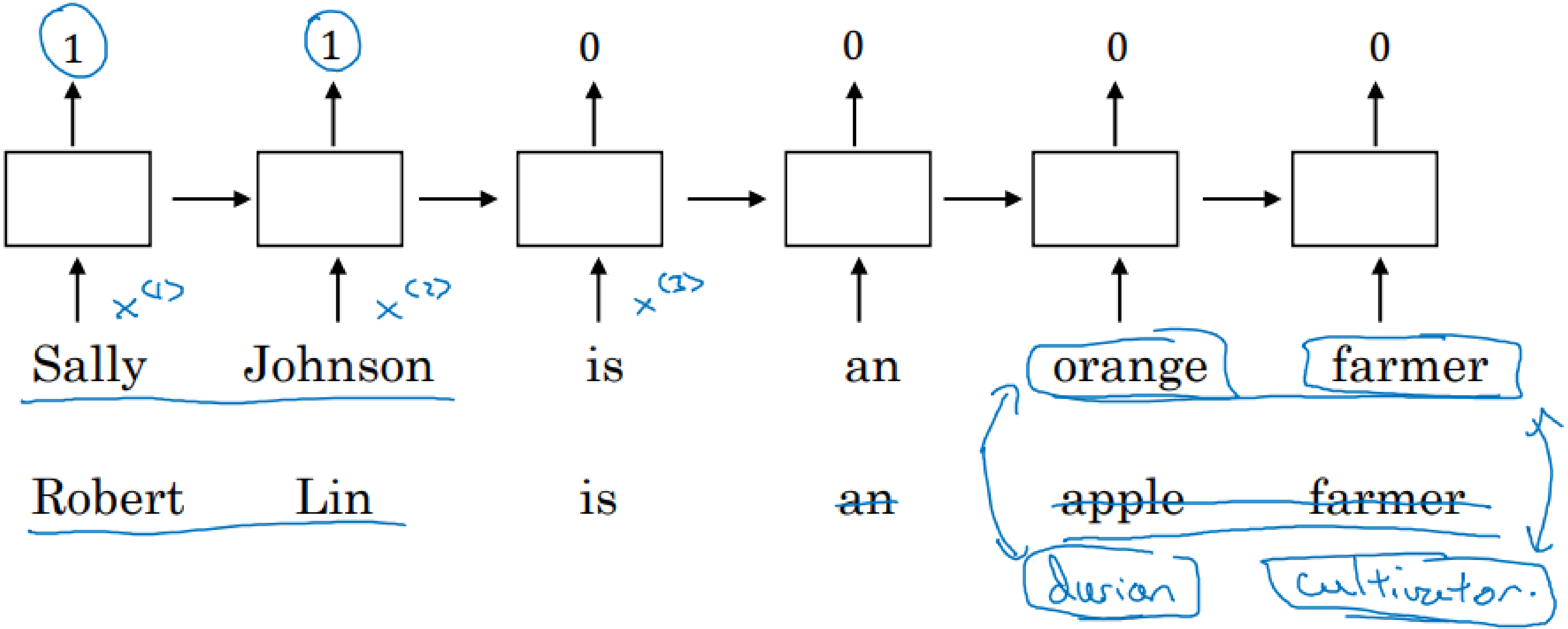
Visualizing word embeddings



[van der Maaten and Hinton., 2008. Visualizing data using t-SNE]

Using word embeddings

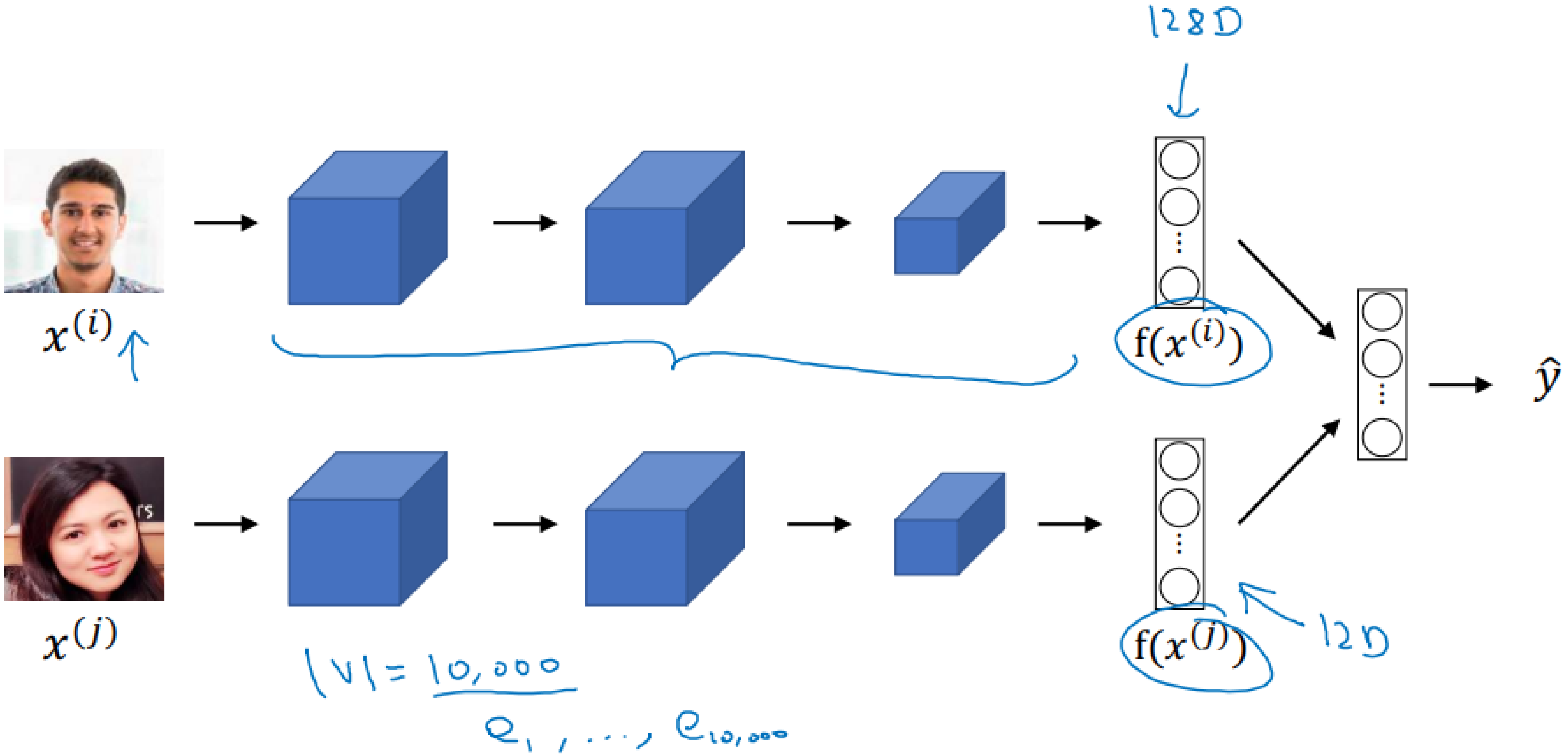
Named entity recognition example



Transfer learning and word embeddings

1. Learn word embeddings from large text corpus. (1-100B words) (Or download pre-trained embedding online.)
2. Transfer embedding to new task with smaller training set. (say, 100k words)
3. Optional: Continue to finetune the word embeddings with new data.

Relation to face encoding



[Taigman et. al., 2014. DeepFace: Closing the gap to human level performance]

Properties of word embeddings

Analogies

جَامعة
المَـنارة

	Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)
Gender	−1	1	-0.95	0.97	0.00	0.01
Royal	0.01	0.02	0.93	0.95	-0.01	0.00
Age	0.03	0.02	0.70	0.69	0.03	-0.02
Food	0.09	0.01	0.02	0.01	0.95	0.97

e_{5391}

e_{man}

e_{woman}

Man → Woman as King →? Queen

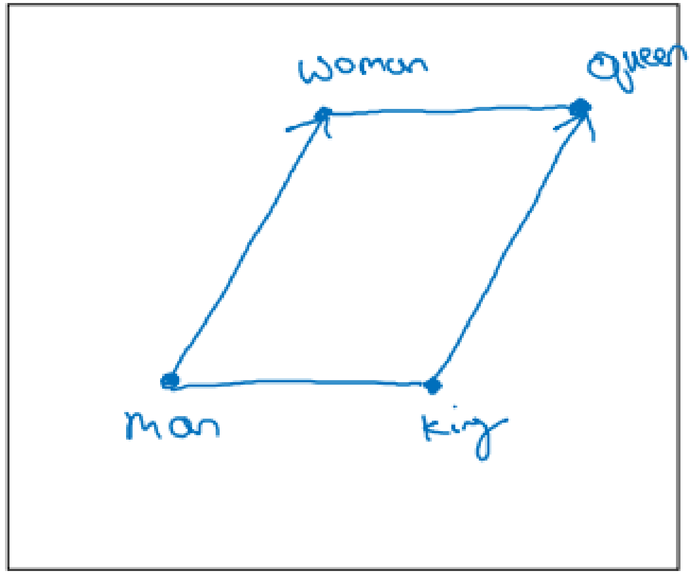
$e_{man} - e_{woman} \approx e_{king} - e_?$ ← Queen

$e_{man} - e_{woman} \approx \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$e_{king} - e_{queen} \approx \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

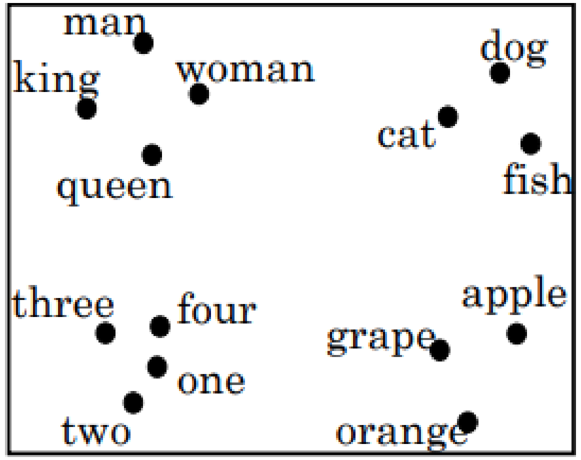
[Mikolov et. al., 2013, Linguistic regularities in continuous space word representations]

Analogies using word vectors



300 D

300D → 2D



t-SNE

$$e_{man} - e_{woman} \approx e_{king} - \cancel{e_?}\ e_w$$

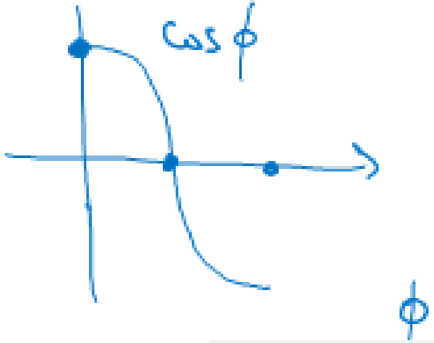
Find word w: $\arg\max_w \boxed{sim(e_w, e_{king} - e_{man} + e_{woman})}$

30 - 75%

Cosine similarity

$$\boxed{sim(e_w, e_{king} - e_{man} + e_{woman})}$$

$$\underline{sim(u,v)} = \frac{u^T v}{\|u\|_2 \|v\|_2}$$



$$\|u - v\|^2$$

Man:Woman as Boy:Girl

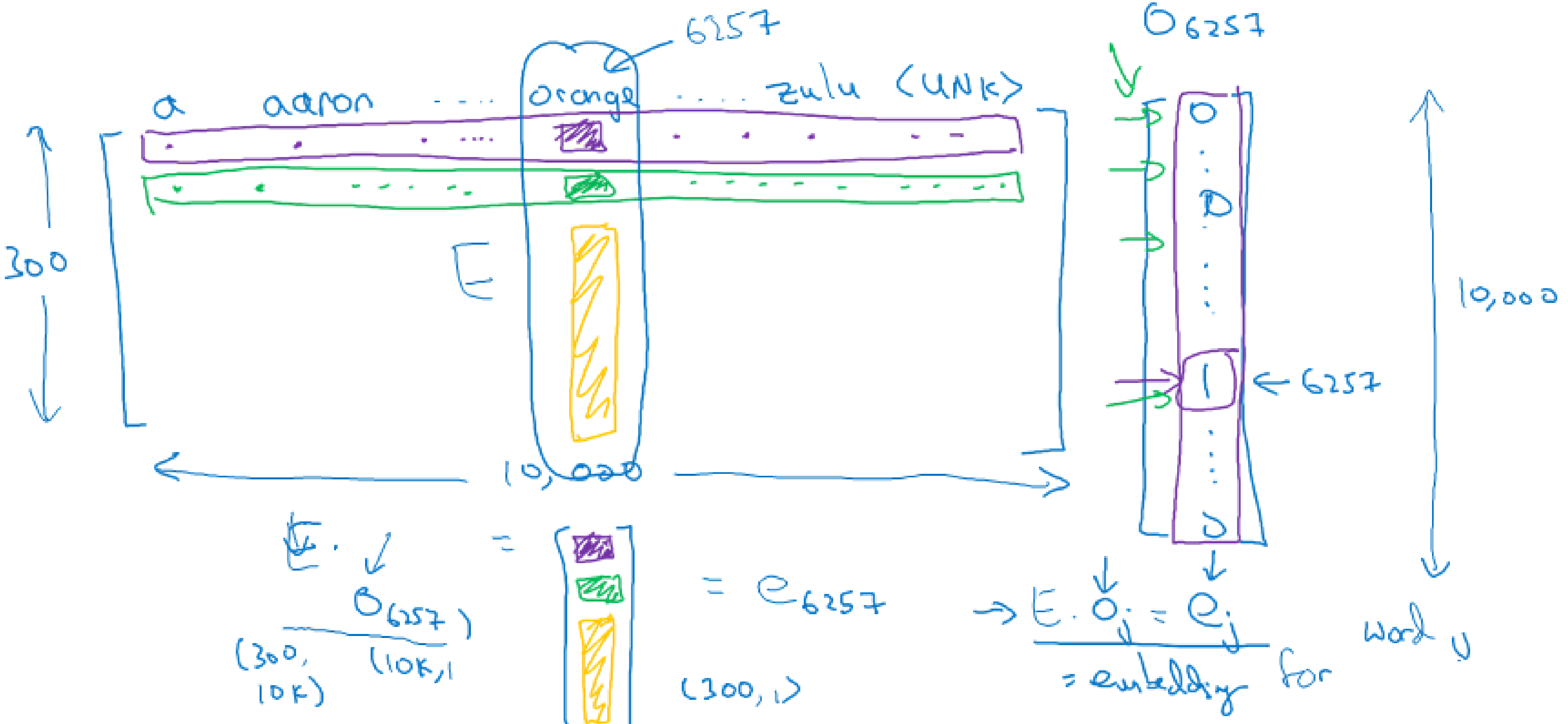
Ottawa:Canada as Nairobi:Kenya

Big:Bigger as Tall:Taller

Yen:Japan as Ruble:Russia

Embedding matrix

Embedding matrix

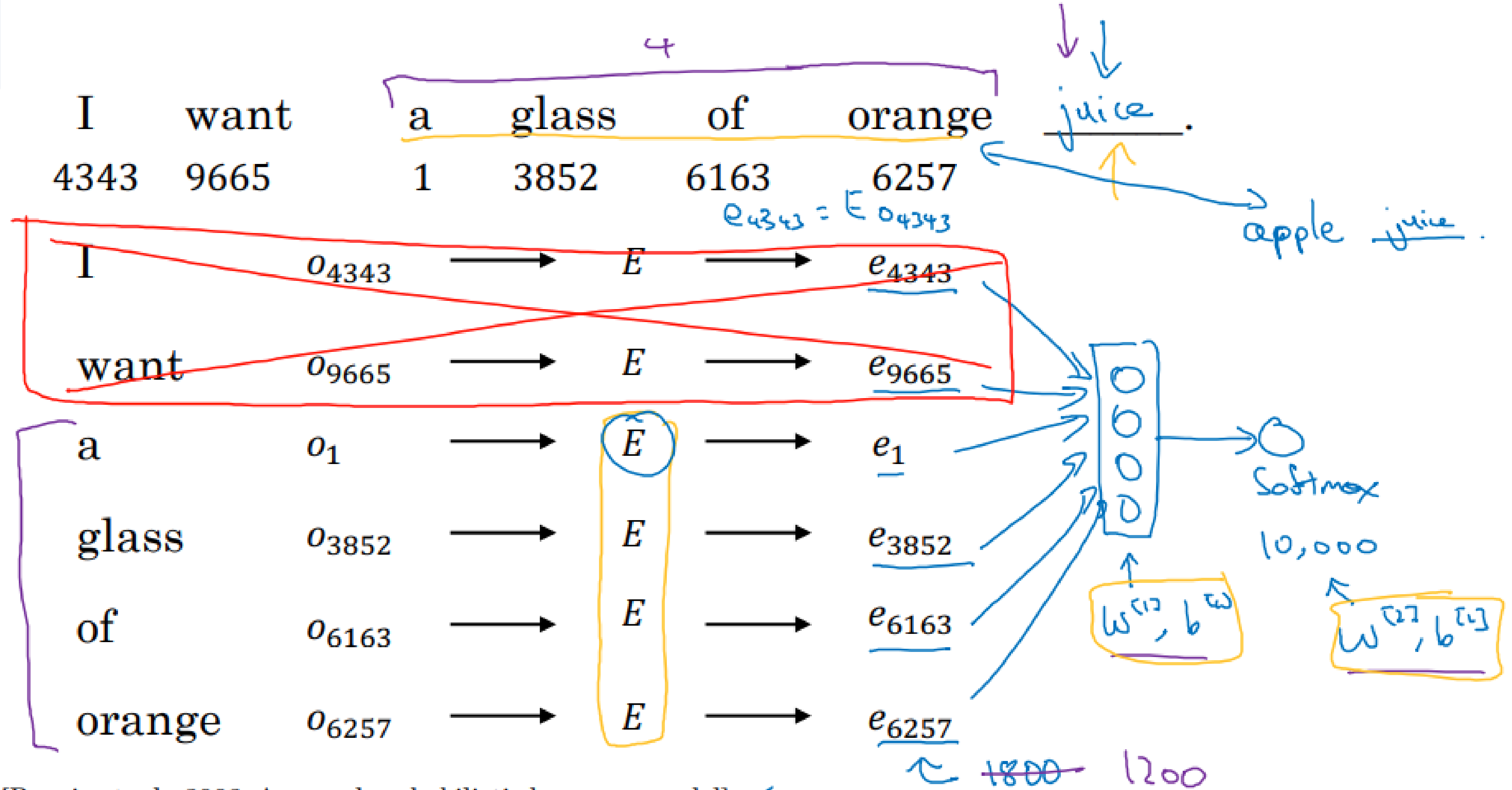


In practice, use specialized function to look up an embedding.

→ Embedding

Learning word embeddings

Neural language model



[Bengio et. al., 2003, A neural probabilistic language model]

Other context/target pairs

جَـامعة
المَـنارة
MANARA UNIVERSITY

I want a $\boxed{\text{glass}}$ of $\boxed{\text{orange}}$ $\underset{\text{target}}{\text{juice}}$ to go along with my cereal.

context

Context: Last 4 words.

- 4 words on left & right — a glass of orange __?__ to go along with
- Last 1 word — orange __?__
- Nearby 1 word — glass __?__

skip gram