



Sparse Usage Graphs as Model for Word Meaning in Context

November 27, 2020

Dominik Schlechtweg Supervisor: apl. Prof. Dr. Sabine Schulte im Walde Institute for Natural Language Processing, University of Stuttgart, Germany

Word Meaning

a word's meaning is the knowledge a word can trigger in a human

(p. 54 Blank, 1997)

Word Meaning

außerspi	achliches Wissen	3a. Konnotationen
	einzelsprachlich-	interne Wort- 2b. vorstellung
lex	lexikalisches Wissen	2a. externe Wort-
	I. einzelsprachlich- sememisches Wissen	1. Semem
		vorstellung
		syntagmatische 2c. Relationen
		3h Weltwissen

Figure 1: Blank (1997)'s levels of word meaning (p. 95).

Word Meaning

- has traditionally been equated to a set of senses ("polysemy", Bréal 1897)
- abstractions over patterns of use (Kilgarriff, 2007)
- each usage of a word in context disambiguates it / "determines" one of its senses (Navigli, 2009)

History in CL

- before second half of 20 century mainly historical (cf. Blank, 1997)
- WSD (Weaver, 1949/1955)
- Cognitive Semantics, Prototype Theory (Rosch & Mervis, 1975)
- \rightarrow discrete to graded (Erk, McCarthy, & Gaylord, 2009, 2013)
 - "I don't believe in the word senses" (Kilgarriff, 1997)

The Blank-Kilgarriff-Continuum

Word senses are a lexicographer's attempt to impose order on a word's role in the language. [...] [The process of a] speaker's understanding of a word [...] is not one which naturally gives rise to a set of distinct senses.

(Kilgarriff, 2007)

[Polysemy is the] consciously experienceable, intersubjectively comprehensible existence of a semantic relationship between two distinct meanings of a word.

(p. 424 Blank, 1997)

The lexicographic process

- 1. gather a corpus of citations for a word;
- divide the citations up into clusters, so that, as far as possible, all the members of each cluster have more in common with any other member of that cluster, than with any member of any other cluster;
- 3. for each cluster, (post-hoc) work out what it is that makes its members belong together; and
- 4. take these conclusions and code into a dictionary definition.

(Kilgarriff, 2007)

The lexicographic process

- lexicographer's criteria in step 2 are not explicit and subjective (Kilgarriff, 1997, 2007)
- we make criteria explicit and inter-subjective
- operationalize Blank's criteria
- gives rise to a graded scale of word meaning in context (Erk et al., 2013)

Advancing the graded view

- graded word sense annotation (Erk et al., 2013)
- Word Usage Graphs (WUGs)
- idea first proposed in McCarthy, Apidianaki, and Erk (2016)
- first operationalized in Schlechtweg, McGillivray, Hengchen, Dubossarsky, and Tahmasebi (2020)
- large-scale, multi-lingual, diachronic, resource of WUGs based on 100,000 judgments

Word Usage Graphs



Figure 2: Usage graph of Swedish *ledning*.

Annotation

- 100–200 changing words selected from etymological dictionaries (OED, 2009; Paul, 2002; Svenska Akademien, 2009)
- pre-annotation (rough filtering by one annotator)
- adding of control words with similar frequency properties
- sample 100 uses (30 for Latin) of each word per time period
- graded word sense annotation (Erk et al., 2013)

Data

	Eminute	die.ods - OpenOffica	e Calis
ut Somer			
anna la san A	A COLORADO	6	D
sentence 1	judgment	comment	sentence 2
Speaking of bread and butter reminds me that we' better eat ours before the coffee gets quite cold.	d		When the meal was over and they had finished their tea after they ate, Wang the Second took the trusty man to his elder brother's gate.
He agreed and began practicing his sleightofhand tricks to the great pleasure of some children , the same ones, I suspect, who had plagued me when was a child.			The daylight had long faded; her child lay calmly sleeping by her side; a candle was burning dimly on the stand.
He came to a crossroad and read the signs; to th south, Kenniston, 20 m.	0		As a result, we are at a crossroad: either school integration efforts will be abandoned in the South, or they will be pursued in the North as well.
His parents had left a lot of money in the bank an now it was all Measle's, but a judge had said that Measle was too young to get it.	1		Sherrell, it is said, was sitting on the bank of the river close by, and as soon as the men had disappeared from sight he jumped on board the schooner.
Hetty was mourned as dead: in every home her name was tenderly and sorrowingly spoken; old memories of her gay and minhful youth, of her cheery and busy womanhood, were revived and dwelt upon.			She does look like a fashionable young lady, but somehow I miss my little Rose, for children dressed like children in my day, "answered Aunt Plenty, peering through her glasses with a troubled look, for she could not imagine the creature before her ever sitting in her lap, running to wait upon her, or making the house gay with a child's bitthe presence."
But I never would return To my cold prison cell; what's life without liberty?			They went their way rejoicing, and with them passed the solitary ray of sunshine that streamed athwart the dark horrors of the emigrant ship, like the wandering pencil of light that sometimes visits the condemned cell of a prison.
Her hand pressed my arm with warm friendliness;			The arm of the statue had a natural form.
+ - fabelet	(a)		

Table 1: Annotation Table.

Scale

Identity Context Variance Polysemy Homonymy

- 4: Identical
- 3: Closely Related
- 2: Distantly Related
- 1: Unrelated

Table 2: Blank (1997)'s continuum of semantic proximity (left) and the DURel relatedness scale derived from it (right).

Identity

My <u>arm</u> hurts.

VS.

She grabbed his <u>arm</u> after he had told her the news.

Context variance

My <u>arm</u> hurts.

VS.

Look at the <u>arm</u> of the statue.

Polysemy

My <u>arm</u> hurts.

VS.

An arm of the sea.

My <u>arm</u> hurts.

VS.

The number of men under <u>arms</u> is no longer the decisive factor in warfare.

Diachronic Data

. . .

- (1) 1830 but I am bound and thrown into a dark cell.
- (2) 1851 ... be fit to burn in a jail; no, not in a condemned **cell**.
- (3) 1990 But I never would return To my cold prison **cell**. What's life without liberty?
- (4) 2006 She searched the bag for her **cell** as we headed toward the door.

Diachronic Data

	A		C	0
1	sentence 1	judgment	comment	sentence 2
2	but I am bound and thrown into a dark cell.			be fit to burn in a jail; no, not in a condemned cell.
3	but I am bound and thrown into a dark cell.			But I never would return To my cold prison cell; what's life without liberty?
4	but I am bound and thrown into a dark cell.			She searched the bag for her cell as we headed toward the door.
5	be fit to burn in a jail; no, not in a condemned cell.			But I never would return To my cold prison cell; what's life without liberty?
	be fit to burn in a jail; no, not in a condemned cell.			She searched the bag for her cell as we headed toward the door.
1	She searched the bag for her cell as we headed toward the door.			But I never would return To my cold prison cell; what's life without liberty?

Table 3: Annotation Table.

Diachronic Data

	A	8	c	D
. 1	sentence 1	judgment	comment	sentence 2
2	but I am bound and thrown into a dark cell.	1	1	be fit to burn in a jail; no, not in a condemned cell.
3	but I am bound and thrown into a dark cell.	4	•	But I never would return To my cold prison cell; what's life without liberty?
4	but I am bound and thrown into a dark cell.	-	2	She searched the bag for her cell as we headed toward the door.
8	be fit to burn in a jail; no, not in a condemned cell.	4	4	But I never would return To my cold prison cell; what's life without liberty?
	be fit to burn in a jail; no, not in a condemned cell.	4	2	She searched the bag for her cell as we headed toward the door.
7	She searched the bag for her cell as we headed toward the door.	1	2	But I never would return To my cold prison cell; what's life without liberty?

Table 4: Annotation Table.

Word Usage Graphs (WUGs)



Figure 3: Graph visualization four uses of cell.

Word Usage Graphs (WUGs)



Figure 4: Graph visualization four uses of cell.

Clustering

►

- correlation clustering (Bansal, Blum, & Chawla, 2004)
- optimization criterion: reduce (weighted) number of cluster-edge conflicts

$$L(C) = \sum_{e \in \phi_{E,C}} W(e) + \sum_{e \in \psi_{E,C}} |W(e)|$$

(i) (ii) (iii) (iv) (v)

finds the optimal number of clusters on its own
handles missing information (non-observed edges)
robust to errors by using the global information
respects the gradedness of word meaning
dominated in simulation study

Clustering



Figure 5: Graph visualization for uses of cell.

Clustering



Figure 6: Graph visualization for uses of *cell*.



Figure 7: Usage graph of Swedish *ledning*.



Figure 8: Usage graph of Swedish *ledning*.



Figure 9: Usage graph of Swedish ledning.



Figure 10: Usage graph of German Eintagsfliege.



Figure 11: Usage graph of German *Eintagsfliege*.



Figure 12: Usage graph of German Eintagsfliege.



Figure 13: Usage graph of German Festspiel.



Figure 14: Usage graph of German zersetzen.



Figure 15: Usage graph of German Abgesang.

The theoretical perspective

- which is the best model for WUGs?
- what is the psychological status of WUGs?

Stochastic Block Model (SBM)

- generative probabilistic model for random graphs with planted clusters
- canonical model for community detection
- simplest model of a graph with communities (Abbe, 2017)

Stochastic Block Model (SBM)



Figure 16: SBM cycle



Figure 17: Usage graph of German Abgesang.

The practical perspective

- how can WUGs be exploited for lexicography?
- how can WUGs be annotated efficiently?
- how can we further automate the lexicographic process?
- DURel annotation interface: https://www.ims.uni-stuttgart.de/data/durel-tool

Example of Use Pair

Home	Kontakt	Richtlinien	Zum Hauptme	nű Über Dutel
Ber	werte die	Semantische Ähnlichkeit der zwei Instanzen des Worte	s Donnerwetter	
Sata Und gien	, Do g; "sieh " her	nnerwetteri komm' her, Gesichmail nef Massikomur seinem Nachbar zu, der so * Geschmas eiti herber und er heter ihre alne Schaube vol versunter, denser, had	ben auch den kürzeren Weg über den vertrockneten Seegrund zu nundlicher Gegenstände unter die Jagen	ı seinem Acker
Sata	nerwetter n	Es wird hebey noch gemeider, daß aus der verbranden it großem Bitzen und Regen eristanden sey.	Gotzenbilderasche in einer Grube ein sehr plözilches und erschre product	skliches
0000	I Identisch Enger Bezu Entfornter	g Nanug		
0.0	: Kann nicht	entscheiden		
	laten absend	en		

Screenshot of visualized usage graph from DURel system's beta version.

Example WUG



Figure 18: Screenshot of annotation interface from DURel system's beta version.



Figure 19: Step 0: No information



Figure 20: Step 1: Initial clustering



Figure 21: Step 2: Cluster comparison



Figure 22: Step 3: Compare non-assignable uses



Figure 23: Step 4: Cluster comparison



Figure 24: Step 5: Cluster comparison

Bibliography

- Abbe, E. (2017). Community detection and stochastic block models: recent developments.
- Bansal, N., Blum, A., & Chawla, S. (2004). Correlation clustering. Machine Learning, 56(1-3), 89–113. doi: 10.1023/B:MACH.0000033116.57574.95
- Blank, A. (1997). Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen. Tübingen: Niemeyer.
- Bréal, M. (1897). Essai de sémantique. Hachette.
- Erk, K., McCarthy, D., & Gaylord, N. (2009). Investigations on word senses and word usages. In Proceedings of the joint conference of the 47th annual meeting of the acl and the 4th international joint conference on natural language processing of the afnlp: Volume 1 (pp. 10–18). Stroudsburg, PA, USA.
- Erk, K., McCarthy, D., & Gaylord, N. (2013). Measuring word meaning in context. Computational Linguistics, 39(3), 511–554.
- Kilgarriff, A. (1997). "I don't believe in word senses". Computers and the Humanities, 31(2).
- Kilgarriff, A. (2007, 11). Word senses. In (p. 29-46). doi: 10.1007/978-1-4020-4809-8_2
- McCarthy, D., Apidianaki, M., & Erk, K. (2016). Word sense clustering and clusterability. Computational Linguistics, 42(2), 245-275.
- Navigli, R. (2009). Word sense disambiguation: a survey. ACM Computing Surveys, 41(2), 1-69.
- OED. (2009). Oxford english dictionary. Oxford University Press.
- Paul, H. (2002). Deutsches Wörterbuch: Bedeutungsgeschichte und Aufbau unseres Wortschatzes (10. ed.). Tübingen: Niemeyer.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7, 573–605.
- Schlechtweg, D., McGillivray, B., Hengchen, S., Dubossarsky, H., & Tahmasebi, N. (2020). SemEval-2020 task 1: Unsupervised Lexical Semantic Change Detection. In Proceedings of the 14th International Workshop on Semantic Evaluation. Barcelona, Spain: Association for Computational Linguistics.
- Svenska Akademien. (2009). Contemporary dictionary of the Swedish Academy. The changed words are extracted from a database managed by the research group that develops the Contemporary dictionary.
- Weaver, W. (1949/1955). Translation. In W. N. Locke & A. D. Boothe (Eds.), Machine translation of languages (pp. 15–23). Cambridge, MA: MIT Press. (Reprinted from a memorandum written by Weaver in 1949.)