Text as Strategic Choice

Yanchuan Sim
Carnegie Mellon University

Thesis Defense
June 27, 2016
Outline

1. Motivation
2. Background
3. Strategic Text in the Scientific Community
4. Strategic Text in the Supreme Court
5. Conclusions
Two Problems in Text Modeling

• Understanding the author

• Understanding the audience’s response
Two Problems in Text Modeling

Author writes Text evokes Response
Two Problems in Text Modeling

Author writes Text

Text evolves Response
An Alternative View

Desired response

Author writes Text

motivates influences evokes

Response
Outline

1. Motivation
2. Background
3. Strategic Text in the Scientific Community
4. Strategic Text in the Supreme Court
5. Conclusions
Text as Choice

• A decision-theoretic problem

• Text is the choice variable

\[ U(\theta) = \text{Response}(\theta) - \text{Cost}(\theta) \]

• Observed text is the result of a utility maximizing action

• Incorporate utility into models of text
Probabilistic Models

• Powerful, versatile statistical framework for encoding our model assumptions

• Well established methods for parameter estimation, i.e., Markov chain Monte Carlo (MCMC) sampling
Topic Models

**Brief**

- 50% criminal procedures
- 20% immigration
- 20% death penalty
- (and many other topics)

**Topics**

- reasonable doubt
- supervised release
- grand jury
- immigration
- immigration law
- due process clause
- death penalty
- habeas corpus
- reasonable doubt
Outline

1. Motivation
2. Background
3. Strategic Text in the Scientific Community
4. Strategic Text in the Supreme Court
5. Conclusions
Modeling Strategic Behavior in the Scientific Community

• A probabilistic model of scientific papers, citation counts and author preferences.

• Researchers write papers to elicit response from the community and yet stay close to their preferences.
Researchers have preferences

- Researchers have preferences, $\eta$
- Prefer to write papers that are closer to their own preferences

$$\text{Cost}(\theta) = \frac{1}{2} \|(\eta + \epsilon) - \theta\|^2_2$$
Researchers desire citations

• Citations are a form of recognition

• We choose to write papers that bring us more citations

\[ \text{Response}(\theta) = \beta^\top \theta \]
Researchers’ Utility Function

\[ U(\theta) = -\frac{1}{2} \| (\eta + \epsilon) - \theta \|_2^2 + \kappa \beta^\top \theta \]

- **Cost**: Closeness to preferences
- **Response**: Expected citation counts

Researchers writes Papers evokes Citations
Co-authorship Utility

\[ U(\theta) = \frac{1}{|\mathcal{A}|} \sum_{a \in \mathcal{A}} -\frac{1}{2} c_a \| (\eta_a + \epsilon_a) - \theta \|_2^2 + \kappa_a \beta^\top \theta \]
Maximizing Utility

\[ \theta \sim \mathcal{N} \left( \sum_{a \in A} c_a \eta_a + \kappa_a \beta, \sum_{a \in A} c_a^2 \right) \]

preferences

author

"contribution"

tradeoff
Temporal Dynamics

• Preferences and research trends shift over time!

• Solution
  - Each year has its own set of parameters
  - Time series regularizer (Yogatama et al., 2011) to share statistical strength across years
Model Estimation

• Probabilistic model
  - Can be estimated using familiar techniques!

• Markov Chain Monte Carlo (MCMC) methods
  - Metropolis-Hastings (Hastings, 1970)
  - Gibbs sampling
  - Slice sampling (Neal, 2003)
Predicting Citation Counts

Mean Absolute Error vs No. of topics

- Yogatama et al (2011)
- Author utility
- TimeLDA

No. of topics: 8, 16, 32, 64, 128

Mean Absolute Error: 3.4, 3.2, 3.0, 2.8, 2.6
Perplexity

Perplexity (lower == better)

Author-Topic
Author utility (–Time)
Author utility

No. of topics

Author-Topic
Author utility (–Time)
Author utility
Trade-off vs. Seniority

![Graph showing median values and mean citations per paper vs. academic age.](image-url)
Outline

1. Motivation
2. Background
3. Strategic Text in the Scientific Community
4. Strategic Text in the Supreme Court
5. Conclusions
Strategic Text in the Supreme Court

- A probabilistic model of the Supreme Court
  - Litigants file **merits briefs**
  - Amici curiae file **amicus briefs**
  - Justices **vote**
  - Justices author **opinions**
Strategic Authors

Amicus writes Briefs

Briefs influences Votes and Opinions
Ideal Points

• Ideal point —A justice/legislator/voter’s preference in a continuous space (ideological continuum).
A 2D view
Scaling Ideal Points Using Text
(Lauderdale & Clark, 2014)

\[ p(v_j) = \sigma(a + \phi_j^\top(b\theta)) \]
Amici Ideal Points (Model A)

\[ p(v_j) = \sigma(a + \phi_j^T (b\theta + c^p \Delta^p + c^r \Delta^r)) \]

Sim et al. (2015)
Persuasive Amici Ideal Point (Model P)

\[ p(v_j) \propto \sigma \left( a + b\psi_j^T \left( \theta + \frac{\chi_j}{|A|} \sum_k \pi_k \Delta_k \right) \right) \]

---

justice
“influenceability”

amicus
“persuasiveness”

---

Amicus briefs
Merits briefs

individual mandate
commerce clause
Vote seeking amicus

• The vote seeking amicus curiae’s utility function is

\[ U^{\text{vote}}(\Delta) = \mathbb{E}[\text{no. of votes} \mid \Delta] - D(\Delta, \theta) \]

\[ = \frac{1}{J} \sum_{j \in J} p(v_j = s) - D(\Delta, \theta) \]

Response

Cost
Vote seeking amicus

• The vote seeking amicus curiae’s utility function is

\[ U^{\text{vote}}(\Delta) = \mathbb{E}[\text{no. of votes} \mid \Delta] - D(\Delta, \theta) \]

• And the model likelihood becomes

\[ \mathcal{L}(\text{vote model}) \times \exp \eta U^{\text{vote}}(\Delta) \]
Vote Prediction

Accuracy (%)

5-fold
2013
2014

Random forest
Model P without utility
Model A
Model P with utility
Opinions Model

Majority opinion

Justice B’s topics

Justice A’s topics

merits topics

amicus topics

Dissenting opinion
Opinion seeking amicus

- The opinion seeking amicus curiae wants to find

\[ U_{\text{opinion}}(\Delta) = D(\Delta, \mathbb{E}[\Omega^s]) - D(\Delta, \theta) \]

Response  Cost
Opinion seeking amicus

• The opinion seeking amicus curiae wants to find

\[ U_{\text{opinion}}(\Delta) = D(\Delta, \mathbb{E}[\Omega^s]) - D(\Delta, \theta) \]

• and like before,

\[ \mathcal{L}(\text{opinion model} \mid \text{vote model}) \times \exp \eta U_{\text{opinion}}(\Delta) \]
Opinion Perplexity

- LDA
- Author-Topic (Rosen-Zvi et al., 2004)
- Opinion model without utility
- Opinion model with utility

Perplexity ($\times 10^3$)

- 5-fold
- 2013
- 2014
Amicus Persuasiveness

![Diagram showing the persuasiveness of different organizations and justices]

The diagram illustrates the persuasiveness of various organizations and justices, with organizations such as ACLU, ABA, and CWFA, and justices such as Bush, Clinton, Obama, Reagan, and Bush Sr. plotted on a log scale.

### Table 5: Justice Values and Average Amicus Citation Rates

<table>
<thead>
<tr>
<th>Officer</th>
<th>Citation Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antonin Scalia</td>
<td>0.162</td>
</tr>
<tr>
<td>Samuel A. Alito</td>
<td>0.268</td>
</tr>
<tr>
<td>Anthony M. Kennedy</td>
<td>0.286</td>
</tr>
<tr>
<td>John G. Roberts</td>
<td>0.468</td>
</tr>
<tr>
<td>Stephen G. Breyer</td>
<td>0.515</td>
</tr>
<tr>
<td>Elena Kagan</td>
<td>0.637</td>
</tr>
<tr>
<td>Sonia Sotomayor</td>
<td>0.714</td>
</tr>
<tr>
<td>Ruth B. Ginsburg</td>
<td>1.590</td>
</tr>
</tbody>
</table>

---

The ideal points model was first introduced by Poole and Rosenthal (1985) and has inspired a variety of IP models in SCOTUS (Lauderdale and Clark, 2014; Gerrish and Blei, 2011; Heckerman et al., 2014). The model measures the relative effect of amicus briefs on justices' decision-making process and the values of justices with how often they cite an amicus brief.

The Spearman's rank correlation coefficient between justices' vote IP and citation rates is 0.678. The latent variable $\pi_j$ refers to different presidential administrations and is large when a justice is shared between all cases that a justice participated in. The performance of more sophisticated methods is better than simpler methods. To estimate utility, we use econometrics and estimates structural utility-based decision-making models in SCOTUS (Lauderdale and Clark, 2014; Gerrish and Blei, 2011; Heckerman et al., 2014). More generally, extensive literature in psychology has shown that people choose media for “slant” to maximize profit (Gentzkow and Shapiro, 2010) and economists choosing research topics maximize certain career outcomes (Jelveh et al., 2008).
Outline

1. Motivation
2. Background
3. Strategic Text in the Scientific Community
4. Strategic Text in the Supreme Court
5. Conclusions
Conclusions

• Explored a novel approach towards text modeling

• Extended the machinery of probabilistic models and decision theory
  • Presented instantiations of our approach in two domains: scientific publishing and judiciary

• Take steps in tackling the challenges of NLP for social science
Future Directions

• Richer models of utility
  - “Exogenous” response functions
  - Collaborative authors
  - Competitive agents

• Richer text representations

• Latent variable neural networks for social science
Acknowledgements

• Thesis committee — Noah, Daniel, Ed, Jiang, Philip

• the Ark research group

• Collaborators

• Friends at CMU and UW
Thank You!