Connotation:
A Dash of Sentiment
Beneath the Surface Meaning
Con•notation

“com-” (“together or with”) | “notare” (“to mark”)
Connotation

“com-” (“together or with”) | “notare” (“to mark”)

• Commonly understood cultural or emotional association that some word carries, in addition to its explicit or literal meaning (denotation).
• Generally described as positive or negative.
Good News? Bad News?

“Decrease in deforestation drives the trend, but emissions from energy and agriculture grow.”
“Decrease in deforestation drives the trend, but emissions from energy and agriculture grow.”

The production and discharge of something
“Decrease in deforestation drives the trend, but emissions from energy and agriculture grow.”
"Decrease in deforestation drives the trend, but emissions from energy and agriculture grow."

The production and discharge of something, esp. gas or radiation
Decrease in deforestation drives the trend, but emissions from energy and agriculture grow.

Negatively connotative in general
**Connotation**

“com-” (“together or with”) | “notare” (“to mark”)

- Commonly understood cultural or emotional *association* that some word carries, in addition to its explicit or literal meaning (denotation).

  Generally described as **positive** or **negative**.
Sentiment vs. Connotation
Sentiment vs. Connotation

joy

sick
Sentiment vs. Connotation

Sentiment

joy

sick

Neutral

surfing
scientist
grid
rose
header
blister
emission
salt
bedbug
...

...
Sentiment vs. Connotation

Sentiment
- Joy
- Sick

Connotation
- Scientist
- Music
- Surfing
- Rose
- Flu
- Deforestation
- Emission

Neutral
- Surfing
- Scientist
- Grid
- Rose
- Header
- Blister
- Emission
- Salt
- Bedbug
- ...
Sentiment vs. Connotation

joy
- scientist
- music
- surfing
- rose

sick
- deforestation
- emission
- flu
- bedbug

Connotation

Neutral
- surfing
- scientist
- grid
- rose
- header
- blister
- emission
- salt
- bedbug
- ...
Learning the General Connotation

- Data
- Linguistic insights
- Graph representation
- Inference algorithms
- Evaluations
Data

- **Web-driven data**
  - Google Web IT (Brants and Franz (2006))
    - N-grams (1 <= n <= 5)
    - Frequency of occurrences
    - Example: “prevent financial malware 4130”

- **Dictionary-drive data**
  - WordNet (George A. Miller (1995))
    - Synsets: synonyms, antonyms
Learning the General Connotation

- Data
- Linguistic insights
- Graph representation
- Inference algorithms
- Evaluations
Diverse Linguistic Insights

- Semantic prosody
- Semantic parallelism of coordination
- Distributional similarity
- Semantic relations
Diverse Linguistic Insights

Semantic prosody

- Semantic parallelism of coordination
- Distributional similarity
- Semantic relations
## Semantic Prosody


<table>
<thead>
<tr>
<th><strong>enjoy music</strong></th>
<th><strong>enjoy blisters</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Web</strong></td>
<td><strong>Web</strong></td>
</tr>
<tr>
<td>Images</td>
<td>Images</td>
</tr>
<tr>
<td>Maps</td>
<td>Maps</td>
</tr>
<tr>
<td>Shopping</td>
<td>Shopping</td>
</tr>
<tr>
<td>About <strong>930,000,000</strong> results (0.28 seconds)</td>
<td>About <strong>2,650,000</strong> results (0.36 seconds)</td>
</tr>
</tbody>
</table>
Selectional Preference on Connotation

\( \text{enjoy}^{\text{Pred}} \rightarrow [\text{enjoy}]^{\text{Pred}} [\text{music}]^{\text{Arg}} \)

\( \text{prevent}^{\text{Pred}} \rightarrow [\text{prevent}]^{\text{Pred}} [\text{bedbugs}]^{\text{Arg}} \)
Selectional Preference on Connotation

enjoy

\rightarrow

\text{[enjoy]} \quad \text{[music]}

prevent

\rightarrow

\text{[prevent]} \quad \text{[bedbugs]}

Candidates
Connotative Predicate: A predicate that has selectional preference on the connotative polarity of some of its semantic arguments.
**Connotative Predicate:**

A predicate that has *selectional preference* on the connotative polarity of some of its semantic arguments.

<table>
<thead>
<tr>
<th>Connotative Predicates</th>
<th>Sentiment of predicate</th>
<th>Preference on arguments</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>suffer</td>
<td>negative</td>
<td>negative</td>
<td>“suffering from cough”</td>
</tr>
<tr>
<td>cure</td>
<td>positive</td>
<td>negative</td>
<td>“cure backache”</td>
</tr>
<tr>
<td>cause</td>
<td>neutral</td>
<td>negative</td>
<td>“caused CO₂ emissions”</td>
</tr>
</tbody>
</table>

Feng et al. 2011
### Connotative Predicate:

A predicate that has *selectional preference* on the connotative polarity of some of its semantic arguments.

<table>
<thead>
<tr>
<th>Connotative Predicates</th>
<th>Sentiment of predicate</th>
<th>Preference on arguments</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>suffer</td>
<td>negative</td>
<td>negative</td>
<td>“suffering from cough”</td>
</tr>
<tr>
<td>cure</td>
<td>positive</td>
<td>negative</td>
<td>“cure backache”</td>
</tr>
<tr>
<td>cause</td>
<td>neutral</td>
<td>negative</td>
<td>“caused CO₂ emissions”</td>
</tr>
</tbody>
</table>

Feng et al. 2011
### 20 Positive Connotative Predicates

- Accomplish
- Achieve
- Advance
- Advocate
- Admire
- Applaud
- Appreciate
- Compliment
- Congratulate
- Develop
- Desire
- Enhance
- Enjoy
- Improve
- Praise
- Promote
- Respect
- Save
- Support
- Win

### 20 Negative Connotative Predicates

- Alleviate
- Accuse
- Avert
- Avoid
- Cause
- Complain
- Condemn
- Criticize
- Detect
- Eliminate
- Eradicate
- Mitigate
- Overcome
- Prevent
- Prohibit
- Protest
- Refrain
- Suffer
- Tolerate
- Withstand

Feng et al. 2011
Diverse Linguistic Insights

- **Semantic prosody** [Corpus: GoogleNgram]
  - \([\text{enjoy}]^{\text{Pred}} [\text{music}]^{\text{Arg}}, [\text{prevent}]^{\text{Pred}} [\text{begbugs}]^{\text{Arg}}\)
Diverse Linguistic Insights

- **Semantic prosody** [Corpus: GoogleNgram]
  - “enjoy *”; “prevent *”

- **Semantic parallelism of coordination** [Corpus: GoogleNgram]
  - Pattern “* and *”, e.g., “music and wine”
Diverse Linguistic Insights

- **Semantic prosody** [Corpus: GoogleNgram]
  - “enjoy *”;“prevent * ”

- **Semantic parallelism of coordination** [Corpus: GoogleNgram]
  - Pattern “ * and * ”, e.g., “music and wine”

- **Distributional similarity** [Corpus: GoogleNgram]
  - “findings”–“potentials” > “findings” –“modifications”
Diverse Linguistic Insights

- **Semantic prosody** [Corpus: GoogleNgram]
  - “enjoy * ”;“prevent * ”
- **Semantic parallelism of coordination** [Corpus: GoogleNgram]
  - Pattern “ * and * ”, e.g., “music and wine”
- **Distributional similarity** [Corpus : GoogleNgram]
  - “findings”–“potentials” > “findings”–“modifications”
- **Semantic relations** [Corpus: WordNet]
  - Synonyms
  - Antonyms
Learning the General Connotation

- Data
- Linguistic insights
- Graph representation
- Inference algorithms
- Evaluations
Graph Representation

- **Graph** $G = (V, E)$
  - $V = \{\text{Pred}\} \cup \{\text{Arg}\}$
  - $E_1: \text{Pred} \rightarrow \text{Arg}$

$$PMI(p, a) = \log_2 \frac{P(p, a)}{P(p)P(a)}$$
Graph Representation

- **Graph** $G = (V, E)$
  - $V = \{\text{Pred}\} \cup \{\text{Arg}\}$
  - $E_1$: Pred – Arg
    
    $$PMI(p, a) = \log_2 \frac{P(p, a)}{P(p)P(a)}$$
  - $E_2$: Arg – Arg
    
    $$CosineSim(\vec{a_1}, \vec{a_2}) = \frac{\vec{a_1} \cdot \vec{a_2}}{||\vec{a_1}|| \cdot ||\vec{a_2}||}$$

![Diagram of a graph representation with nodes and edges connecting different words such as enjoy, thank, avoid, prevent, profit, investment, reading, aid, writing, etc.](image-url)
Graph Representation

- **Graph G = (V, E)**
  - V = \{Pred\} U \{Arg\}
  - E1: Pred – Arg

\[ PMI(p, a) = \log_2 \frac{P(p, a)}{P(p)P(a)} \]

- E2: Arg – Arg

\[ CosineSim(a_1, a_2) = \frac{a_1 \cdot a_2}{||a_1|| \cdot ||a_2||} \]

- “a_1 and w_1” $\Rightarrow$ PMI(a_1, w_1)
- “a_1 and w_2” $\Rightarrow$ PMI(a_1, w_2)
- …
- “a_1 and w_3” $\Rightarrow$ PMI(a_1, w_n)

\[ a_1 := \begin{bmatrix} PMI(a_1, w_1) \\ PMI(a_1, w_2) \\ \vdots \\ PMI(a_1, w_n) \end{bmatrix} \]
Graph Representation

- prevent
- suffer
- enjoy
- thank

- tax
- loss
- writing
- profit
- preventing
- flu
- bonus
- investment
- gain
- cold
- flu

- prosody
- synonyms
- coordination
- antonyms
Learning the General Connotation

- Data
- Linguistic insights
- Graph representation
- Inference algorithms
- Evaluations
Inference Algorithm

Algorithms
- HITS/PageRank
- Label Propagation
- Integer Linear Programming
- Belief Propagation

Linguistic Insights
- Semantic Prosody
- Semantic Parallelism of Coordination
- Distributional Similarity
- Semantic Relations
For each unlabeled word $i$, solve $x_i, y_i, z_i \in \{0, 1\}$,

- $x_i \rightarrow$ positive,
- $y_i \rightarrow$ negative,
- $z_i \rightarrow$ neutral;

$x_i + y_i + z_i = 1$. 

ILP: Problem Formulation
ILP: Problem Formulation

- For each unlabeled word \( i \), solve \( x_i, y_i, z_i \in \{0, 1\}, \)
  - \( x_i \rightarrow \text{positive}, \)
  - \( y_i \rightarrow \text{negative}, \)
  - \( z_i \rightarrow \text{neutral}; \)
  - \( x_i + y_i + z_i = 1. \)

- **Initialization (Hard constraints)**
  - Positive seed predicates (e.g. "achieve") \( \rightarrow x_i = 1 \)
  - Negative seed predicates (e.g. "prevent") \( \rightarrow y_i = 1 \)
Maximize

\[ F = \Phi^{\text{prosody}} + \Phi^{\text{coord}} + \Phi^{\text{neu}} \]
ILP: Objective Function

Maximize

\[ F = \Phi_{\text{prosody}} + \Phi_{\text{coord}} + \Phi_{\text{neu}} \]
ILP: Objective Function

Maximize

$$F = \Phi_{\text{prosody}} + \Phi_{\text{coord}} + \Phi_{\text{neu}}$$
ILP: Objective Function

\[ F = \Phi_{prosody} + \Phi_{coord} + \Phi_{neu} \]

\[ \Phi_{prosody} = \sum_{i,j}^{\mathcal{R}_{pred}} w_{i,j}^{pred} (d_{i,j}^{++} + d_{i,j}^{--} - d_{i,j}^{+-} - d_{i,j}^{-+}) \]

\[ \Phi_{coord} = \sum_{i,j}^{\mathcal{R}_{coord}} w_{i,j}^{coord} (d_{i,j}^{++} + d_{i,j}^{--} + d_{i,j}^{00}) \]

\[ \Phi_{neu} = \alpha \sum_{i,j}^{\mathcal{R}_{pred}} w_{i,j}^{pred} \cdot z_j \]
ILP: Objective Function

\[ F = \Phi_{\text{prosody}} + \Phi_{\text{coord}} + \Phi_{\text{neu}} \]

\[
\Phi_{\text{prosody}} = \sum_{i,j} \mathcal{R}^{\text{pred}}_{i,j} \cdot w_{i,j}^{\text{pred}} \cdot (d_{i,j}^{++} + d_{i,j}^{--} - d_{i,j}^{+-} - d_{i,j}^{-+})
\]

\[
\Phi_{\text{coord}} = \sum_{i,j} \mathcal{R}^{\text{coord}}_{i,j} \cdot w_{i,j}^{\text{coord}} \cdot (d_{i,j}^{++} + d_{i,j}^{--} + d_{i,j}^{00})
\]

\[
\Phi_{\text{neu}} = \alpha \sum_{i,j} \mathcal{R}^{\text{pred}}_{i,j} \cdot w_{i,j}^{\text{pred}} \cdot z_{i,j}
\]
ILP: Soft Constraints

 Predicate – Argument

\[ w^{\text{pred}}(p, a) = \frac{\text{freq}(p, a)}{\sum_{(p, x) \in \mathcal{R}^{\text{pred}}} \text{freq}(p, x)} \]

 Argument – Argument

\[ w^{\text{coord}}(a_1, a_2) = \cos\text{Sim}({a_1}, {a_2}) = \frac{{a_1} \cdot {a_2}}{||{a_1}|| \cdot ||{a_2}||} \]

\[ \Phi^{\text{prosody}} = \sum_{i,j} w_{i,j}^{\text{pred}} (d_{i,j}^{++} + d_{i,j}^{-} - d_{i,j}^{+-} - d_{i,j}^{--}) \]

\[ \Phi^{\text{coord}} = \sum_{i,j} w_{i,j}^{\text{coord}} (d_{i,j}^{++} + d_{i,j}^{-} + d_{i,j}^{00}) \]
ILP: Hard Constraints

- **Semantic relations**
  - Antonym pairs will not have the same positive or negative polarity.
    \[
    \forall (i, j) \in R^{ant}, \quad x_i + x_j \leq 1, \quad y_i + y_j \leq 1
    \]
  - Synonym pairs will not have the opposite polarity.
    \[
    \forall (i, j) \in R^{syn}, \quad x_i + y_j \leq 1, \quad x_j + y_i \leq 1
    \]
Inference Algorithm

Algorithms

- HITS/PageRank
- Label Propagation
- Integer Linear Programming
- Belief Propagation

Linguistic Insights

- Semantic Prosody
- Semantic Parallelism of Coordination
- Distributional Similarity
- Semantic Relations
Graph

- Lemmas (115K)
- Synsets (63K)

Types of Edges
1. Predicate-argument (179K)
2. Argument-argument (144K)
3. Argument-synset (126K)
4. Synset-synset (34K)
Problem Formulation

- $G^{\text{Lemma + Sense}} = (V, E)$

- Nodes (random variables)
  - $V = \{v_1, v_2, \ldots, v_n\}$
  - Unobserved variables: $\Upsilon = \{Y_1, Y_2, \ldots, Y_n\}, y_i$

- Typed edges
  - $E = \{e(v_i, v_j, v_k) \mid v_i, v_j \in V, \ T_k \in T\}$
  - $T = \{\text{pre-arg}, \text{arg-arg}, \text{arg-syn}, \text{syn-syn}\}$

- Neighborhood function
  - $N_v = \{u \mid e(u, v) \in E\}$

- Labels
  - $L = \{+,-\}, y_i$ denotes the label of $Y_i$. 
Pairwise Markov Random Fields
Objective Function

\[ P(y \mid x) = \frac{1}{Z(x)} \prod_{y_i \in Y} \psi_i(y_i) \prod \psi_{ij}^t(y_i, y_j) \]
Objective Function

An assignment to all the unobserved variables

\[ P(y \mid x) = \frac{1}{Z(x)} \prod_{y_i \in Y} \psi_i(y_i) \prod \psi_{ij}^t(y_i, y_j) \]
Objective Function

\[ P(y \mid x) = \frac{1}{Z(x)} \prod_{y_i \in Y} \psi_i(y_i) \prod_{y_i, y_j} \psi^t_{ij}(y_i, y_j) \]

An assignment to all the unobserved variables

Variables with known labels
Objective Function

\[ P(y | x) = \frac{1}{Z(x)} \prod_{y_i \in Y} \psi_i(y_i) \prod \psi_{ij}^t(y_i, y_j) \]

Prior mapping:
y_i refers to \( Y_i \)'s label
\( \psi_i \) is prior mapping.
Objective Function

\[ P(y | x) = \frac{1}{Z(x)} \prod_{y_i \in Y} \psi_i(y_i) \prod_{y_i, y_j} \psi_{ij}^t(y_i, y_j) \]

Prior mapping: \( L \rightarrow \mathbb{R}_{\geq 0} \)
Compatibility mapping: \( L \times L \rightarrow \mathbb{R}_{\geq 0} \)
Objective Function

An assignment to all the unobserved variables
Loopy Belief Propagation

- Message passing

\[
m_{i \rightarrow j}(y_i) = \alpha \sum_{y_i \in L} (\psi_{ij}(y_i, y_j)\psi_i(y_i) \prod_{Y_k \in N_i \cap Y \setminus Y_i} m_{k \rightarrow i}(y_i)), \forall y_j \in L
\]

- Belief

\[
b_i(y_i) = \beta \psi_i(y_i) \prod_{Y_j \in N_i \cap Y} m_{j \rightarrow i}(y_i), \forall y_i \in L
\]
Loopy Belief Propagation

- Initialize “message” between all node pairs connected by an edge.
- Initialize priors for all nodes.
- Integrative message passing until all messages stop changing

\[
m_{i\rightarrow j}(y_i) = \alpha \sum_{y_i \in L} (\psi_{ij}^t(y_i, y_j)\psi_i(y_i) \prod_{Y_k \in Ni \cap \Upsilon \setminus Y_i} m_{k\rightarrow i}(y_i)), \forall y_j \in L
\]

- Compute beliefs.

\[
b_i(y_i) = \beta \psi_i(y_i) \prod_{Y_j \in Ni \cap \Upsilon} m_{j\rightarrow i}(y_i), \forall y_i \in L
\]

- Assign the label

\[
L_i \leftarrow \max_i b_i(y_i)
\]