

Attitude Beyond the Word Level

2013 NSF-CDI Specialist Meeting Knowledge Discovery in
Cyberspace and Big Data (SDSU)
Pranav Anand · UC Santa Cruz · *panand@ucsc.edu*

August 7, 2013

Preview

Why we might care about sentiment

- ▶ **private states**: judgment/emotion an evaluative **source** bears towards a **target** (entity, event, or proposition) (Weibe et al. 2003)
- ▶ sentiment analysis has been applied to a wide variety of “siding” problems in opinion analysis
 - ▶ election prediction
 - ▶ product review abstraction
 - ▶ political position profiling

Next generation opinion analysis

- ▶ **Clarity**: clear, grounded theories of what “sentiment” and “opinion” are
- ▶ **Coverage**: opinion extraction from more sophisticated, exotic (naturalistic?) texts
- ▶ **Context**: models of
 - ▶ genre
 - ▶ domain
 - ▶ pragmatic environment
 - ▶ semantic composition

Two approaches to sentiment

Two varieties of modeling sentiment

- ▶ **Lexical**: words are associated with constant values in relevant sentiment space

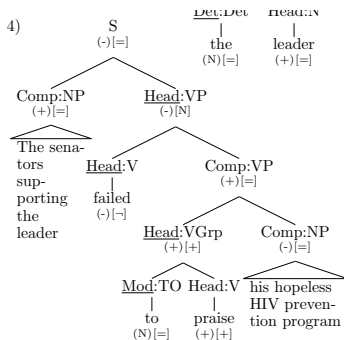
```
type=weaksubj len=1 word1=abdicate pos1=verb stemmed1=y priorpolarity=n  
type=strongsubj len=1 word1=aberration pos1=adj stemmed1=n priorpolarit:  
type=strongsubj len=1 word1=aberration pos1=noun stemmed1=n priorpolarit:  
type=strongsubj len=1 word1=abhor pos1=anypos stemmed1=y priorpolarity=n  
type=strongsubj len=1 word1=abhor pos1=verb stemmed1=y priorpolarity=n  
type=strongsubj len=1 word1=abhorred pos1=adj stemmed1=n priorpolarity=n  
type=strongsubj len=1 word1=abhorrence pos1=noun stemmed1=n priorpolarit:  
type=strongsubj len=1 word1=abhorrent pos1=adj stemmed1=n priorpolarity:n  
type=strongsubj len=1 word1=abhorrently pos1=anypos stemmed1=n priorpol:  
type=strongsubj len=1 word1=abhors pos1=adj stemmed1=n priorpolarity=n  
type=strongsubj len=1 word1=abhors pos1=noun stemmed1=n priorpolarity=n  
type=strongsubj len=1 word1=abidance pos1=adj stemmed1=n priorpolarity:n  
type=strongsubj len=1 word1=abidance pos1=noun stemmed1=n priorpolarity:
```

Wilson et al. (2005)

- ▶ presence of word: presence of sentiment
- ▶ more carefully: sentiment triples of ⟨source, target, sentiment⟩

Two varieties of modeling sentiment

- ▶ **Compositional**: words are associated with n -ary functors into sentiment space
 - ▶ **flipper**: opposite polarity to its argument (fail)
 - ▶ **preserver**: same polarity to its argument (to)
 - ▶ **constant**: lexically determined polarity (praise)



Moilanen & Pulman (2007)

Two varieties of modeling sentiment

- ▶ **Lexical**: words are associated with constant values in relevant sentiment space
- ▶ **Compositional**: words are associated with n -ary functions over sentiment space
- ▶ Choice point: determining **phrasal sentiment**

- (1) a. John isn't helpful.
b. John defeated Mary.

logical operators

event-level evaluativity

Event-level evaluativity

- ▶ Sentiments towards participants impacts sentiment of event

(2) a. **The hero** defeated **the villain**.

default positive

b. **The villain** defeated **the hero**.

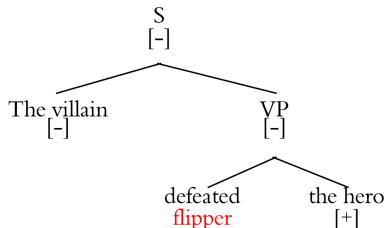
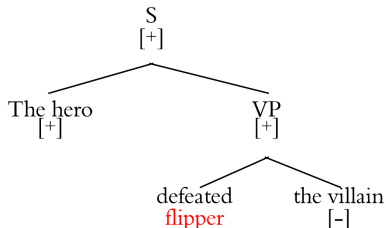
default negative

Event-level evaluativity

- ▶ Sentiments towards participants impacts sentiment of event

- (3) a. **The hero** defeated **the villain**. default positive
- b. **The villain** defeated **the hero**. default negative

- ▶ Explicable in a compositional system



Event-level evaluativity

- ▶ Event sentiment determined by verb **outcome type** (Anand & Reschke 2010, Neviarouskaya et al. 2010)
- ▶ verbs are **flippers** or **preservers**, given final state of event
 - ▶ injure/destroy affected object: **flippers**
 - ▶ create/benefit affected object: **preservers**
 - ▶ lack object: **flippers**
 - ▶ have object: **preservers**
- ▶ Outcome type estimated from proxies (Reschke & Anand 2011)
- ▶ heuristic #1: MPQA (Wilson et al. 2005)
 - ▶ 380 + words \mapsto **preservers** (abound, absolve, accomplish, ...)
 - ▶ 869 - words \mapsto **flippers** (abandon, abase, abolish, ...)
- ▶ heuristic #2: Framenet (Ruppenhofer et al., 2006)
 - ▶ 124 frames manually extracted based on outcome type
 - ▶ 146 **preservers**: abet, begin, have, ...
 - ▶ 341 **flippers**: leave, lack, ...

Compositional Advantage: Top-Down Inference

- ▶ In a compositional system, known event-level evaluativity can allow inference to entity sentiment (Reschke & Anand 2011)

(4) We liked_{e_{like}} him because he so clearly disliked_{e_{dislike}} Margaret Thatcher.

LEXICAL EVALUATIVITY: $\langle \text{we, him, +} \rangle$; $\langle \text{he, M.T., -} \rangle$

PRAGMATIC INFERENCE: $\langle \text{we, e}_d\text{islike, +} \rangle$ (via *because*)

COMPOSITIONAL INFERENCE: $\langle \text{we, M.T., -} \rangle$

Comparing the strategies

Is this approach useful in practice?

- ▶ But perhaps people tend to choose verbs that lexically convey sentiment.

(5) The hero $\left\{ \begin{array}{l} \text{killed} \\ \text{eliminated} \\ \text{took down} \\ \text{took care of} \end{array} \right\}$ the monster.

Problem #1: Mixing of perspectives

- ▶ Neviarouskaya et al. (2010): distinguish between *internal* and *external* sources; distinct composition rules
- ▶ Moilanen & Pulman (2007) conflate Narrator source and Participant source, creating inconsistent lexicon
 - ▶ *defeat* marked as **flipper**
 - ▶ flips for Narrator
 - ▶ constant + for subject
 - ▶ *hate* marked as **constant** +
 - ▶ flips (?) for Narrator
 - ▶ constant + for subject

Problem #2: Lack of relevant evaluation data

- ▶ Moilanen & Pulman evaluate on the Headlines Corpus (Strapparava and Mihalcea 2007)
- ▶ annotators asked to provide polarity for whole headline
- ▶ problem: gestalt measures don't necessarily target event-level evaluativity

(6) Outcry at N Korea 'nuclear test'

(7) Minister slams airlines on carbon

- ▶ uncontrolled nuclear weapons and pollution are bad
- ▶ but are public outcry and calling out bad events bad?

Problem #2: Lack of relevant evaluation data

- ▶ Neviarouskaya et al. (2010) evaluate on 1000 sentences culled from online confessional site (Experience Project)
- ▶ but evaluation materials only annotated for internal (participant) perspective

Problem #2: Lack of relevant evaluation data

- ▶ MPQA corpus (Wiebe et al. 2005)
 - ▶ separates Narrator and participant sources
 - ▶ provides annotations for both NPs and VPs (among others)
 - ▶ but not all NPs are annotated
- (8) He [Sharon] has destroyed the peace process and the Oslo agreements.
 - a. destroyed: —
 - b. he has destroyed...agreements: —
 - c. no annotations for inanimates (peace process, Oslo agreements)

Political Ads (Reschke & Anand 2012)

- ▶ 161 TV political ads from 2008 Presidential campaign
 - ▶ short
 - ▶ designed to be engaging
 - ▶ high density of event-level evaluativity
 - ▶ important sentiment genre
 - ▶ text & video easily available

Sample

In tough times, who will help Michigan's auto industry? Barack Obama favors loan guarantees to help Detroit retool and revitalize. But John McCain refused to support loan guarantees for the auto industry. Now he's just paying lip service. Not talking straight. And McCain voted repeatedly for tax breaks for companies that ship jobs overseas, selling out American workers. We just can't afford more of the same.

Sample

In [tough times], who will [help [[Michigan]’s auto industry]]? [Barack Obama] [favors [loan guarantees] to [help [Detroit] retool and revitalize]]. But [John McCain] [refused to [support [loan guarantees] for [the auto industry]]]. Now [he]’s just [paying [lip service]]. Not [talking straight]. And [McCain] [voted repeatedly for [tax breaks] for [[companies] that [ship [jobs] overseas]]], [selling out [American workers]]. [We] just can’t [afford [more of the same]].

- ▶ All NPs & VPs extracted via Stanford Parser
- ▶ VPs headed by auxiliaries and modals ignored

4 Perspectival Questions

1. **Narrator** : How does the narrator want you to feel about **XP** ?
 2. **You** : How do you to feel about **XP** ?
 3. **Generic** : How do people in general feel about **XP** ?
 4. **Controversiality** : How controversial is **XP** ?
- ▶ Included 'Doesn't Make Sense' button for all response categories

Instructions

When answering the questions, it is important to focus on the highlighted text. Here's an example:

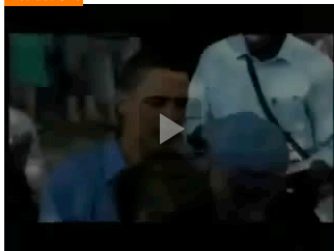
Obama doesn't care about our troops .

Even though the whole sentence is a criticism of Obama, the highlighted portion is positive because “caring about our troops” is a good thing.

Consequently, the answer to question 1 (how the narrator feels) should be positive. You should answer questions 2-4 by asking what you and society in general thinks of “caring about our troops.”

Annotation Setup

instructions



Barack Obama: "I'm Barack Obama, and I approve this message." Announcer: In tough times, who will help Michigan's auto industry? Barack Obama favors loan guarantees to help Detroit retool and revitalize. But John McCain refused to support loan guarantees for the auto industry. Now he's just paying lip service. Not talking straight. And McCain voted repeatedly for tax breaks for companies that ship jobs overseas, selling out American workers. We just can't afford more of the same. PFB: OBAMA FOR AMERICA

Back

Next

Submit

6 of 32

How does **the narrator** want you to feel about the highlighted expression?



Doesn't Make Sense

How do **you** feel about the highlighted expression?



Doesn't Make Sense

In your opinion, how controversial is the highlighted expression in American society?

no agreement

complete agreement

Doesn't Make Sense

In your opinion, how do **people in general** feel about the highlighted expression?



Doesn't Make Sense

Annotation Setup

- ▶ Mechanical Turk
 - ▶ \$0.40 per HIT, 4-5 annotators per transcript
 - ▶ restricted to U.S.
 - ▶ no qualifications or training
- ▶ 206 workers annotated the 161 transcripts ($\mu = 4.8/\text{transcript}$)
 - ▶ 3945 NP tokens, 24892 total annotations
 - ▶ 1549 VP tokens, 9800 total annotations
- ▶ filtering
 - ▶ 74 low quality annotators (removed for analysis)
 - ▶ 132 high quality annotators ($\mu = 3.9/\text{transcript}$)

Experimental Setup

- ▶ 635 VPs with NP direct object selected as well as associated NPs
- ▶ 272 (44%) in our lexicons above
- ▶ lexical: 65.1% accuracy; compositional: 84.2%

	LEXICAL	
<u>COMPOSITIONAL</u>	CORRECT	INCORR.
CORRECT	155	74
INCORR.	22	21

More naturalistic data

Is this enough?

- ▶ Do the results of election ads carry over to the Big Data zoo?

Is this enough?

- ▶ Do the results of election ads carry over to the Big Data zoo?
- ▶ Short answer: I don't know
- ▶ Long answer: I have doubts that the compositional model scales easily

Case Study: Informal online political discussion

- ▶ Online debate sites: portals for debate “poll” creation and participation (voting + commenting)
 - ▶ comments are pithy position statements
 - ▶ comments are automatically sided as pro or anti
- ▶ 1113 two-sided debates on 14 topics from Convinceme.net (Walker et al. 2012)

Ninjas versus Pirates

History



Jan 03, 2007



219

125

53

24

23

22

20

12

11

237 debaters

993 votes

+ Add Argument

577

Ninjas historically have kicked Pirates' butts, and will continue to do so

VOTE



king

Jan 03, 2007

112 convinced



[Rebuttal](#)

Ninjas have the advantage on land. I mean, you can't even see them until it's too late, and then you'll never see them because you have no head.

But on sea, pirates are dangerous. But a ninja could sneak on board. And pirates are too loud. Yell and say yar, and are usually drunk, so it is all ninja!



mknorpp

Jan 04, 2007

53 convinced



[Rebuttal](#)

Real pirates are and were historically horrible people killing and stealing from innocent people. Ninjas were people fighting for what was right - their freedoms and families. And they are way hotter. (Johnny Depp is the exception)



king

Jan 11, 2007

31 convinced



[Rebuttal](#)

OH, come one people!! Obvious answer is the ninja. Not as "cool" as the pirate, but think about this way:

When you were a kid what did you want to be? Everyone I know wanted to be a ninja. GI Joe? No pirates there. Lot's of ninjas though. Oh the countless number of Russians I killed in my backyard pretending I was a ninja.

How cool was it when your friend, or you yourself bought/made your first shakonen (spelling). And you threw it at some board until it had too many holes, or it accidentally hit your sister. Peter Pan fought pirates, but he wishes he was a ninja.

+ Add Argument

416

Pirates historically have kicked Ninjas' butts, and will continue to do so

VOTE

Wilbur

Jan 03, 2007

69 convinced



[Rebuttal](#)

Ah, the anachronism that is the ninja! Are there even any ninjas left in the world, to fight the pirates? I am pretty sure that the true ninjas went the way of the samurai (speaking of which, in the movie THE LAST SAMURAI, ninjas and samurais duke it out, and I am pretty sure that was anachronistic too) I say NAY!

Yet modern pirates abound! See The Outlaw Sea: A World of Freedom, Chaos, and Crime by William Langewiesche. Pirates still kick trash, even in a world of atom bombs and blackberries. In fact, pirates probably use blackberries. Blackberries, eyepatches and parrots. Yar, matey!



king

Jan 03, 2007

48 convinced



[Rebuttal](#)

Although, I need to add that pirates have numbers. A whole crew of people. Not only that, but they are not bound by morality like ninjas and their codes. Pirates will kill anyone, at anytime, for any reason.

Wilbur

Jan 05, 2007

33 convinced



[Rebuttal](#)

That is just prejudice against Monocolurism! One-eyed people are perfectly capable of seeing ninja stars.

Wilbur

Jan 04, 2007

20 convinced



[Rebuttal](#)

Two problems with the last pro-ninja argument... 1) it implies that ninjas are NOT horrible people historically... which is NOT TRUE!!

Corpus details

Topic	Posts	Rebuttals	P/A	A > 1p	Length
Cats v. Dogs	148	40%	1.68	26%	242
Firefox vs. IE	218	40%	1.28	16%	167
Mac vs. PC	126	47%	1.85	24%	347
Superman/Batman	140	34%	1.41	21%	302
2nd Amendment	134	59%	2.09	45%	385
Abortion	594	70%	2.82	43%	339
Climate Change	202	69%	2.97	40%	353
Communism vs. Capitalism	212	70%	3.03	47%	348
Death Penalty	324	62%	2.44	45%	389
Evolution	798	76%	3.91	55%	430
Exist God	844	77%	4.24	52%	336
Gay Marriage	505	65%	2.12	29%	401
Healthcare	110	80%	3.24	56%	280
Marijuana Legalization	214	52%	1.55	26%	423

Siding results

- ▶ siding classifiers per topic using
 - ▶ unigrams
 - ▶ sentiment word classes
 - ▶ general word classes (LIWC)
 - ▶ part of speech
 - ▶ grammatical dependencies
- ▶ this is not easy (nowhere near human performance)
- ▶ domain effect: results vary by topic
- ▶ tough to beat unigrams, though POS dependencies do best

Siding results

	Human	Uni	LIWC	dep	GdepP	GdepO	GdepL	All
Cats v. Dogs	94	66	56	63	67	61	70	65
Firefox v. IE	74	52	63	43	52	63	55	56
Mac v. PC	76	47	45	54	60	57	60	47
Superman v. Batman	89	56	42	53	55	38	52	58
2nd Amendment	69	60	53	56	47	55	53	65
Abortion	75	51	52	59	55	50	55	53
Climate Change	66	58	56	42	50	50	49	58
Comm. v. Capitalism	68	48	53	46	56	51	50	50
Death penalty	79	49	54	49	52	47	52	49
Evolution	72	54	48	57	57	56	55	55
Existence of God	73	52	51	54	55	53	53	54
Gay Marriage	88	61	56	61	59	62	61	60
Healthcare	86	46	48	47	58	54	55	45
MJ Legalization	81	53	58	45	51	49	54	55
Weighted avg.	77	54	52	54	56	53	55	55

Three reasons sentiment didn't help

- ▶ **discourse context**: many texts are replies (esp. in contentious domains)
 - ▶ including previous post elevates results

Siding results with context

	Human	Uni	LIWC	dep	GdepP	GdepO	GdepL	All
Cats v. Dogs	94	69	65	75	70	69	72	70
Firefox v. IE	74	57	66	60	57	60	57	58
Mac v. PC	76	55	55	52	55	55	60	56
Superman v. Batman	89	60	59	61	58	49	55	61
2nd Amendment	69	69	58	52	60	56	55	69
Abortion	75	58	57	65	65	60	64	60
Climate Change	66	65	59	53	54	51	56	67
Comm. v. Capitalism	68	61	63	59	67	56	51	61
Death penalty	79	58	61	61	61	61	55	57
Evolution	72	59	54	58	64	58	61	61
Existence of God	73	58	54	59	57	54	55	59
Gay Marriage	88	65	56	66	63	64	65	65
Healthcare	86	63	61	62	60	69	65	64
MJ Legalization	81	60	55	64	66	62	62	55
Weighted avg.	77	60	57	61	61	58	60	61
No context avg.	77	54	52	54	56	53	55	55

Three reasons sentiment didn't help

- ▶ **discourse context**: many texts are replies (esp. in contentious domains)
 - ▶ including previous post elevates results
- ▶ **local pragmatic context**: people reference others' positions
 - ▶ term usage isn't enough

Referencing via quotation

“Exactly, Evolution is no more science than Creation”

I honestly don't know if this post is a troll.

1) “Creation” is faith-based. A religious book says it is so, and followers of that religion then take it on faith. Collecting evidence from the world is not required. This has 0 to do with science.

2) “Evolution”, agree with it or not, is our current best guess at how species came into being based on evidence that has been collected. This embodies the definition of science - it is creating explanations for natural phenomenon through observation and critical thought.

If you disagree, please explain by what method you believe the theory of evolution has come to exist.

Referencing throughout a discourse

Even if you want a baby, it's unbelievably hard to go through and something your body never fully recovers from....You can't force somebody to go through this **life-threatening condition** that is always *physically harmful*.

“**Life threatening condition** that is always *physically harmful*”?
What a giant load of steamy BS. Rarely is pregnancy *physically harmful* and even rarer is it **life threatening**.

Pregnancy is ALWAYS **physically harmful**. You try carrying a load of extra weight about and see what that does to your heart.

So, tell me, if it is ALWAYS **life-threatening**, why are there so few **life-threatening conditions** mentioned above?

Three reasons sentiment didn't help

- ▶ **discourse context**: many texts are replies (esp. in contentious domains)
 - ▶ including previous post elevates results
- ▶ **local pragmatic context**: people reference others' positions
 - ▶ term usage isn't enough
 - ▶ attempts to model writer's **commitment** to a text span environment (conditionals, questions, repetitions, strongly veridical contexts, doubt contexts) have little effect

Three reasons sentiment didn't help

- ▶ **discourse context**: many texts are replies (esp. in contentious domains)
 - ▶ including previous post elevates results
- ▶ **local pragmatic context**: people reference others' positions
 - ▶ term usage isn't enough
 - ▶ attempts to model writer's **commitment** to a text span environment (conditionals, questions, repetitions, strongly veridical contexts, doubt contexts) have little effect
- ▶ **lack of term coverage**: even in highly invested domains, people resort to side-indicating “emotional” language rarely
 - ▶ unigrams beaten by POS-dependencies because have reasonable training data
 - ▶ 40% of posts have no reliable sentiment term!

Sentiment vs. event-level evaluativity

You don't have to walk them or bathe them because they're smart enough to figure out all that stuff on their own. Plus, they have the common courtesy to do their business in the litter box, instead of all over your house and yard. Just one of the many reasons cats rule and dogs, quite literally drool!

Say, you had a bad day at work, or a bad breakup, you just wanna go home and cry. A cat would just look at you like "oh ok, you're home" and then walk away. A dog? Let's see, the dog would most likely wiggle its tail, with tongue sticking out and head tilted - the "you're home! i missed you so much, let's go snuggle in front of the TV and eat ice-cream" look. What more do I need to say?

What we need, practically

- ▶ Reliable text span environment segmenter
- ▶ A good default procedure for guessing the sentiment type of neutral terms in context
- ▶ Training data
 - ▶ massive amounts of within-topic data
 - ▶ a better way to generalize words to relevant labels

What is needed, more generally

- ▶ **Clarity**: Is sentiment tendential or constant? Inferred or stated
- ▶ **Coverage**: How do we deal with the huge sentiment-less text subject to Clarity?
- ▶ **Context**: How do we incorporate
 - ▶ the type of conversation
 - ▶ the topic
 - ▶ the commitment environment
 - ▶ the local compositional system?