Vector semantics and embeddings

@manuel__pita
Dr. Manuel Pita

UNIVERSIDADE
LUSÓFONA

What do words mean?

- N-gram or text classification methods we've seen so far
  - Words are just strings (or indices w_i in a vocabulary list)
  - That's not very satisfactory!
- Introductory logic classes:
  - The meaning of "dog" is DOG; cat is CAT
  - $\forall x \, . \, DOG(x) \implies MAMMAL(x)$
- Old linguistics joke by Barbara Partee in 1967:
  - Q: What's the meaning of life?
  - A: LIFE
- That seems hardly better!

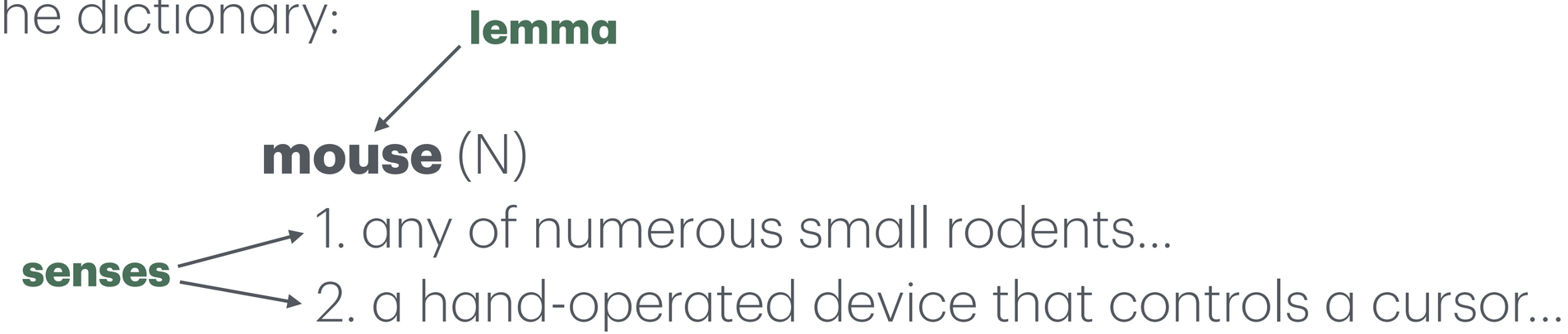
Desiderata

What should a theory of word meaning do for us? Let's look at some desiderata

From **lexical semantics**, the linguistic study of word meaning

Desiderata

From the dictionary:



A sense or "**concept**" is the meaning component of a word

Lemmas can be **polysemous** (have multiple senses)

Relations between senses

synonymy

- Synonyms have the same meaning in some or all contexts.
  - filbert / hazelnut
  - couch / sofa
  - big / large
  - automobile / car
  - vomit / throw up
  - water / H20

Relations between senses

synonymy

Notice that there are probably no examples of perfect synonymy, even if many aspects of meaning are identical, word use may differ based on politeness, slang, register, genre, etc.

> **water/H20**. Would it be strange to use ”H20" in a surfing guide?

> **big/large**: "my big sister” is not the same as “my large sister"

Linguistic principle of contrast:

difference in form -> difference in meaning

So what? No perfect synonyms.

Relations between senses

Similarity

- Words with similar meanings. Not synonyms, but sharing some element of meaning
  - car, bicycle
  - cow, horse

If we ask humans about similarity...

word1	word2	similarity
vanish	disappear	9.8
behave	obey	7.3
belief	impression	5.95
muscle	bone	3.65
modest	flexible	0.98
hole	agreement	0.3

Relations between senses

Relatedness

- Also called "word association"
- Words can be related in any way, perhaps via a semantic frame or field
  - coffee, tea: **similar**
  - coffee, cup: **related**, not similar

Semantic field

Words that cover a particular semantic domain and bear structured relations with each other.

hospitals

surgeon, scalpel, nurse, anaesthetic, hospital

restaurants

waiter, menu, plate, food, menu, chef

houses

door, roof, kitchen, family, bed

Relations between senses

Antonymy

- Senses that are opposites with respect to only one feature of meaning. Otherwise, they are very similar!
- dark/light short/long fast/slow rise/fall
- hot/cold up/down in/out
- More formally: antonyms can
  - define a binary opposition or be at opposite ends of a scale
  - long/short, fast/slow
  - Be reversives:
  - rise/fall, up/down

Connotation

Sentiment

- **Words have affective meanings**
  - Positive connotations (happy)
  - Negative connotations (sad)
- **Connotations can be subtle:**
  - Positive connotation: copy, replica, reproduction
  - Negative connotation: fake, knockoff, forgery
- **Evaluation (sentiment!)**
  - Positive evaluation (great, love)
  - Negative evaluation (terrible, hate)

Connotation

Sentiment

- Words seem to vary along 3 affective dimensions:
  - **valence**: the pleasantness of the stimulus
  - **arousal**: the intensity of emotion provoked by the stimulus
  - **dominance**: the degree of control exerted by the stimulus

	Word	Score		Word	Score
Valence	love	1.000		toxic	0.008
	happy	1.000		nightmare	0.005
Arousal	elated	0.960		mellow	0.069
	frenzy	0.965		napping	0.046
Dominance	powerful	0.991		weak	0.045
	leadershi	0.983		empty	0.081

So far...

- Concepts or word senses
  - Have a complex many-to-many association with words (homonymy, multiple senses)
- Have relations with each other
  - Synonymy
  - Antonymy
  - Similarity
  - Relatedness
  - Connotation

Computational models of word meaning

Can we build a theory of how to represent word meaning, that accounts for at least some of the desiderata?

We'll introduce vector semantics

- The standard model in language processing
- Handles many of our goals

“The meaning of a word is its use in the language”

Ludwig Wittgenstein

Defining words according to their use

One way to define "usage":

words are defined by their **environments** (the words around them)

Zellig Harris (1954):

If A and B have almost identical environments we say that they are synonyms.

Defining words according to their use

Suppose you see these sentences:

- Ong choi is delicious sautéed with garlic.
- Ong choi is superb over rice
- Ong choi leaves with salty sauces

And you've also seen these:

- ...spinach sautéed with garlic over rice
- Chard stems and leaves are delicious
- Collard greens and other salty leafy greens

Conclusion:

Ongchoi is a leafy green like spinach, chard, or collard greens

We could conclude this based on words like "leaves" and "delicious" and "sauteed"

In fact Ongshoi is a type of water spinach

Idea 1

Defining meaning by linguistic distribution

Let's define the meaning of a word by its distribution in language use, meaning its neighbouring words or grammatical environments.

Idea 2

Meaning as a point in space

3 affective dimensions for a word

valence: pleasantness

arousal: intensity of emotion

dominance: the degree of control exerted

	Word	Score		Word	Score
Valence	love	1.000		toxic	0.008
	happy	1.000		nightmare	0.005
Arousal	elated	0.960		mellow	0.069
	frenzy	0.965		napping	0.046
Dominance	powerful	0.991		weak	0.045
	leadershi	0.983		empty	0.081

Hence the connotation of a word is a vector in 3-space

Idea 1: Defining meaning by linguistic distribution

Idea 2: Meaning as a point in multidimensional space

Defining meaning as a point in space based on distribution

Each word = a vector (not just "good" or w_{45})

Similar words are "nearby in semantic space"

We build this space automatically by seeing which words are nearby in text

Word meaning is a vector

- Called an "embedding" because it's embedded into a vectorial space
- The standard way to represent meaning in NLP
- Every modern NLP algorithm uses embeddings as the representation of word meaning
- Fine-grained model of meaning for similarity

Word2Vec

- Inspired by the BoW vector representation of documents and words
- Motivated by having words in a vector space that allowed more semantic inference
- Using machine learning to train word vectors by neighbouring words

Mikolov, T., *et al*. 2013). In *ADV NEU INF PROC SYS* (pp. 3111-3119).

Word2Vec

How are word embeddings (vectors) trained?

Training examples:

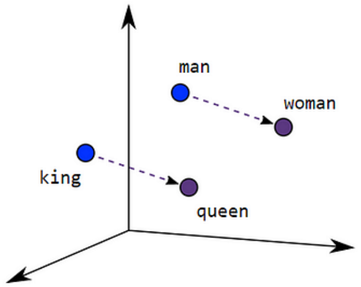
1. She went to the bank of the river : `(bank, river, go)`

2. He deposited some money in the bank: `(bank, deposit, money)`

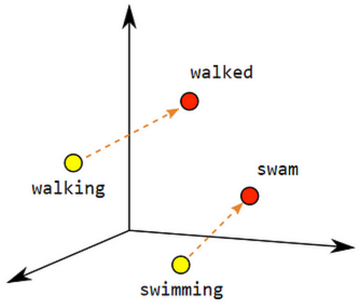
In these triplets, bank is the target word, the other words are context words

The machine learning algorithm adjusts the vector for bank to get it more aligned to new context words

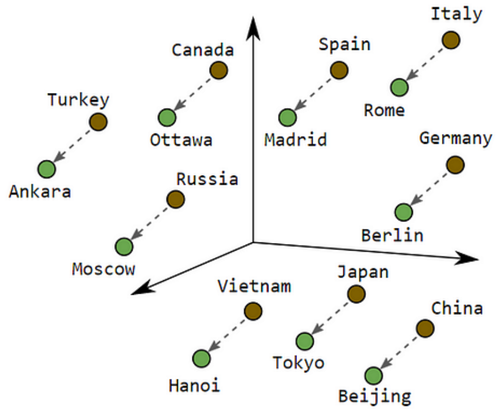
That means that analogous relationships are well captured in the vector field



Male-Female

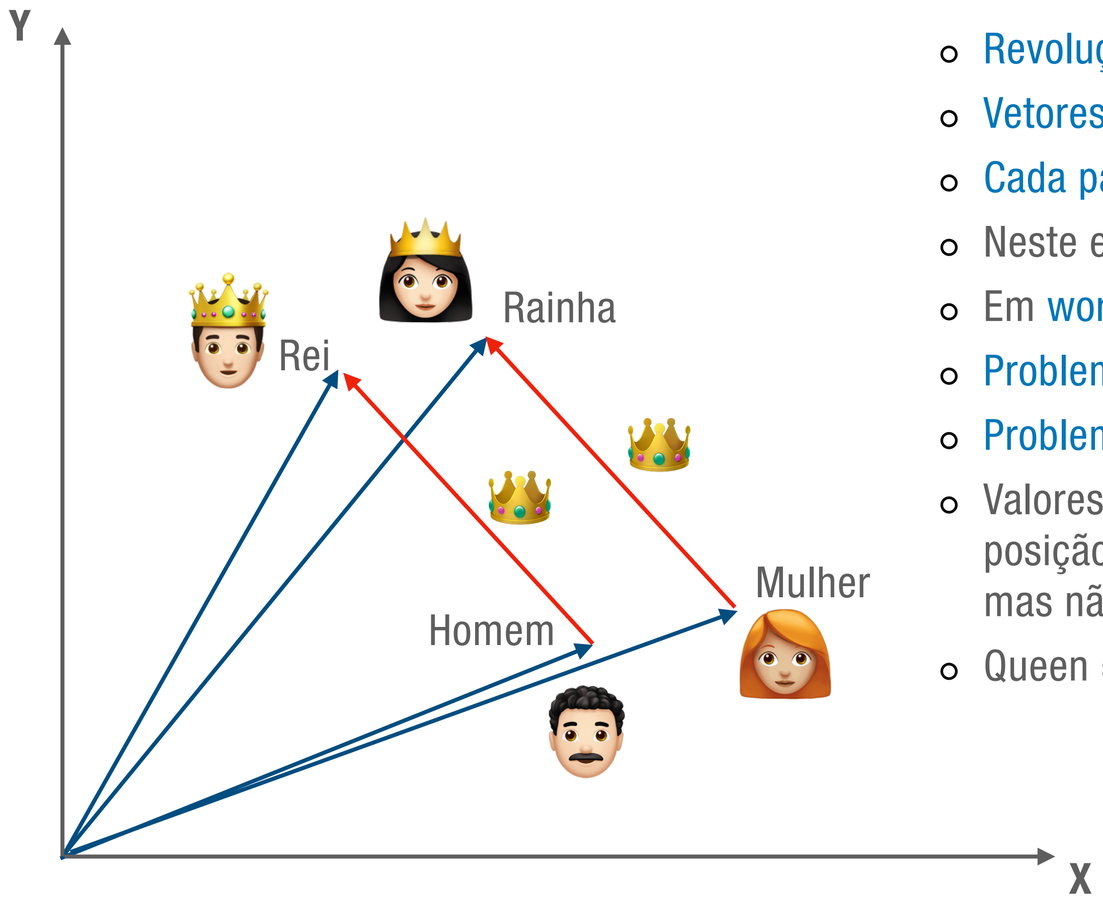


Verb Tense



Country-Capital

Word2Vec



- Revolução no processamento de linguagem natural em 2013
- Vetores com muitas dimensões treinados com grandes volumes de dados
- Cada palavra se posiciona em relação às outras
- Neste exemplo os vetores ilustrativos existem em duas dimensões X e Y
- Em word2vec os vetores têm 300 dimensões para ter suficiente flexibilidade
- Problema 1: palavras em contexto
- Problema 2: representações de frases, parágrafos, etc.
- Valores das componentes dos vetores (embeddings) são usados para ajustar posição do vetor de cada palavra em relação aos vetores de outras palavras, mas não têm significado específico
- Queen = King - Man + Woman

Mikolov, T., *et al*. 2013). In *ADV NEU INF PROC SYS* (pp. 3111-3119).

Word2Vec

- Each word has only one vector representation
- If there are many different contexts, word vector is in the middle between them
- This means that contextual semantics is hard in the Word2Vec model

Mikolov, T., *et al*. 2013). In *ADV NEU INF PROC SYS* (pp. 3111-3119).