CHAPTER6: Vector Semantics and Embeddings

Instructor: PhD. Nguyen Thi Quy

Group 5: 6.1 - 6.2

December 14, 2023

Presentation Overview

Presentation Overview

1. **Lexical Semantics**
2. Vector Semantics

What do words mean?

N-gram or text classification methods we've seen so far

- Words are just strings (or indices wi in a vocabulary list)
- That's not very satisfactory!

Introductory logic classes:

The meaning of "dog" is **DOG**; cat is **CAT**

$$\forall x, DOG(x) \longrightarrow MAMMAL(x)$$

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 lemma

mouse (N)

sense

1. any of numerous small rodents...
2. a hand-operated device that controls a cursor...

Modified from the online thesaurus WordNet

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 / H_2O

Relations between senses: Synonymy

Note that there are probably no examples of perfect synonymy.

- Even if many aspects of meaning are identical
- Still may differ based on politeness, slang, register, genre, etc.

water / $\mathrm{H_2O}$

"$\mathrm{H_2O}$" in a surfing guide?

big / large

my big sister $\neq$ my large sister

The Linguistic Principle of Contrast

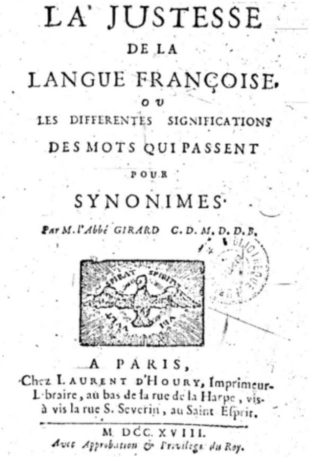
Difference in form $\rightarrow$ difference in meaning

Abbé Gabriel Girard 1718t

Re: "exact" synonyms

"je ne crois pas qu'il y ait de mot synonime dans aucune Langue"

[I do not believe that there is a synonymous word in any language]


LA JUSTESSE
DE LA
LANGUE FRANÇOISE.
OU
LES DIFFERENTES SIGNIFICATIONS
DES MOTS QUI PASSENT
POUR
SYNONIMES.
Par M. l'Abbé GIRARD C. D. M. D. D. B.

A PARIS,
Chez LAURENT D'HOURY, Imprimeur-Libraire, au bas de la rue de la Harpe, vis-à vis la rue S. Severin, au Saint Esprit.
M. DCC. XVIII.
Avec Approbation & Privilege du Roy.

Relation: **Similarity**

Words with similar meanings. Not synonyms, but sharing some element of meaning

- car, bicycle
- cow, horse

Relation: **Similarity**

Ask humans how similar 2 words are

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

SimLex-999 dataset (Hill et al., 2015)

Relation: Word 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
- 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

Relation: 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

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
	leadership	0.983		empty	0.081

Values from NRC VAD Lexicon (Mohammad 2018)

Word Meaning

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

Presentation Overview

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!

Ludwig Wittgenstein

PI #43:

"The meaning of a word is its use in the language"

Let's define words by their usages

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.

What does recent English borrowing *ongchoi* mean?

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"

Ongchoi: *Ipomoea aquatica "Water Spinach"*

空心菜
kangkong
rau muống
...

Idea 1: Defining meaning by linguistic distribution

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

Idea 2: Meaning as a point in space (Osgood et al. 1957)

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
	leadership	0.983		empty	0.081

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

Vector Semantics

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**

We define meaning of a word as a vector

Called an "embedding" because it's embedded into a space (see textbook)

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

Intuition: why vectors?

Consider sentiment analysis:

- With **words**, a feature is a word identity
- Feature 5: 'The previous word was "terrible"'
- requires **exact same word** to be in training and test

With **embeddings**:

- Feature is a word vector
- The previous word was vector [35,22,17...]
- Now in the test set we might see a similar vector [34,21,14]
- We can generalize to **similar but unseen** words!!!

Intuition: why vectors?

We'll discuss 2 kinds of embeddings

Tf-idf

- Information Retrieval workhorse!
- A common baseline model
- **Sparse** vectors
- Words are represented by (a simple function of) the **counts** of nearby words

Word2vec

- **Dense** vectors
- Representation is created by training a classifier to **predict** whether a word is likely to appear nearby
- Later we'll discuss extensions called contextual embeddings

Computing with meaning representations

From now on: Computing with meaning representations instead of string representations

荃者所以在鱼，得鱼而忘荃	Nets are for fish; Once you get the fish, you can forget the net.
言者所以在意，得意而忘言	Words are for meaning; Once you get the meaning, you can forget the words
	庄子(Zhuangzi), Chapter 26

Reference

Speech and Language Processing (3rd ed. draft)

Dan Jurafsky and James H. Martin

Part I: Fundamental Algorithms, *Chapter 6: Vector Semantics and Embeddings*

Thanks for listening!

Q&A section