

# Sense and Sentiment

Francis Bond



Palacký University  
bond@ieee.org

Merrick Choo Yeu Herng



Nanyang Technological University  
merrick.emrys@gmail.com

13th Edition of its Language Resources and Evaluation Conference.  
LREC 2022-06-20



Faculty  
of Arts

# A sense-based sentiment lexicon

- Sentiment analysis and lexicon-based sentiment systems have proven to be very useful
- Many different lexicons exist, but mainly word-based
- However, sentiment depends on the meaning of a word in context:  
*plot*
  - ▶ “plan secretly, usually something illegal”
  - ▶ “devise the sequence of events in a literary work”
- Adding sentiment to a wordnet allows us to make finer-grained decisions
- It also allows us to measure the effect of semantic relations
- So we built one, based on existing corpus annotation, word-based lexicons and the semantic network:


**the Sentimental Wordnet**



Faculty  
of Arts

# Existing Lexicons: Description

We compare several existing English lexicons that have been shown to perform well and are available for research:

- The Glasgow Norms (GLAS) are a set of normative ratings for 5,553 English words on nine psycholinguistic dimensions: including , valence, concreteness, imageability, familiarity, and more.
- The labMT Lexicon created ~10,000 word sentiment lexicons for 10 languages including English
- The sentiment 140 (S140) lexicon was made by learning sentiment from a large twitter corpus
- The SO-CAL lexicon was made by a single researcher, and checked by a committee of three. The lexicon includes part of speech.
- Vader is a sentiment analysis system based on a large hand-built lexicon and a few heuristics.
- The WKW-SCI Sentiment Lexicon contains common words hand  tagged by 3 annotators

# Existing Lexicons: Features

Name	Size	<i>good</i>	Misc	#	Licence
GLAS	5,553	0.63	wrd	33	CC-BY
labMT	10,222	0.55	wrd	50	CC-BY-NC-SA
S140	62,468	0.16	ML	0	Research
SOCAL	6,091	0.60	lex+pos	1	CC-BY-NC-SA
VADER	7,506	0.47	wrd	10	MIT
WKW	28,955	0.67	lex+pos	3	CC-BY-NC-SA

This table shows the size of the lexicon, the score it gives for the word *good*, the type (wrd is word based, lex+pos is based on the lexeme + part of speech and ML is machine learned), the number of raters and the license. There is quite a bit of variance in both values and coverage!



# Existing Lexicons: Comparison

Name	labMT	S140	SOCAL	VADER	WKW
GLAS	0.95	0.48	0.88	0.96	0.81
labMT		0.41	0.86	0.95	0.73
S140			0.51	0.57	0.35
SOCAL				0.90	0.86
VADER					0.90

- Calculated using Pearson's  $\rho$  for all entries that appear in each pair of lexicons
- The lexicons built with multiple annotators correlate well
- The lexicons with fewer annotators correlate OK
- The ML lexicon is very different:  
it fits very well to the domain it was trained on, but does not generalize

# A sense and sentiment tagged corpus: NTU-MC

- stories taken from the Sherlock Holmes canon
- Annotation has proceeded in three phases.
  - 1 the stories *The Adventure of the Dancing Men* and *The Adventure of the Speckled Band* were annotated in Chinese, English and Japanese and the results compared across languages. For example, in (1), the appropriate senses of *false*<sub>a</sub> and *villain*<sub>n</sub> are annotated as -0.34 and -0.64 respectively.

*If we make one false move the villain may escape us yet .*
  - 2 the sense annotation was done by multiple students, and the sentiment annotation was done by one RA (*The Redheaded League*, *A Scandal in Bohemia* and half of the novel *The Hound of the Baskervilles*)
  - 3 both sense and sentiment were annotated by multiple student annotators, with a round of comparison and harmonization (*The Adventure of the Final Problem*, *The Adventure of the Naval Treaty* and the rest of *The Hound of the Baskervilles*)



# NTU-MC Summary

Corpus	Sents	Words	Concepts	Pos.	Neg.	$\rho$	Overlap
Phase1	1,199	23,093	13,077	983	1,244	0.78	130
Phase2	3,021	54,698	30,287	805	869	0.68	225
Phase3	3,250	60,702	33,407	1,556	4,710	0.80	240
NTUMC	7,470	138,493	76,771	3,344	6,823	0.75	339

- Neg and Pos show the number of concepts with positive and negative sentiment (above a threshold of 0.05)
- $\rho$  is the agreement with MICRO-WNOP
- Overlap is the number of entries in MICRO-WNOP annotated in the corpus

MICRO-WNOP is a very small **sense-tagged** sentiment lexicon used to evaluate SentiWordNet and build ML-SentiCon. We used version 3.0, which has 1,054 entries, of which 579 are non-zero.



# Effects of semantic relations

- Semantic relations have been used in creating resources such as SentiWordNet and ML-SentiCon, but there has been no empirical analysis of the effects of semantic relations on sentiment score.
- We use the corpus annotations to test the effects of both concept and sense-based relations





# Effects of concept-based semantic relations

Relation	All	Score	Non-Zero	Score
similar	833	+0.109	450	+0.202
hyponym	851	+0.075	312	+0.206
holo location	0	+nan	0	+nan
holo member	24	+0.007	2	+0.089
holo part	160	+0.013	12	+0.171
holo portion	0	+nan	0	+nan
holo substance	8	+0.021	1	+0.170
entails	58	+0.057	23	+0.143
causes	24	+0.093	9	+0.249

- We consider the difference in sentiment of two concepts linked by a relation
- In general they are close, but still a little unpredictable



# Effects of sense-based semantic relations

Relation	All	Score	Non-Zero	Score
synonym	1,408	+0.069	551	+0.184
antonym	249	+0.217	116	+0.467
ant opposite	249	+0.081	116	+0.175
also	1	+0.000	0	+nan
derivation	1,293	+0.071	507	+0.180
pertainym	183	+0.110	115	+0.176

- We consider the difference in sentiment of two senses linked by a relation
- In general they are closer
- If we consider antonym to link to the opposite score, then it is also close



# Error Analysis

- Some annotation was wrong:
  - ▶ *His dark eyes, glaring out of the white mask of his face, were full of horror and astonishment as he gazed from Sir Henry to me.*
  - ▶ should have been “anemic looking from illness or emotion” not “being of the achromatic color of maximum lightness”
- Umbrella terms spoil the hyponym similarity
  - ▶ *love* 0.64 and *hate* -0.95 are both hyponyms of *emotion* 0.0



# The Sentimental Wordnet

- Seed the lexicons

- ▶ senses in the corpus take the average of their annotation
  - ★ *good* “having desirable or positive qualities” = 0.53
- ▶ monosemous senses take values from the word based lexicons
  - ★ *damnable* “deserving a curse” = -0.425
- ▶ word+pos monosemous lexemes take values from the lex+pos based lexicons
  - ★ *perk* noun “an incidental benefit” = -0.2
  - ★ *perk* verb “gain or regain energy” = 0.0

- Propagate the values

- ▶ Senses with values linked by **derivation**, **pertainym** and **antonym** propagate their values
- ▶ Senses in the same concept (synonyms) take the same values
- ▶ E.g. *agreeability* (0.47), *agreeably* (0.37) and *disagreeable* (-0.64) propagate to *agreeable* giving it a value of 0.49



# Initial Lexicons

Method	Size	$\neq 0$	$\rho$	Cover	$\neq 0$
VADER	1,700	1,700	0.95	40	39
GLAS	842	836	0.87	35	24
labMT	1,545	1,535	0.77	36	17
WRD	3,550	3,535	0.88	79	55
SOCAL	2,078	2,078	0.84	85	78
WKW	14,668	5,134	0.87	198	140
POS	15,002	5,744	0.85	207	148
LEX	16,499	8,179	0.86	217	153
NTUMC	11,154	2,989	0.75	339	200
ALL	26,325	10,793	0.81	471	296

- WRD is monosemous seeds from the word based lexicons
- POS is monosemous seeds from the lex+pos based lexicons
- LEX is all the lexicons together
- NTUMC is the corpus based lexicon



# Propagate from Lexicons

Method	Size	$\neq 0$	$\rho$	Cover	$\neq 0$
LEX	16,499	8,179	0.86	217	153
LEX P	40,477	21,964	0.85	366	243
LEX P <sup>2</sup>	52,740	29,608	0.84	476	321
LEX P <sup>3</sup>	60,576	35,603	0.83	561	374
LEX P <sup>4</sup>	65,226	39,630	0.82	612	401
LEX P <sup>5</sup>	67,880	42,367	0.82	634	415
LEX P <sup>6</sup>	69,394	44,329	0.81	648	423
LEX P <sup>7</sup>	70,224	45,688	0.81	656	430
LEX P <sup>8</sup>	70,719	46,596	0.81	659	430
LEX P <sup>9</sup>	71,010	47,251	0.81	659	430
LEX P <sup>10</sup>	71,161	47,661	0.81	662	432
LEX P <sup>11</sup>	71,242	47,955	0.81	662	432

Human correlation in MICRO-WNOP is 0.88, 579 non-zero values



# Propagate from Lexicons and Corpus

Method	Size	$\neq 0$	$\rho$	Cover	$\neq 0$
ALL	26,325	10,793	0.81	471	296
ALL P	57,987	27,643	0.82	614	378
ALL P <sup>2</sup>	69,788	35,683	0.82	695	431
ALL P <sup>3</sup>	75,177	40,967	0.83	742	457
ALL P <sup>4</sup>	77,546	44,125	0.82	758	464
ALL P <sup>5</sup>	78,483	46,149	0.82	765	466
ALL P <sup>6</sup>	78,905	47,454	0.82	767	467
ALL P <sup>7</sup>	79,075	48,225	0.82	768	467
ALL P <sup>8</sup>	79,131	48,772	0.82	768	467
ALL P <sup>9</sup>	79,150	49,092	0.82	768	467
ALL P <sup>10</sup>	79,155	49,343	0.82	768	467
ALL P <sup>11</sup>	79,155	49,466	0.82	768	467

Human correlation in MICRO-WNOP is 0.88, 579 non-zero values



# Conclusions

- We have created a new sense-based sentiment lexicon **the Sentimental Wordnet**
- It has sentiment values for around 79,000 senses of which over half have some sentiment.
- Correlation with the gold standard is 0.82 (Pearson's  $\rho$ ) slightly below human (at 0.88)
- The data and code are released under an open licence at <https://github.com/bond-lab/sentimental> (MIT)
- We have built a sense-based sentiment system **sensitive** to go with this lexicon, available at <https://github.com/bond-lab/sensitive>





# Acknowledgements

- We would like to thank all the students involved in the tagging, from HG2002 (2011–2015, 2018–2021) and HG8011 (2016, 2018, 2019, 2021), especially Vigneswari Thirumoolan who did the Phase 2 tagging.
- This work was supported by the European Regional Development Fund Project “Sinophone Borderlands –Interaction at the Edges”, CZ.02.1.01/0.0/0.0/16\_019/0000791 and co-financed by the Ministry of Education, Youth and Sports, Czech Republic.



EUROPEAN UNION  
European Structural and Investment Funds  
Operational Programme Research,  
Development and Education



Faculty  
of Arts