Language, Mind, and Vision
- Learning to Read Deception
- Learning to Describe the Visual World

Yejin Choi
Computer Science & Engineering

UNIVERSITY of WASHINGTON
Natural Language Processing (NLP) - a quick overview
What is NLP?

Fundamental goal: deep understand of human language
– Not just string processing or keyword matching!
What is NLP?

- Simple: spelling correction, text categorization...
- Complex: speech recognition, machine translation, dialog interfaces, question answering...
- Unknown: human-level comprehension (is this just NLP?)
Semantic Ambiguity

At last, a computer that understands you like your mother.

• Direct Meanings:
  – It understands you like your mother (does) [presumably well]
  – It understands (that) you like your mother
  – It understands you like (it understands) your mother

• But there are other possibilities, e.g. mother could mean:
  – a woman who has given birth to a child
  – a stringy slimy substance consisting of yeast cells and bacteria; is added to cider or wine to produce vinegar

• Context matters, e.g. what if previous sentence was:
  – Wow, Amazon predicted that you would need to order a big batch of new vinegar brewing ingredients. 😊

[Example from L. Lee]
A phone that understands our questions

“What's the best movie to see this weekend”

That would probably start an argument. But here's a list of highly-regarded movies:

25 MOVIES

NORTH BY NORTHWEST
Released July 17, 1959
100%

THE TREASURE OF THE SIERRA
Released January 6, 1948
100%

What can I help you with?

“You need to start understanding me Siri”

I'll make a note of that.

“Yeah you better make a note of that”

Noted:

Of that
US Cities: Its largest airport is named for a World War II hero; its second largest, for a World War II battle.
Business

Latest News

- The exchange of financial stocks fell slightly prominent lower
12 stocks in Tokyo, ahead of sell orders from the backlash of higher yesterday, with slightly lower values. Nikkei ... ... (11:13) [Full article]

- Negotiation and integration of Japan Sompo Japan興亜 to aggregate in three large camps
Sompo Japan Insurance and it's five to start the negotiations for the merger of NIPPONKOA Insurance Co., Ltd. No. 12, 2007, minutes ... ... (10:33) [Full...
Natural Language Processing (NLP)
- a quick overview
Natural Language Processing (NLP)
- recent research (of our own)
films are if anyone wants to help dig under the snow for them.”

Soon a small party with a lantern dashed out into the howling darkness where Blackie’s memory suggested that a box of film had been left during the rush to get settled for the winter. Working like wild men to beat the cold, they dug a hole six feet deep into the snow and finally located the missing box.

The show, an old Charlie Chaplin release, was given right there in the mess hall where a stove and the kitchen filled half of one side of the room and bunks lined the other side. In the center was a long table and on either side of this were benches. Those who could not sit anywhere else stretched out on the upper bunks where they could drop things on the heads of those below.

What was said about the actors and actresses would have made them forget their cues could they but have heard. Comments were rough. If the members of the expedition didn’t like anyone on the screen they told him so in unmistakable terms of disapproval. Often they named the actors after some of those present, and yells of derision greeted their appearance on the screen. For instance, “Bill” Vander Veer, on account of his...
Three Different Layers of Reading

Information

Intent

Identity

Reading the author’s mind
films are if anyone wants to help dig under the snow for them."

Soon a small party with a lantern dashed out into the howling darkness where Blackie's memory suggested that a box of film had been left during the rush to get settled for the winter. Working like wild men to beat the cold, they dug a hole six feet deep into the snow and finally located the missing box.

The show, an old Charlie Chaplin release, was given right there in the mess hall where a stove and the kitchen filled half of one side of the room and bunks lined the other side. In the center was a long table and on either side of this were benches. Those who could, then, lay across the tables, while the upper bunk where they could try to get to the ends of some below.

What was said about the actors and actresses would have made them forget their cues could they but have heard Comments were rough. If the members of the expedition didn't like anyone on the screen they told him so in unmistakable terms of disapproval. Often they named the actors after some of those present, and yells of derision greeted their appearance on the screen. For instance, "Bill" Vander Veer, on account of his
films are if anyone wants to help dig under the snow for them."

Soon a small party with a lantern dashed out into the howling
darkness where Blackie’s memory suggested that a box of film
had been left during the rush to get settled for the winter. Work-
ing like wild men to beat the cold, they dug a hole six feet deep
into the snow and finally located the missing box.

The show, an old Charlie Chaplin release, was given right
there in the mess hall where a stove and the kitchen filled half
of one side of the room and bunks lined the other side. In the
center was a long table and on either side of this were benches.
Those who could not sit in the bunks stretched out on the
upper bunks where they could be seen only on the heads of those
below.

What was said about the actors and actresses would have
made them forget their cues could they but have heard. Com-
mments were rough. If the members of the expedition didn’t like
anyone on the screen they told him so in unmistakable terms of
disapproval. Often they named the actors after some of those
present, and yells of derision greeted their appearance on the
screen. For instance, “Bill” Vander Veer, on account of his

[14]
framing in media & political discourse
(Yano et al., 2010)
(Recasens et al., 2013)

dodging
(Nguyen et al. 2013)

hedging
(Choi et al. 2012)
(Ganter and Strube, 2009)
(Kilicoglu and Bergler 2008)

syntactic packaging
"My toy broke" instead of "I broke my toy"
(Greene and Resnik 2009)

dodging
(Nguyen et al. 2013)

hedging
(Choi et al. 2012)
(Ganter and Strube, 2009)
(Kilicoglu and Bergler 2008)

syntactic packaging
"My toy broke" instead of "I broke my toy"
(Greene and Resnik 2009)

defection
fake online reviews

"Eunsol" Choi

"My toy broke" instead of "I broke my toy"
(Greene and Resnik 2009)
films are if anyone wants to help dig under the snow for them.

Soon a small party with a lantern dashed out into the howling darkness where Blackie's memory suggested that a box of film had been left during the rush to get settled for the winter. Working like wild men to beat the cold, they dug a hole six feet deep into the snow and finally located the missing box.

The show, an old Charlie Chaplin release, was given right there in the mess hall where a stove and the kitchen filled half of one side of the room and bunks lined the other side. In the center was a long table, and on either side of this were benches. Those who could not sit in the benches trooped out on the upper bunks where they could look over the heads of those below.

What was said about the actors and actresses would have made them forget their cues could they but have heard. Comments were rough. If the members of the expedition didn't like anyone on the screen they told him so in unmistakable terms of disapproval. Often they named the actors after some of those present, and yells of derision greeted their appearance on the screen. For instance, "Bill" Vander Veer, on account of his

[14]
authorship verification

authorship obfuscation

demographics: gender, nationality, age, vocation

personality, psychological state: happy, authoritative, depressed...

intellectual traits & development: literary success
From Language to the Mind
From Language to the Mind

Is it even possible? (without full semantic understanding)

• It is more about “HOW” it is said than “WHAT” is said.

“HOW” it is said
i.e., Writing Style

Information
“WHAT”

Intent
“WHY”

Identity
“WHO”
From Language to the Mind

Is it even possible? (without full semantic understanding)

• It is more about “HOW” it is said than “WHAT” is said.
• We --humans– also often rely on “overall impression”.

Computers at times can do better than humans!

“How” it is said i.e., Writing Style

Information “WHAT”

Intent “WHY”

Identity “WHO”
“So how can you spot a fake review? Unfortunately, it’s difficult, but with some technology, there are a few warning signs:”

“To obtain a deeper understanding of the nature of deceptive reviews, we examined the relative utility of three potentially complementary framings of our problems.

“As online retailers increasingly depend on reviews as a sales tool, an industry of fibbers and promoters has sprung up to buy and sell raves for a pittance.”

What is “Writing Style”?


Research Paper (ACL, 2011)

The New York Times

Blog Post
What is “Writing Style”?

Genre Categorization:
- Petrenz and Webber, 2011; Finn et al., 2006; Argamon et al., 2003; Kessler et al., 1997

Authorship Attribution:
- Holmes 1985, Raghavan et al., 2010; Koppel and Shler, 2004; Gamon, 2004;

Many more possibilities...
- Swanson and Charniak, 2012; Xu et al., 2012; Iyyer et al., 2014; Hardisty et al., 2010

“HOW” it is said i.e., **Writing Style**

**Intent** “WHY”

**Identity** “WHO”

Alan Ritter
From Language to the Mind

Unconventional Case Studies:
I. Deceptive Reviews (ACL 2011)
II. Success of Novels (EMNLP 2013)

"HOW" it is said i.e., Writing Style

Intent "WHY"
Identity "WHO"
Motivation

Online reviews = shopping tool

Commercial impact → potential target for deceptive reviews
“My husband and I stayed at the James Chicago Hotel for our anniversary. This place is fantastic! We knew as soon as we arrived we made the right choice! The rooms are BEAUTIFUL and the staff very attentive and wonderful! The area of the hotel is great, since I love to shop I couldn’t ask for more! We will definitely be back to Chicago and we will for sure be back to the James Chicago.”

Deceptive or Truthful?
"My husband and I stayed at the James Chicago Hotel for our anniversary. This place is fantastic! We knew as soon as we arrived we made the right choice! The rooms are BEAUTIFUL and the staff very attentive and wonderful! The area of the hotel is great, since I love to shop I couldn’t ask for more! We will definitely be back to Chicago and we will for sure be back to the James Chicago."

"I have stayed at many hotels traveling for both business and pleasure and I can honestly say that The James is tops. The service at the hotel is first class. The rooms are modern and very comfortable. The location is perfect within walking distance to all of the great sights and restaurants. Highly recommend to both business travellers and couples."
Gathering Data

• Label existing reviews?
  – Can’t manually do this
Gathering Data

• Label existing reviews?
  – Can’t manually do this

☐ Instead, create new reviews
  – By hiring people to write fake positive reviews
  – Amazon Mechanical Turk
    • 20 hotels
    • 20 reviews / hotel
    • Offer $1 / review
    • 400 reviews
How good are humans in detecting deceptive reviews?

- 80 truthful and 80 deceptive reviews
- 3 undergraduate judges
Human Performance

→ Aligns with previous studies in deception literature: humans typically perform barely better than chance. trained experts may perform at ~70%

Accuracy

- Judge 1: 61.9
- Judge 2: 56.9 (performed at chance, p-value = 0.1)
- Judge 3: 53.1 (performed at chance, p-value = 0.5)
How Well Can Computers Do?
By analyzing *only* the distribution of part-of-speech (e.g., nouns, verbs, adjectives), already performs much better than human judges!
Classifier Performance (SVM with 5-fold CV)

Accuracy

- **Best Human Variant**: 61.9%
- **Classifier: Part-of-Speech**: 73%
- **Classifier: Words**: 89.8%

→ No human performs at this level in deception literature!
Data-driven Discovery of Insights into Deceptive Writings
Informative writing (left) --- nouns, adjectives, prepositions

Imaginative writing (right) --- verbs, adverbs, pronouns

Rayson et. al. (2001)
<table>
<thead>
<tr>
<th>Category</th>
<th>Variant</th>
<th>Weight</th>
<th>Category</th>
<th>Variant</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOUNS</td>
<td>Singular</td>
<td>0.008</td>
<td>DECEPTIVE/IMAGINATIVE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plural</td>
<td>0.002</td>
<td></td>
<td>Base</td>
<td>-0.057</td>
</tr>
<tr>
<td></td>
<td>Proper, singular</td>
<td>-0.041</td>
<td></td>
<td>Past tense</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>Proper, plural</td>
<td>0.091</td>
<td></td>
<td>Present participle</td>
<td>-0.089</td>
</tr>
<tr>
<td>ADJECTIVES</td>
<td>General</td>
<td>0.002</td>
<td></td>
<td>Singular, present</td>
<td>-0.031</td>
</tr>
<tr>
<td></td>
<td>Comparative</td>
<td>0.058</td>
<td></td>
<td>Third person</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>Superlative</td>
<td>-0.164</td>
<td></td>
<td>singular, present</td>
<td>-0.063</td>
</tr>
<tr>
<td>PREPOSITIONS</td>
<td>General</td>
<td>0.064</td>
<td></td>
<td>Modal</td>
<td></td>
</tr>
<tr>
<td>DETERMINERS</td>
<td>General</td>
<td>0.009</td>
<td></td>
<td>General</td>
<td>0.001</td>
</tr>
<tr>
<td>COORD. CONJ.</td>
<td>General</td>
<td>0.094</td>
<td></td>
<td>Comparative</td>
<td>-0.035</td>
</tr>
<tr>
<td>VERBS</td>
<td>Past participle</td>
<td>0.053</td>
<td></td>
<td>Personal</td>
<td>-0.098</td>
</tr>
<tr>
<td>ADVERBS</td>
<td>Superlative</td>
<td>-0.094</td>
<td></td>
<td>Possessive</td>
<td>-0.303</td>
</tr>
<tr>
<td>PRONOUNS</td>
<td></td>
<td></td>
<td></td>
<td>Pre-determiners</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Truthful Reviews ≈ Informative Writing (Journalism)

Deceptive Reviews ≈ Imaginative Writing (Novels)
STRONG DECEPTIVE INDICATORS

A focus on who they were with
In this example, “My husband;” also words like “family.”

Greater use of first-person singular
Fake reviews tend to use “I” and “me” more often.

Direct mention of where they stayed
Hotel and city names were less common in truthful reviews, which focus more on details about the hotel itself, like “small” or “bathroom.”

“My husband and I stayed in the [hotel name] Chicago and had a very nice stay! The rooms were large and comfortable. The view of Lake Michigan from our room was gorgeous. Room service was really good and quick, eating in the room looking at that view, awesome! The pool was really nice but we didn’t get a chance to use it. Great location for all of the downtown Chicago attractions such as theaters and museums. Very friendly staff and knowledgable, you can’t go wrong staying here.”

SLIGHT DECEPTIVE INDICATORS

High adverb use
“Very” and “really” are both used twice; “here” is used once.

High verb use
“Get”, “go”, “use”, “can’t”, “didn’t”, “eating”, “had”, “looking”, “stayed”, “was” (three times), “were.”

Use of “!?” and positive emotion
Deceptive reviews tend to use exclamation points, while truthful reviews used more punctuation of other kinds, including “!”.
A focus on who they were with
In this example, “My husband;” also words like “family.”

Greater use of first-person singular
Fake reviews tend to use “I” and “me” more often.

Direct mention of where they stayed
Hotel and city names were less common in truthful reviews, which focus more on details about the hotel itself, like “small” or “bathroom.”

“My husband and I stayed in the [hotel name] Chicago and had a very nice stay! The rooms were large and comfortable. The view of Lake Michigan from our room was gorgeous. Room service was really good and quick, eating in the room looking at that view, awesome! The pool was really nice but we didn’t get a chance to use it. Great location for all of the downtown Chicago attractions such as theaters and museums. Very friendly staff and knowable, you can’t go wrong staying here.”

- lack of spatial, sensorial details (Vrij et al., 2009)
- lack of descriptive adjectives: low, small, shiny
- less use of prepositions
instead, story telling:

-- why they were there: “vacation”, “business”, “anniversary”
-- whom they were with: “husband”, “family”
Exaggeration, words over the top: “fantastic”, “luxurious”, “gorgeous”, “awesome”
Superlatives: “the most”, “best”, “ever”
Certainty: “absolutely”, “definitely”, “for sure”

The room was clean and comfortable. The view of Lake Michigan from our room was gorgeous. Room service was really good and quick, eating in the room looking at that view, awesome! The pool was really nice but we didn’t get a chance to use it. Great location for all of the downtown Chicago attractions such as theaters and museums. Very friendly staff and knowledgeable, you can’t go wrong staying here.”
Increased level of “first person singular” “I”, “me”, “my”, “mine”

In contrast to psychological distancing (Newman et al., 2003)
 ➔ deception cues are domain dependent
What happened after then ( = 2011) ?
1. We built better detection models

① Syntax Improves Deception Detection
(Feng et al., ACL 2012)
--- 3 product review dataset
--- 1 essay dataset (Mihalcea and Strapparava (2009))

② Natural V.S. Distorted Distributions of Opinions
(Feng et al., ICWSM 2012, best paper runner up)
2. We excited other researchers

185 citations
3. Been featured by media outlets
(Highlights 2011-2014)


![Images of media outlets logos]


- [EMNLP 2013] Where Not to Eat? Improving Public Policy by Predicting Hygiene...
4. We hope NLP for Social Good

- When our work was first published in 2011, no clear legal regulations against fake reviews.
- Not any more! New York law enforcement charged 19 firms $350,000 for facilitating fake reviews (Sep 2013).
  - (not based on automatic detection)
Conclusion (Part I – Deception)

- Learning to read the “intent” of the author, even a hidden one.
- Humans not good at this task.
- Computers can at times perform better than humans, even without full blown semantic understanding.
- Data-driven discovery of insights to complement hypothesis-driven research

Ganganath, Jurafsky, McFarland (EMNLP 2009)

- computers predict flirtation intention better than humans can, despite humans having access to vastly richer information (visual features, gesture, etc.).
From Language to the Mind

Unconventional Case Studies:

I. Deceptive Reviews (ACL 2011)
II. Success of Novels (EMNLP 2013)

"HOW" it is said i.e., Writing Style

Information "WHAT"
Intent "WHY"
Identity "WHO"
Predicting the success of novels

- Novelty
- Style of writing
- Story line
- Social context
- Luck!
Describing the Visual World in Natural Language
Task:
Learning to Describe Images in Natural Language

Two approaches:

I. **BabyTalk**  
   - Formulaic image description  
     - CVPR 2011

II. **TreeTalk**  
   - Expressive image description  
     - TACL 2014 (in submission), ACL 2013, ACL 2012

How people write

“A butterfly having lunch”

Web Imagery
This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair, and near the potted plant.
“This picture shows one person,
“This picture shows one person, one grass,
“This picture shows one person, one grass, one chair,
“This picture shows one person, one grass, one chair, and one potted plant.
“This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass,
“This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair.”
“This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair, and near the potted plant.”
Methodology Overview

This is a photograph of one person and one brown sofa and one dog. The person is against the brown sofa. And the dog is near the person, and beside the brown sofa.
Conditional Random Fields (CRF)
Potential Functions for CRF

\[ \psi(\text{object }_i) \]  
\[ \psi(\text{attribute }_i) \]  
\[ \psi(\text{preposition }_ij) \]  

unary potentials

\[ \psi(\text{attribute }_i , \text{object }_i) \]  
\[ \psi(\text{object }_i , \text{preposition }_ij , \text{object }_j) \]  

relational (binary & ternary) potentials
Potential Functions for CRF

Practical challenge of relational potentials:
→
observing all possible combinations of variables unlikely
(limited corpus with detailed visual annotations)

unary potentials

relational (binary & ternary) potentials

ψ(object_i)
ψ(attribute_i)
ψ(preposition_rijk)

visual potentials

ψ(attribute_i, object_i)

ψ(object_i, preposition_rijk, object_j)

textual potentials
Computer: “This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair, and near the potted plant.”

Human (UIUC Pascal dataset):
A. A Lemonade stand is manned by a blonde child with a cookie.
B. A small child at a lemonade and cookie stand on a city corner.
C. Young child behind lemonade stand eating a cookie.
This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair, and near the potted plant.

Computer: “This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair, and near the potted plant.”

How can we reduce the gap between these two?

Human (UIUC Pascal dataset):

A. A Lemonade stand is manned by a blonde child with a cookie.
B. A small child at a lemonade and cookie stand on a city corner.
C. Young child behind lemonade stand eating a cookie.
Welcome to Amazon.com Books!

One million titles, consistently low prices.

(If you explore just one thing, make it our personal notification service. We think it's very cool!)

Spotlight! -- August 16th
These are the books we love, offered at Amazon.com low prices. The spotlight moves EVERY day so please come often.

One Million Titles
Search Amazon.com's million title catalog by author, subject, title, keyword, and more... Or take a look at the books we recommend in over 20 categories... Check out our customer reviews and the award winners from the Hugo and Nebula to the Pulitzer and Nobel... and bestsellers are 30% off the publishers list.

Eyes & Editors, A Personal Notification Service
Like to know when that book you want comes out in paperback, or when your favorite author's next book is published... we'll tell you in a heartbeat.

Computer Shares Lift Stocks Again; Bonds Are Weak
By Dave Pettit
Money & Investing Update

<table>
<thead>
<tr>
<th>MARKETS DIARY</th>
<th>5 p.m. EDT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Res. 10-year</td>
<td>6.0%</td>
</tr>
<tr>
<td>10-Year Treasury Note</td>
<td>6.1%</td>
</tr>
<tr>
<td>30-Year Treasury Bond</td>
<td>6.6%</td>
</tr>
<tr>
<td>Eurodollar (per $100,000)</td>
<td>92.52</td>
</tr>
<tr>
<td>German mark (per $100)</td>
<td>1.4750</td>
</tr>
</tbody>
</table>

What's News —
Business and Finance
Web Today: Increasingly Visual
-- social media, news media, online shopping

- Facebook.com has over 250 billion images uploaded as of Jun 2013
- 1.15 billion users uploading 350 million images a day on average
Task:
Learning to Describe Images in Natural Language

Two approaches:

I. **BabyTalk**  
   - Formulaic image description  
     - CVPR 2011

II. **TreeTalk**  
   - Expressive image description  
     - TACL 2014 (in submission), ACL 2013, ACL 2012
Given a query image (& an object)

① **Harvest** tree branches

② **Compose** a new tree by combining tree branches

SBU Captioned Photo Dataset (Ordonez et al. 2011)
Example Composition:

the dirty sheep meandered along a desolate road in the highlands of Scotland through frozen grass
A cow was staring at me in the grass in the countryside.
Sentence Composition :=

1. Select a subset of harvested phrases
2. Decide the ordering of the selected phrases

A cow
in the grass
was staring at me
in the countryside

A cow
was staring at me
in the grass
in the countryside
Sentence Composition :=

1. Select a subset of harvested phrases
2. Decide the ordering of the selected phrases

Tree Structure --- Probabilistic Context Free Grammars (PCFG)

Target Image:
A cow in the grass was staring at me in the countryside

Object (NP):
```
NP
  DT  NN
    a  cow
```

Action (VP):
```
VP
  VBD  VBG
    was  staring
      IN  NP
        at  PRP
          me
```

Stuff (PP):
```
PP
  IN
    in
      DT  NN
        the  grass
```

Scene (PP):
```
PP
  IN
    in
      DT  NN
        the  countryside
```
In the grass --- was staring at me --- a cow
A cow --- was staring at me --- in the countryside
Sentence Composition :=

In the grass --- was staring at me --- a cow

SINV

VP

PP

IN

the

NP

grass

VP

VBD

was

VBG

staring

IN

at

NP

PRP

me

NP

different from parsing because we must consider different choices of subtree selection and re-ordering simultaneously

Object (NP)

Action (VP)

Stuff (PP)

Scene (PP)
Sentence Composition as Constraint Optimization using Integer Linear Programming


Sentence Composition:

In the grass --- was staring ---

\[ \begin{array}{c}
\text{Object (NP)} \\
\text{Action (VP)} \\
\text{Stuff (PP)} \\
\text{Scene (PP)}
\end{array} \]

\( \Rightarrow \) different from parsing because we must consider different choices of subtree selection and re-ordering simultaneously

\( \Rightarrow \) finding the optimum selection+ordering = NP-hard (≈ TSP)
Sentence Composition as Constraint Optimization using Integer Linear Programming

decision variable: \( \alpha_{ijk} = 1 \) iff phrase \( i \) of type \( j \) selected for position \( k \in [0, N) \)

objective function: \( F = \sum_{ij} F_{ij} \times \sum_{k=0}^{N-1} \alpha_{ijk} \)

Content selection score based on visual recognition and matching

Object (NP)

Stuff (PP)

Scene (PP)
Sentence Composition as Constraint Optimization using Integer Linear Programming

Decision variable: 
\[ \alpha_{ijk} = 1 \quad \text{iff} \quad \text{phrase } i \text{ of type } j \text{ selected for position } k \in [0, N) \]

\[ \alpha_{ijkpq(k+1)} = 1 \quad \text{iff} \quad \alpha_{ijk} = \alpha_{pq(k+1)} = 1 \]

Objective function: 
\[ F = \sum_{ij} F_{ij} \times \sum_{k=0}^{N-1} \alpha_{ijk} + \sum_{ijpq} F_{ijpq} \times \sum_{k=0}^{N-2} \alpha_{ijkpq(k+1)} \]

Object (NP)

Content selection score based on visual recognition and matching

Language model score for local linguistic cohesion

Google Web 1-T Dataset
Sentence Composition as Constraint Optimization using Integer Linear Programming

Decision variable:  
\[ \alpha_{ijk} = 1 \quad \text{iff} \quad \text{phrase \ i \ of \ type \ j \ selected} \]

for position \( k \in [0, N) \)

\[ \alpha_{ijkpq(k+1)} = 1 \quad \text{iff} \quad \alpha_{ijk} = \alpha_{pq(k+1)} = 1 \]

Objective function:  
\[ F = \sum_{ij} F_{ij} \times \sum_{k=0}^{N-1} \alpha_{ijk} + \sum_{ijpq} F_{ijpq} \times \sum_{k=0}^{N-2} \alpha_{ijkpq(k+1)} \]
decision variable:  
\[ \alpha_{ijk} = 1 \quad \text{iff} \quad \text{phrase } i \text{ of type } j \text{ selected} \]
\[ \text{for position } k \in [0, N) \]
\[ \alpha_{ijkpq(k+1)} = 1 \quad \text{iff} \quad \alpha_{ijk} = \alpha_{pq(k+1)} = 1 \]

objective function:  
\[ F = \sum_{ij} F_{ij} \times \sum_{k=0}^{N-1} \alpha_{ijk} \ + \sum_{ijpq} F_{ijpq} \times \sum_{k=0}^{N-2} \alpha_{ijkpq(k+1)} \]
In this section we describe a constraint optimization approach to tree composition using extracted tree fragments, as above. We extract the above set of phrases for each object prior to the phrase extraction and composition. In scenes, making it possible to obtain richer visual linguistic regularities with respect to objects, actions, and scenes, considering both the parse structure and n-gram cohesion.

decision variable:
\[ \alpha_{ijk} = 1 \quad \text{iff phrase } i \text{ of type } j \text{ selected for position } k \in [0, N) \]
\[ \alpha_{ijkpq(k+1)} = 1 \quad \text{iff } \alpha_{ijk} = \alpha_{pq(k+1)} = 1 \]

objective function:
\[ F = \sum_{ij} F_{ij} \times \sum_{k=0}^{N-1} \alpha_{ijk} + \sum_{ijpq} F_{ijpq} \times \sum_{k=0}^{N-2} \alpha_{ijkpq(k+1)} \]
Objective function:

\[
F = \sum_{ij} F_{ij} \times \sum_{k=0}^{N-1} \alpha_{ijk} + \sum_{ijpq} F_{ijpq} \times \sum_{k=0}^{N-2} \alpha_{ijklpq(k+1)}
\]

Decision variable:

\[
\alpha_{ijk} = 1 \quad \text{iff} \quad \text{phrase } i \text{ of type } j \text{ selected for position } k
\]

\[
\beta_{ij} = 1 \quad \text{iff} \quad \text{cell } i,j \text{ of the matrix is assigned with PCFG tag } s
\]

\[
\beta_{ijk} = 1 \quad \text{iff} \quad \beta_{ijk} = \beta_{ikp} = \beta_{(k+1)jq} = 1
\]

Language model score for global parse tree structure:

\[
+ \sum_{ij} \sum_{k=i}^{j-1} \sum_{r} F_r \times \beta_{ijk}
\]

Language model score for local linguistic cohesion:

\[
+ \sum_{ijpq} \sum_{k=0}^{N-2} F_{ijpq} \times \sum_{r} \alpha_{ijklpq(k+1)}
\]

Global sentence structure:
Sentence Composition as Constraint Optimization using Integer Linear Programming

Objective function:

\[
F = \sum_{ij} F_{ij} \times \sum_{k=0}^{N-1} \alpha_{ijk} \\
+ \sum_{ijpq} F_{ijpq} \times \sum_{k=0}^{N-2} \alpha_{ijkpq(k+1)} \\
+ \sum_{ij} \sum_{k=i}^{j-1} \sum_{r \in R} F_r \times \beta_{ijk\text{r}}
\]

(Content selection ~ Visual Rec)

(Sequential cohesion ~ Lang Model)

(Tree structure ~ PCFG Model)

Constraints:

Consistency between sequence variables \(\alpha_{ijk}\) & tree leaf variables \(\beta_{ij}s\)

\[
\forall_{ijk}, \alpha_{ijk} \leq \sum_{s \in S^j} \beta_{kks}
\]

\[
\forall_{k}, \sum_{ij} \alpha_{ijk} = \sum_{s \in S} \beta_{kks}
\]

Valid PCFG parse tree

\[
\forall_{ij}, \sum_{s \in S} \beta_{ij}s \leq 1
\]

\[
R_h = \{ r \in R : r = h \rightarrow pq \}
\]

\[
\forall_{i,j>h,h}, \beta_{ijh} = \sum_{k=i}^{j-1} \sum_{r \in R_h} \sum_{r \in R} \beta_{ijk\text{r}}
\]

\[
\forall_{k \in [1,N]}, \sum_{s \in S} \beta_{kks} \leq \sum_{i=k}^{N-1} \sum_{s \in S} \beta_{0ts}
\]

\[
\forall_{ij}, \sum_{k} \gamma_{ijk} \leq 1
\]

Decision variable:

\(\alpha_{ijk}\) \(\alpha_{ijkpq(k+1)}\) (Sequential)

\(\beta_{ij}s\) \(\beta_{ijk\text{r}}\) (Tree structure)
Machine Caption VS Human Caption (forced choice w/ Amazon Mechanical Turk)

- Final system (seq + tree + pruning): 24% win

\[\text{Bleu@1！}\]

Machine Translation: From Images to Text

- Machine Caption VS Human Caption (forced choice w/ Amazon Mechanical Turk)

\[\text{Final system (seq + tree + pruning): 24% win}\]
The duck sitting in the water.

The flower was so vivid and attractive.

This window depicts the church.

Blue flowers are running rampant in my garden.

Good Examples

Correct choice of an action verb

Highly expressive!

Interesting choice of an abstract verb!
Mini Turing Test: our system wins in ~ 24% cases!

Scenes around the lake on my bike ride.

Spring in a white dress.

Blue flowers have no scent. Small white flowers have no idea what they are.

Almost poetic, situationally relevant

This horse walking along the road as we drove by.

Maybe the most common bird in the neighborhood, not just the most common water fowl in the neighborhood!

The duck was having a feast.
Examples with Mistakes

- The couch is definitely bigger than it looks in this photo.
- My cat laying in my duffel bag.
- Yellow ball suspended in water.
- A high chair in the trees.
Examples with Mistakes

A cat looking for a home. The other cats are making the computer room. ???

The castle known for being the home of Hamlet in the Shakespeare play.
Acknowledgements

- **My PhD**: Ritwik Banerjee, Song Feng, Jun Seok Kang, Polina Kuznetsova
- **Other PhD**: Vikas Ashok, Ritwik Bose, Jianfu Chen, Vicente Ordonez, Karl Stratos, Siming Li,
- **MS**: Manoj Harpalani, Kailash Gajulapalli, Anupam Gogar, Rohith Menon, Ruchita Sarawgi, Sandesh Singh, Longfei Xing, Girish Kulkarni, Sagnik Dhar, Visruth Premraj
- **Undergrad**: Alyssa Mensch, Jesse Dodge
- **Professor**: Hal Daumé III, Jia Deng, Alex Berg, Tamara Berg, Claire Cardie, Jeffrey Hancock, Rob Johnson, Michael Luca
- **Industry**: Margaret Mitchell, Sujith Ravi, Ravi Kumar, Michael Hart, Myle Ott, Amit Goyal
Natural Language Processing
Artificial Intelligence
Machine Learning
Computer Vision

(algorithms
+ statistics
+ probabilities
+ programming
+ ...)
Question?
From Language to the Mind

Unconventional Case Studies:
I. Deceptive Reviews (ACL 2011)
II. Success of Novels (EMNLP 2013)

"HOW" it is said i.e., Writing Style

Information "WHAT"
Intent "WHY"
Identity "WHO"
Predicting the success of novels

Novelty
Style of writing
Story line
Social context
Luck!
Publishers do make mistakes

Paul Harding’s “Tinkers” that won 2010 Pulitzer Prize for Fiction was rejected couple times before publication.

Rejected ~12 times before publication.
Can Computers Predict the Success of Novels without Really Reading the Book?

- based only on writing style
- stylistic correlates of successful novels?
How to define success
How to quantify success
Popularity v.s. Literary Quality

Best Seller amazon.com

Project Gutenberg

Downloads

2013-10-10

last 7 days

last 30 days

THE NEW YORK TIMES BOOK REVIEW
Best Sellers

FICTION

1. BUDGET, by Pamela McLaughlin (Warner). $2.95. The murder of a cardinal leads a Yale professor and an underworld model to the Middle East, where they uncover clues to a conspiracy kept hidden by the Shiners.

2. KNIGHTS OF DARKNESS, by G.K. Easton. $2.95. A golden Southpaw attempts to reclaim the throne from the wicked Scaly clan.

3. THE BASTARD TABBY, by T.S. Dwayne. $2.95. The murder of a cardinal leads a Yale professor and an underworld model to the Middle East, where they uncover clues to a conspiracy kept hidden by the Shiners.

4. GREAT FISH, by Liz Martin (Simon & Schuster). $2.95. The biblical story of Jonah, retold from the point of view of the whales.

5. NICK BOYE'S SHOCK BLADE: LYNCHPIN, by Ivan Moskowitz (Broadman & Holman). $2.95. After a coup by Admiral Cho, nine thousand to destroy the internet, the shockblade team is forced to ally with their Chinese rivals.

NONFICTION

1. CRACKED LIKE TERRA, by Dorothy Egan (Morrow). $2.95. A memoir of petty crime, drunkenness, and recovery, by a woman who was addicted to paint thinner by age nine.

2. EMPATHY IN WORCESTER, by James Weinbach. $2.95. A memoir of the author's childhood and his experiences with a boy who was addicted to paint thinner by age nine.


Dataset

• Project Gutenberg
  – free ebooks.
  – Title, author, genre, download count.
• 50 books per class, 8 genres.
Dataset

- Project Gutenberg
  - offers over 40,000 free ebooks.
  - Title, author, genre, download count.
- 50 books per class, 8 genres.
- <=2 books per author.

Authorship attribution
Prediction:
(based on best performing features, 5-fold CV with SVM)

<table>
<thead>
<tr>
<th>Category</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adventure</td>
<td>84</td>
</tr>
<tr>
<td>Mysterious</td>
<td>75</td>
</tr>
<tr>
<td>Fiction</td>
<td>75</td>
</tr>
<tr>
<td>History</td>
<td>61</td>
</tr>
<tr>
<td>Love</td>
<td>82</td>
</tr>
<tr>
<td>Poetry</td>
<td>76</td>
</tr>
<tr>
<td>Sci-fi</td>
<td>77</td>
</tr>
<tr>
<td>Short story</td>
<td>78</td>
</tr>
</tbody>
</table>

Average accuracy: **77.2%**
This is Surprising Because…

• Not considering any other influencing factor, not actually understanding the story, only looking at writing styles

• Different writers have wildly different writing styles. Should there even be stylistic commonalities shared by those different individuals?

• Testing: only the books by previously unseen authors (who presumably have his/her own unique writing style)
Secret Elements in Successful Novels

(only as correlates, not to be confused as causality)
Writing Style of Journalism (Douglas and Broussard 2000, Rayson et al. 2001)
Readability & Literary Success

Easier to Read

Harder to Read

More Successful

Less Successful
Readability & Literary Success
Success in Academic Journals (best paper awards)

Easier to Read

More Successful

Harder to Read

Less Successful

Sawyer et al (2008) @ Journal of Marketing
Readability & Literary Success

Easier to Read

Harder to Read

More Successful

Less Successful
Readability & Literary Success

1. Increased use of VP= better readability (Pitler and Nenkova (2008))
2. Readability Indices:

<table>
<thead>
<tr>
<th>METRIC</th>
<th>More Successful</th>
<th>Less Successful</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOG index</td>
<td>9.88</td>
<td>9.80</td>
</tr>
<tr>
<td>Flesch index</td>
<td>87.48</td>
<td>87.64</td>
</tr>
</tbody>
</table>
Less successful:

- verbs that are explicitly descriptive of actions and emotions: want, went, took, promise, cry, shout, jump, glare, urge
- extreme words: never, very, breathless, absolutely, perfectly
- cliche: love (desires, affair), body parts (face, arms, skin), obvious locations (beach, room, boat, avenue)

More successful:

- verbs that describe thought-processing: recognized, remembered
- verbs for reports or quotes: said
- prepositions: up, into, out, after, in, within
- discourse connectives: and, which, though, that, as, after

except for “think”, which is a more direct and general word
From Language to the Mind

Unconventional Case Studies:
I. Deceptive Reviews
   (ACL 2011)
II. Success of Novels
    (EMNLP 2013)

Information “WHAT”
Intent “WHY”
Identity “WHO”

intellectual traits (~ cognitive identity)
Bibliography (2011 – 2013)

I. Deception & Public Opinion
   - EMNLP 2013 Where Not to Eat? Improving Public Policy by Predicting Hygiene...
   - ICWSM 2012 Distributional Footprints of Deceptive Product Reviews.
   - ACL 2012 Syntactic Stylometry for Deception Detection

II. Authorship & Writing Style
    - EMNLP 2012 Characterizing Stylistic Elements in Syntactic Structure.
    - CoNLL 2011 Gender Attribution: Tracing Stylometric Evidence Beyond Topic...
    - ACL 2011 Language of Vandalism: Improving Wikipedia Vandalism Detection..

III. Connotation
     - ACL 2013 Connotation Lexicon: A Dash of Sentiment Beneath the Surface Meaning.
     - EMNLP 2011 Learning General Connotation of Words using Graph-based Algorithms.

IV. Literary Success & Linguistic Creativity
    - EMNLP 2013 Success with Style: Using Writing Style to Predict the Success of Novels.
    - EMNLP 2013 Understanding and Quantifying Creativity in Lexical Composition.
Research Outlook

1. Many more surprising and impactful applications --- yet to be discovered, formulated, and explored!
2. Computers may at times perform better than humans.
3. NLP for Digital Humanities (... and for Humanities) --- Data-driven discovery of insights vs. hypothesis-driven

"HOW" it is said i.e., **Writing Style**

Information "WHAT"

Intent "WHY"

Identity "WHO"