Word representation

V = [a, aaron, ..., zulu, <UNK>]

$|V| = 10,000$

1-hot representation

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

O_{5391} O_{9853}
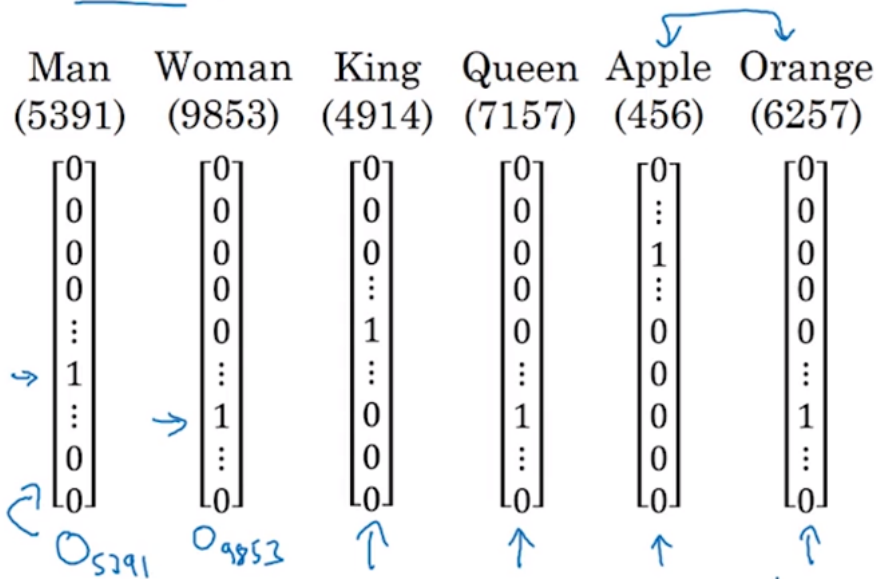
I want a glass of orange juice.

I want a glass of apple____?____.

Andrew Ng

One of the weaknesses of this representation is that it treatd each word as a thing onto itself and it doesn't allow an algorithm to easily generalize the cross words.

And you could get distance between any pair of these vectors is also the same. So it just doesn't know that somehow apple and orange are much more similar than king and orange or queen or orange.
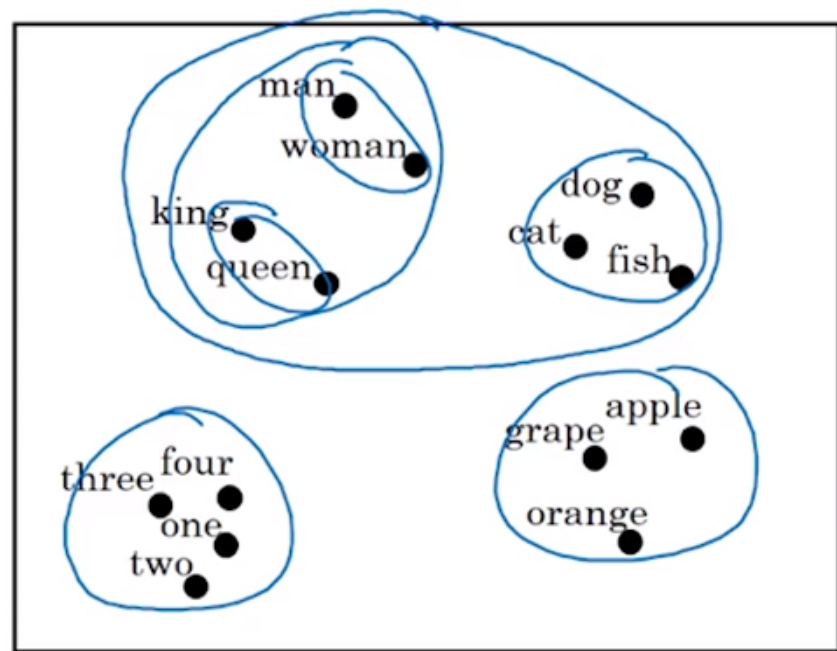
Featurized representation: word embedding

	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	0.09	0.01	0.02	0.01	0.95	0.97

300

size
cost
alive   verb

e_{5391} e_{9853}

I want a glass of orange juice.

I want a glass of apple juice.

Andrew Ng

Visualizing word embeddings

$\rightarrow 300D$

$\downarrow$

$2D$

man

woman

king

queen

dog

cat fish

four

three

one

two

grape apple

orange

$t - SNE$

$300D$

apple orange

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

Andrew Ng