Sentiment Annotation

in the NTU Multilingual Corpus (NTU-MC)

2nd Wordnet Bahasa Workshop (WBW2016)

Francis Bond, Tomoko Ohkuma,
Luis Morgado Da Costa, Yasuhide Miura, Rachel Chen,
Takayuki Kuribayashi, **Wang Wenjie**

January 15, 2016

NANYANG TECHNOLOGICAL UNIVERSITY

Presentation Roadmap

1. Introduction
2. Concept-level Annotation
3. Chunk-level Annotation
4. The Annotation Tools
5. Q & A

NANYANG TECHNOLOGICAL UNIVERSITY

Introduction

- Sentiment annotation of NTU-MC
  - Concept level
  - Chunk level
- Goal: Look at how different languages encode sentiment
  - How the results compare to those of existing resources

NANYANG TECHNOLOGICAL UNIVERSITY

Introduction: Text

- *The Adventure of the Speckled Band** (Arthur Conan Doyle)
  - English (original)
  - Chinese
  - Japanese
  - Next: Indonesian

NANYANG TECHNOLOGICAL UNIVERSITY

Introduction: Corpus

Language	Sentences	Words	Concepts	Distinct Concepts
English	599	11,741	6,513	2,273
Chinese	619	12,662	8,367	2,547
Japanese	702	13,906	4,817	1,694

NANYANG TECHNOLOGICAL UNIVERSITY

Introduction: Other Sentiment Corpora and Resources

- There exist other sentiment corpora and lexical resources
  - Stanford Sentiment Treebank (Socher et al., 2013)
  - Sentiwordnet (Baccianella et al., 2010)
  - MLSentiCon (Cruz et al., 2014)
- Few multi-lingual ones
- None for Asian languages

NANYANG TECHNOLOGICAL UNIVERSITY

Sentiment Annotation in NTU-MC

- Using IMI (A Multilingual Semantic Annotation Environment)
- One annotator per language
- Continuous Scale of [-100 to 100]
  - Seven (7) main steps:

Steps	1	2	3	4	5	6	7
Score	-95	-64	-34	0	34	64	95

NANYANG TECHNOLOGICAL UNIVERSITY

Sentiment Annotation in NTU-MC: Guidelines

Score	Example	Example	Example	Corpus Examples
95	fantastic	very good		perfect, splendidly
64	good	good		soothing, pleasure
34	ok	sort of good	not bad	easy, interesting
0	beige	neutral		puff
-34	poorly	a bit bad		rumour, cripple
-64	bad	bad	not good	hideous, death
-95	awful	very bad		deadly, horror-stricken

NANYANG TECHNOLOGICAL UNIVERSITY

Concept-level Annotation

Concept-level Annotation

- Annotate concepts that show clear positive or negative sentiment:
  - Eg: Happy, sad, angry
- Context independent
- Concepts can be single- or multi-word expressions
  - Eg: *give rise* "produce", *break down* "to no longer function"
- Modifiers such as ***very*** and ***not*** are not tagged

NANYANG TECHNOLOGICAL UNIVERSITY

Concept-level Annotation: Cross-lingual Comparisons

Correlation between the different language pairs:

Pair	ρ	**# samples**
Chinese-English	.70	3,204
Chinese-Japanese	.78	1,795
English-Japanese	.70	1,862

- Matched by concept, and get their average sentiment score per language
- Repeated by their minimum frequency in both languages

NANYANG TECHNOLOGICAL UNIVERSITY

Concept-level Annotation: Cross-lingual Comparisons

- Agreement in polarity was high between languages
- But rarely identical — differences mainly in intensity/magnitude

NANYANG TECHNOLOGICAL UNIVERSITY

Concept-level Annotation: Cross-lingual Comparisons (cont'd)

- Agreement in polarity was high between languages
- But rarely identical — differs **mainly in intensity/magnitude**

- 02433000-a "showing the wearing effects of overwork or care or suffering"
  - (EN) *drawn* and *haggard*: **-64** (average)
  - (CH) 憔悴 qiáocuì: **-34**
- 01386883-a "relatively large in size or number or extent"
  - (EN) *great*: **+45.2**
  - (JP) 大きい "big": **+0** (neutral)

NANYANG TECHNOLOGICAL UNIVERSITY

Concept-level Annotation: Cross-lingual Comparisons (cont'd)

- Agreement in polarity was high between languages
- But rarely identical — differs **mainly in intensity/magnitude**
- Inter-annotator differences; subjective
- Lemmata in the same synset might have slightly different sentiment ratings (Eg: ***grotesque*** and ***fantastic*** in 00967646-a)
- A lemma might have senses in other synsets, which might be more/less positive. This could cause cross-concept "bleeding".
  - This cross-concept sensitivity might differ from language to language

NANYANG TECHNOLOGICAL UNIVERSITY

Concept-level Annotation: Comparing to Other Resources

- Compared to **Sentiwordnet** and **MLSentiCon**
  - Both are automatically-generated resources
- Compared at synset level, for concepts that appear at least once in any language, and averaged over all occurrences.

Concept-level Annotation: Comparing to Other Resources (cont'd)

Correlation between the different resources:

Pair	ρ	**# samples**	
SentiWN-MLSenticon	.48	4,202	(.42 for all 121,861)
NTUMC-SentiWN	.41	4,202	
NTUMC-MLSenticon	.45	4,202	

They don't seem to agree very well!

NANYANG TECHNOLOGICAL UNIVERSITY

Concept-level Annotation: Comparing to Other Resources (cont'd)

- None of the resources agree very well with each other
- The other resources agree more with each other than with NTU-MC
  - But only barely.

Concept-level Annotation: Comparing to Other Resources (cont'd)

- Neutral (+0) in NTU-MC, but non-neutral in the other resources
  - ***be*** *is* ***+0.125*** *in Sentiwordnet*
  - ***April*** is **-0.125** in *MLSenticon*
- Strong score in NTU-MC, but neutral in the other resources
  - ***Violence****, which is* ***-64*** *in NTU-MC, was neutral in the other two resources*

NANYANG TECHNOLOGICAL UNIVERSITY

Chunk-level Annotation

Chunk-level Annotation

- Units larger than a concept (phrase, sentence)
- Context-dependent
- Chunks are "built up" with other chunks (usually contiguous), if there is a change of strength and/or polarity
- *Operators* (***very***, ***not***, etc) can change the strength and/or polarity
- Words like ***think***, ***feel***, etc, reduce the strength of the sentiment
- Questions are rated neutral

NANYANG TECHNOLOGICAL UNIVERSITY

Chunk-level Annotation: Examples

Example 1:

"Do they think this is very good?"

+64	good	
+95	very good	Operator ***very*** increases strength
+95	this is very good	
+90	they think this is very good	***Think*** reduces* strength
+0	Do they think this is very good?	Questions are rated neutral

Chunk-level Annotation: Examples (cont'd)

Example 2:

"I do not understand"

+34	understand	
-34	not understand	Operator ***not*** flips polarity
-34	I do not understand	Overall sentence senti-rating

Chunk-level Annotation: Examples (cont'd)

Example 3:

"It is not beautiful"

+64	beautiful	
-34	not beautiful	Polarity flips, but is not mirrored
-34	It is not beautiful	Overall sentence senti-rating

Chunk-level Annotation: Examples (cont'd)

Example 4:

"The horse raced past the barn"

+0 The horse raced past the barn Neutral

Chunk-level Annotation: Annotator Agreement

Sentence correlation between the different languages:

Pair	ρ	**# samples**
English-Chinese	.55	561
English-Japanese	.61	450
Chinese-Japanese	.70	390

Only for sentences that aligned one-to-one.

NANYANG TECHNOLOGICAL UNIVERSITY

Chunk-level Annotation: Annotator Agreement

- There was less agreement for chunk-level annotations
- Majority of the sentences were neutral (+0)
- Annotators found it hard to decide chunk boundaries
  - When in doubt, more-is-more approach adopted
  - Stanford's sentiment treebank creates chunks using every word in the sentence

NANYANG TECHNOLOGICAL UNIVERSITY

Supra-Chunk Level Annotation?

- Devices operating above the surface chunk
- Eg: Sarcasm (can sometimes reverse polarity)
  - The sentences are tagged as if they are non-sarcastic.
  - However, annotators are instructed to indicate "SARCASM" in the comments box
- Pragmatics? [including questions?]

NANYANG TECHNOLOGICAL UNIVERSITY

Supra-Chunk Level Annotation? (cont'd)

- Devices operating above the surface chunk
- Interlocutors can affect senti-rating
- Eg: **"He's dead."**
  - Sententially, rather morbid.
  - However!
  - More positive sentiment if "his" death is good news to the interlocutors.
- Presently, this is not taken into account when annotation is performed

NANYANG TECHNOLOGICAL UNIVERSITY

Greetings, Interjections, Vituperatives

- Greetings
  - Good morning!, G'day!, Hello!
  - Good evening vs. Good night
  - Slightly positive in English
  - Neutral in Japanese
- Interjections
  - Good grief! Hell's Bells!
  - Yikes! Zoinks! Great moons of Neptune!

NANYANG TECHNOLOGICAL UNIVERSITY

Greetings, Interjections, Vituperatives, Curses

- Greetings
  - Good morning!, G'day!, Hello!
  - Good evening vs. Good night
  - Slightly positive in English
  - Neutral in Japanese
- Interjections
  - Good grief! Hell's Bells!
  - Yikes! Zoinks! Great moons of Neptune!
- Curses
  - ****! ******!
  - !@#, vs. ****?

NANYANG TECHNOLOGICAL UNIVERSITY

Sentiment Annotation Demo

Sentiment Annotation Demo (S.A.D)

Ongoing Work

- Extend to other languages
  - Bahasa Indonesian being the most upcoming
- Extend to other texts
  - Other texts in the NTU-MC
  - Software reviews
- Reflect sentiment holder and/or targets of the sentiment
- Train and automate process

NANYANG TECHNOLOGICAL UNIVERSITY

Q & A

NANYANG TECHNOLOGICAL UNIVERSITY

References

Baccianella, S., Esuli, A., and Sebastiani, F. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Chair), N. C. C., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., and Tapias, D., editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

Cruz, F. L., Troyano, J. A., Pontes, B., and Javier Ortega, F. (2014). Building layered, multilingual sentiment lexicons at synset and lemma levels. *Expert Systems with Applications*.

Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., and Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *EMNLP*.

NANYANG TECHNOLOGICAL UNIVERSITY