



# Annotation of Subjective Language

Nianwen Xue

7/29/2011

LSA Summer Institute



## What is subjectivity?

- The linguistic expression of somebody's **opinions, sentiments, emotions, evaluations, beliefs, speculations** (*private states*)
- **Private state:** state that is not open to objective observation or verification  
*Quirk, Greenbaum, Leech, Svartvik (1985). A Comprehensive Grammar of the English Language.*
- Subjectivity analysis classifies content in **objective** or **subjective**

## Question Answering

Q: What is the international reaction to the reelection of Robert Mugabe as President of Zimbabwe?

African observers **generally approved** of his victory while Western Governments **denounced** it.



## Application: Information Extraction

“The Parliament **exploded** into fury against the government when word leaked out...”

**Observation:** subjectivity often causes false hits for IE

Goal: augment the results of IE



## More Applications

- **Product review mining:** What features of the ThinkPad T43 do customers like and which do they dislike?
- **Review classification:** Is a review positive or negative toward the movie?
- **Tracking sentiments toward topics over time:** Is anger ratcheting up or cooling down?
- **Prediction (election outcomes, market trends):** Will Clinton or Obama win?
- **Expressive text-to-speech synthesis**
- **Text semantic analysis**
- **Text summarization**



## Corpus

- [MPQA: www.cs.pitt.edu/mqpa/databaserelease](http://www.cs.pitt.edu/mqpa/databaserelease) (version 2)
- English language versions of articles from the world press (*187 news sources → 425 documents*)
  - 15,991 subjective expressions in 8,984 sentences
- Themes of the instructions:
  - No rules about how *particular words* should be annotated.
  - Don't take expressions out of context and think about what they *could* mean, *but judge them as they are used in that sentence.*



## Subjectivity annotation in MPQA

- Subjectivity refers to the linguistic expression of *private states*
- Direct subjective frames
  - The US *fears* a spill-over of the anti-terrorist campaign
  - Italian senator Renzo Gubert *praised* the Chinese government's efforts
  - “The report is full of absurdities,” he *said*.
- Expressive subjective frames
  - “The report is *full of absurdities*,” he said.
- All expressive subject frames are potential sentiment expressions, but not all direct subjective frames



## Private state

- A private state is a state not open to observation or verification
- Private states can be viewed as states of “experiencers” holding “attitudes” optionally towards some “targets”





## Three types of private state expressions

- Explicit mentions of private states
- Speech events expressing private states
- Expressive subjective elements
  - Expressive subjective elements are used by people to express their frustration, anger, wonder, positive sentiment, mirth, etc., without explicitly stating that they are frustrated, angry, etc.



## Three types of private state expressions

- Explicit mentions of private states
  - “The U.S. **fears** a spill-over,” said Xirao-Nima.
- Speech events expressing private states
  - “The report is full of absurdities,” Xirao-Nima **said**.
- Expressive subjective elements
  - “The report is **full of absurdities**,” Xirao-Nima said.

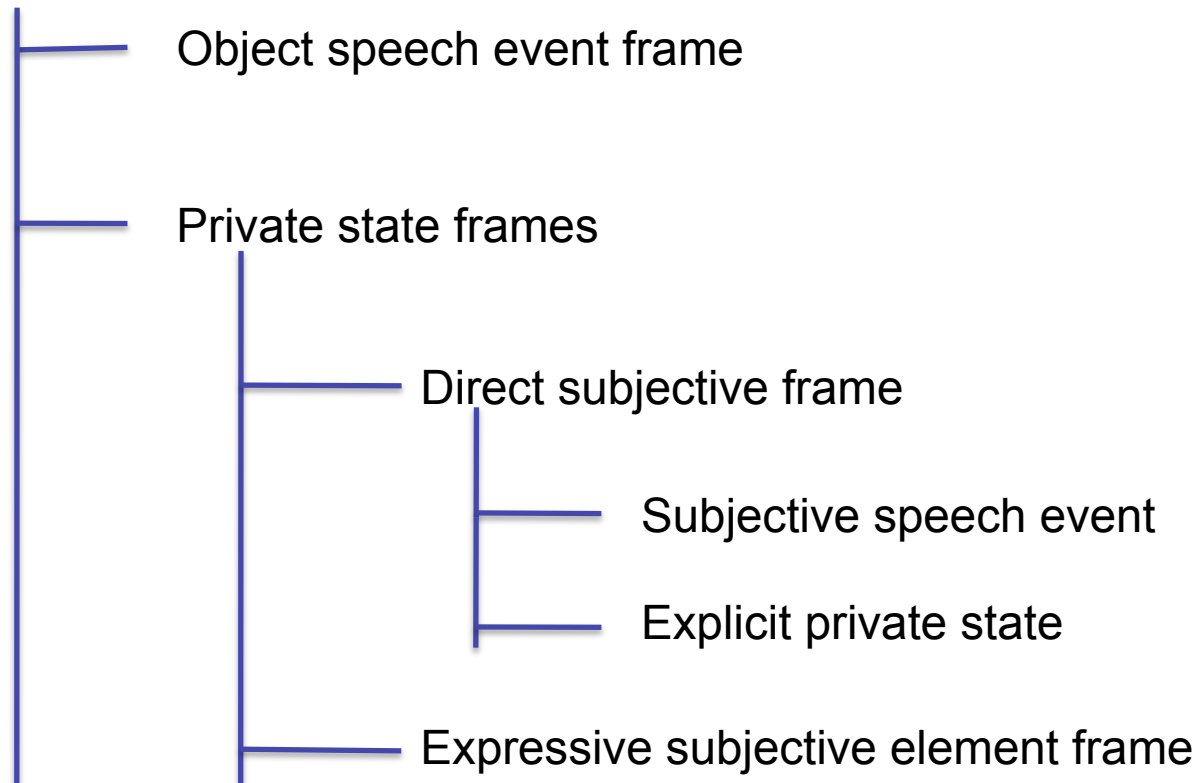


## Two types of private state frames

- Direct subject frames
  - Subjective speech events
  - Explicitly mentioned private states
  - Private state actions also folded into this frame:
    - As the long line of would-be voters marched in, those near the front of the queue began to spontaneously **applaud** those who were far behind them.
- Expressive subject element frames



## Complete annotation scheme





## What's in a **direct subjective frame**?

- Text anchor
- Source: the person or entity that is expressing the private state, possibly the writer.
- Target: the target or topic of the private state, i.e., what the speech event or private state is about
- Properties
  - Intensity (low, medium, high, extreme)
  - Expression intensity (low, medium, **neutral**, high, extreme)
  - Insubstantial (true, false)
  - Attitude type (positive, negative, other, none)



## What's in a **expressive subjective element frame**?

- Text anchor
- Source: the person or entity that is expressing the private state, possibly the writer.
- Target: the target or topic of the private state, i.e., what the speech event or private state is about
- Properties
  - Intensity (low, medium, high, extreme)
  - Expression intensity (low, medium, neutral, high, extreme)
  - Insubstantial (true, false)
  - Attitude type (positive, negative, other, none)



## What's in a **objective speech event frame**?

- Text anchor
- Source: the speaker or writer
- Target: the target or topic of the speech event, i.e., the content of what is said..
- Properties
  - Intensity (low, medium, high, extreme)
  - Expression intensity (low, medium, neutral, high, extreme)
  - Insubstantial (true, false)
  - Attitude type (positive, negative, other, none)



## Objective speech event frame: an example

- Sargeant O'Leary **said** the incident took place at 2:00pm.
- Text anchor?
- Source?
- Target?



## Agent frames as source

(1) “I have been to Tibet many times. I have seen the truth there, which is very different from what some US politicians with ulterior motives have described,” said [Xirao-Nima](#), who is a Tibetan. [“US Human Rights Report Defies Truth,” 2002-02-11, By Xiao Xin, Beijing China Daily, Beijing, China]

(2) Some Westerners who have been there have also seen the ever-improving human rights in the Tibet Autonomous Region, [he](#) added. [“US Human Rights Report Defies Truth,” 2002-02-11, By Xiao Xin, Beijing China Daily, Beijing, China]

Agent:

Text anchor: Xirao-Nima in (1)

Source: nima

Agent:

Text anchor: he in (2)

Source: nima

**Coreference!**



## Nested source

The Foreign Ministry said Thursday that it was “surprised, to put it mildly” by the U.S. State Department’s [criticism](#) of Russia’s human rights record and objected in particular to the “odious” section on Chechnya. [“Ministry Criticizes ‘Odious’ U.S. Report,” 2002-03-08, Moscow Times, Moscow, Russia]

Source: writer, Foreign Ministry, Foreign Ministry, U.S. State Department



## Nested source

“The U.S. **fears** a spill-over,” said Xirao-Nima.

Source (writer, Xirao-Nima, U.S.)

Source for “said”?

Source for the entire sentence?



## Putting things together

“The US **fears** a spill-over,” **said** Xirao-Nima, a professor of foreign affairs at the Central University for Nationalities.

Objective speech event:

Text anchor: **the entire sentence**

Source: <writer>

Implicit: true

Objective speech event:

Text anchor: **said**

Source: <writer,Xirao-Nima>

Direct subjective:

Text anchor: **fears**

Source: <writer,Xirao-Nima,U.S.>

Intensity: medium

Expression intensity: medium

Attitude type: negative



## Try it out:

“The report is full of absurdities,” Xirao-Nima said.

Objective speech event:

Text anchor: the entire sentence

Source: ???

Implicit: ???

Direct subjective:

Text anchor: said

Source: ???

Intensity: ??

Expression intensity: ??

Target: ??

Attitude type: ??

Expressive subjective element:

Text anchor: full of absurdities

Source: ??

Intensity: ??

Attitude type: ??



## Try it out:

“The report is full of absurdities,” Xirao-Nima said.

Objective speech event:

Text anchor: the entire sentence

Source: writer

Implicit: true

Direct subjective:

Text anchor: said

Source: <writer,Xirao-Nima>

Intensity: high

Expression intensity: neutral

Target: report

Attitude type: negative

Expressive subjective element:

Text anchor: full of absurdities

Source: <writer, Xirao-Nima>

Intensity: high

Attitude type: negative



## Subjective vs objective speech events

- Subjective speech event
  - “The report is full of absurdities,” Xirao-Nima **said**.
- Objective speech event
  - “The U.S. fears a spill-over,” **said** Xirao-Nima.



## (Explicit) Subjective vs objective speech events

- “We foresaw electoral fraud but not daylight robbery,”  
Tsvangirai **said**.

**subjective**

- Medical Department head Dr Hamid Saeed **said** the  
patient’s blood had been sent to the Institute for Virology  
in Johannesburg for analysis.

**objective**





## (Implicit) Subjective vs objective speech events

- The report is flawed and inaccurate.

**subjective**

- Bell Industries Inc. increased its quarterly to 10 cents from 7 cents a share.

**objective**

Critical test:

Given a speech event  $S$  with nested source  $X_1, X_2, X_3$ , if  $S$  reveals the private state of  $X_3$ , then  $S$  is subjective. Otherwise it is objective.



## Context dependency of subjectivity

- Smoking causes cancer.

Subjective? Objective?



## Mixture of private states and speech

- berate, object, praise, and criticize
- Motivated the merging of subjective speech events and explicit private states into one single direct subjective frame
- Motivated the inclusion of expression intensity



## Levels

“The United States was slandering China again,” said Xirao-Nima, a professor of Tibetan history at the Central University for Nationalities.

Expressive subjective element:

Text anchor: slandering

Source: <writer,xirao-nima>

Intensity: high

Attitude type: negative

Direct subjective:

Text anchor: slandering

Source: <writer, xirao-nima,US>

Target: China

Intensity: high

Expression intensity: high



## Ambiguity

- achieved (2 subjective, 4 objective);
- against (15 subjective, 40 objective);
- considering (3 subjective, 7 objective);
- difficult (7 subjective, 8 objective);
- fact (14 subjective, 7 objective);
- necessary (2 subjective, 2 objective);
- pressure (4 subjective, 4 objective);
- thousands (2 subjective, 5 objective);
- victory (3 subjective, 9 objective);
- world (13 subjective, 51 objective).



## Try it yourself

Let us not put too fine a point on it: Thursday's House vote on Speaker John Boehner's debt ceiling proposal is a joke. If it passes the House, Harry Reid has said it is dead on arrival in the Senate. If it somehow passes the Senate, which it won't, President Obama will veto it. It is, therefore, a symbolic act that is wasting precious time. It follows last week's Republican theatrics, the passage of the Cut and Demolish Act (or whatever they called it), which also was a waste of time. These are the actions of a party that has completely lost track of reality—and of a leader, John Boehner, who has lost the support of his party.



## Try it yourself

Let us not put too fine a point on it: Thursday's House vote on Speaker John Boehner's debt ceiling proposal **is a joke**. If it passes the House, Harry Reid has **said** it is **dead on arrival** in the Senate. If it somehow passes the Senate, which it won't, President Obama will veto it. It is, therefore, **a symbolic act that is wasting precious time**. It follows last week's **Republican theatrics**, the passage of the **Cut and Demolish Act** (or whatever they called it), which also **was a waste of time**. These are the actions of a party that has **completely lost track of reality**—and of a leader, John Boehner, who has **lost the support of his party**.



## What is sentiment analysis?

- Also known as opinion mining
- Attempts to identify the opinion/sentiment that a person may hold towards an object
- It is a finer-grained analysis compared to subjectivity analysis

Sentiment Analysis	Subjectivity analysis
Positive	Subjective
Negative	
Neutral	Objective



## Components of an opinion

- Basic components of an **opinion**:
  - **Opinion holder**: The person or organization that holds a specific opinion on a particular object.
  - **Object**: on which an opinion is expressed
  - **Opinion**: a view, attitude, or appraisal on an object from an opinion holder.





## Opinion mining tasks

- At the **document (or review) level**:
  - Task: sentiment classification of reviews
  - Classes: positive, negative, and neutral
  - Assumption: each document (or review) focuses on a single object (not true in many discussion posts) and contains opinion from a single opinion holder.
- At the **sentence level**:
  - Task 1: identifying subjective/opinionated sentences
    - Classes: objective and subjective (opinionated)
  - Task 2: sentiment classification of sentences
    - Classes: positive, negative and neutral.
    - Assumption: a sentence contains only one opinion; not true in many cases.
    - Then we can also consider clauses or phrases.



## Opinion Mining Tasks (cont.)

- At the **feature level**:
  - *Task 1: Identify and extract object features that have been commented on by an opinion holder (e.g., a reviewer).*
  - *Task 2: Determine whether the opinions on the features are positive, negative or neutral.*
  - *Task 3: Group feature synonyms.*
    - Produce a feature-based opinion summary of multiple reviews.
- **Opinion holders**: identify holders is also useful, e.g., in news articles, etc, but they are usually known in the user generated content, i.e., authors of the posts.



## Facts and Opinions

- Two main types of textual information on the Web.
  - Facts and Opinions
- Current search engines search for facts (assume they are true)
  - Facts can be expressed with topic keywords.
- Search engines do not search for opinions
  - Opinions are hard to express with a few keywords
    - How do people think of Motorola Cell phones?
  - Current search ranking strategy is not appropriate for opinion retrieval/search.



## Applications

- **Businesses and organizations:**
  - product and service benchmarking.
  - market intelligence.
  - Business spends a huge amount of money to find consumer sentiments and opinions.
    - Consultants, surveys and focused groups, etc
- **Individuals:** interested in other's opinions when
  - purchasing a product or using a service,
  - finding opinions on political topics
- **Ads placements:** Placing ads in the user-generated content
  - Place an ad when one praises a product.
  - Place an ad from a competitor if one criticizes a product.
- **Opinion retrieval/search:** providing general search for opinions.



## Two types of evaluations

- Direct Opinions: sentiment expressions on some objects, e.g., products, events, topics, persons.
  - E.g., “the picture quality of this camera is great”
  - Subjective
- Comparisons: relations expressing similarities or differences of more than one object. Usually expressing an ordering.
  - E.g., “car x is cheaper than car y.”
  - Objective or subjective.



## Sentiment annotation based on the MPQA Corpus

- Sentiment annotation based on the MPQA Corpus (Wilson et al 2008)
- A sentiment as a positive or negative emotion, evaluation, or stance.
- Sentiment expressions are a subset of subjectivity expressions
  - Expressive subjective elements
  - Direct subjective annotations with a non-neutral expression intensity



## Examples

	Positive sentiments	Negative sentiments
Emotion	I'm happy	I'm sad
Evaluation	Great idea!	Bad idea!
Stance	She supports the bill	She's against the bill





## Annotation scheme

- Positive: marks positive sentiment
- Negative: marks negative sentiment
- Both: expressions in which both positive and negative sentiments are expressed
- Neutral: expressions that are neither positive or negative.

## Annotation Scheme

- A sentiment is a positive or negative emotion, evaluation, or stance
- Mark polarity of subjective expressions as **positive**, **negative**, **both**, or **neutral**

African observers **generally approved** of his victory while Western governments **denounced** it.

**positive**

**negative**

Besides, politicians refer to **good and evil**...

**both**

Jerome says the hospital **feels** no different than a hospital in the states.

**neutral**



## Example annotation

- Thousands of coup supporters celebrated (**positive**) overnight, waving flags, blowing whistles . . .
- The criteria set by Rice are the following: the three countries in question are repressive (**negative**) and grave human rights violators (**negative**) . . .
- Besides, politicians refer to good and evil (**both**) only for purposes of intimidation and exaggeration.
- Jerome says the hospital feels (**neutral**) no different than a hospital in the States.



## Context dependence

- Judge the contextual polarity of sentiment *ultimately* being conveyed

They **have not succeeded, and will never succeed**, in breaking the will of this valiant people.

**positive**



“Annotators are asked to judge the contextual polarity of the sentiment that is ultimately being conveyed by the subjective expression, that is, once the sentence has been fully interpreted.”

## Prior Polarity versus Contextual Polarity

- From a lexicon of positive and negative words

**Prior polarity:** out of context, positive or negative

*beautiful* → positive

*horrid* → negative

- A word may appear in a phrase that expresses a different polarity in context

**Contextual polarity**

“Cheers to Timothy Whitfield for the **wonderfully horrid** visuals.”



## Example

Philip Clap, President of the National Environment Trust, sums up well the general thrust of the reaction of environmental movements: there is no reason at all to believe that the polluters are suddenly going to become reasonable.



## Another example

Philip Clap, President of the National Environment **Trust**, sums up **well** the general thrust of the reaction of environmental movements: there is no **reason** at all to believe that the **polluters** are suddenly going to become **reasonable**.

Example

Philip Clap, President of the National Environment ~~Trust~~, sums up well the general thrust of the reaction of environmental movements: there is no reason at all to believe that the ~~polluters~~ are suddenly going to become reasonable.

prior polarity

Contextual polarity







## Polarity influencers

- Negation
  - Not good, no one thinks that is good, not only good but amazing
- Modality
  - No reason at all to believe
- Word sense
  - Environmental Trust, He has won the people's trust
- Syntactic function
  - Polluters are vs they are polluters
- Diminishers
  - Little truth



## Polarity influencers

- Domain/topic
  - Cool car/cool demeanor
  - High fever
- Perspective
  - Israel failed to defeat Hezbollah

## Annotation procedure

**Need:** sentiment expressions with contextual polarity

→ positive and negative expressions of emotions, evaluations, stances

**Had:** subjective expression annotations in MPQA Opinion Corpus

<http://nrrc.mitre.org/NRRC/publications.htm>

→ words/phrases expressing emotions, evaluations, stances, speculations, etc.

sentiment expressions  $\subseteq$  subjective expressions

**Decision:** annotate subjective expressions in MPQA Corpus with their contextual polarity



## Agreement Study

- 10 documents with 447 subjective expressions
  - Kappa: 0.72 (82%)
- Remove uncertain cases → at least one annotator marked *uncertain* (18%)
  - Kappa: 0.84 (90%)

(But all data included in experiments)



## Try it yourself

Our nation's economy and international reputation as the world's presiding grownup has already been badly damaged. It is a self-inflicted wound of monumental stupidity. I am usually willing to acknowledge that Democrats can be as silly, and hidebound, as Republicans—but not this time. There is zero equivalence here. The vast majority of Democrats have been more than reasonable, more than willing to accept cuts in some of their most valued programs. Given the chance, there was the likelihood that they would have surrendered their most powerful weapon in next year's election—a Medicare campaign—by agreeing to some necessary long-term reforms in that program. The President, remarkably, proposed raising the age of eligibility for Medicare to 67.



## Try it yourself

Our nation's economy and international reputation as the world's presiding grownup has already been **badly damaged**. It is a self-inflicted wound **of monumental stupidity**. I am usually willing to acknowledge that Democrats can be **as silly**, and **hidebound**, as Republicans—but not this time. There is zero equivalence here. The vast majority of Democrats have been **more than reasonable**, **more than willing** to accept cuts in some of their most valued programs. Given the chance, there was the likelihood that they would have surrendered their most powerful weapon in next year's election—a Medicare campaign—by agreeing to some necessary long-term reforms in that program. The President, **remarkably**, proposed raising the age of eligibility for Medicare to 67.



## References

- Janyce Wiebe, Theresa Wilson, and Claire Cardie (2005). [Annotating expressions of opinions and emotions in language.](#) *Language Resources and Evaluation*, volume 39, issue 2-3, pp. 165-210.
- Theresa Wilson, Janyce Wiebe, and Paul Hoffmann (2009). [Recognizing Contextual Polarity: An exploration of features for phrase-level sentiment analysis.](#) *Computational Linguistics*, 35:3, pages 399-433.