Opportunities and Challenges for a Bayesian Approach to Language Processing

Andrew Kehler
UC San Diego

Past, Present, Future

* Pre-90s: Handcrafted systems, world knowledge, rules, inference, etc

* 90s-present: Machine learning, annotated data, etc

* Future: Reverse engineering raw language data to extract knowledge with which to perform (statistical) inference

* A challenge problem: Detecting invited inferences ('elicitures')

  The boss fired the employee who was hired in 2002.

  The boss fired the employee who was embezzling money.

Bayesian Pronoun Interpretation (Kehler et al. 2008)

✤ Bayesian formulation:

$$P(\text{referent} \mid \text{pronoun}) = \frac{P(\text{pronoun} \mid \text{referent})\, P(\text{referent})}{\sum\limits_{\text{referent} \in \text{referents}} P(\text{pronoun} \mid \text{referent})\, P(\text{referent})}$$

Interpretation → $P(\text{referent} \mid \text{pronoun})$

Production (Subject Bias) → $P(\text{pronoun} \mid \text{referent})$

Prior Expectation (Semantics/Coherence) → $P(\text{referent})$

✤ Data is consistent with a scenario in which semantics/coherence-driven biases primary affect probability of *next-mention,* whereas grammatical biases affect *choice of referential form*

✤ Results in the counterintuitive prediction that production biases are insensitive to a set of factors that affect the ultimate interpretation bias

Testing the Theory: Inferred Causes

* Passage completion study:

    The boss fired the employee who was hired in 2002. He _____ [Control]

    The boss fired the employee who was embezzling money. He _____ [ExplRC]

    The boss fired the employee who was hired in 2002. _____ [Control]

    The boss fired the employee who was embezzling money. _____ [ExplRC]

* Analyze:

    * Coherence relations (Explanation or Other)

    * Next-mentioned referent (Subject or Object)

    * Form of Reference (free-prompt condition; Pronoun or Other)

Predictions

RC Type

[ExplRC] *The boss fired the employee **who was embezzling money.***
[Control] *The boss fired the employee **who was hired in 2002.***

Coherence Relations

ExplRC: fewer Explanations

Production Bias
P(pronoun | referent)

Next-Mention Biases
P(referent)

ExplRC: fewer object next-mentions
(i.e., more subject references)

Subjects: more pronouns
ExplRC: no effect

Interpretation Bias
P(referent | pronoun)

ExplRC: fewer object refs (= more subjects)
Pronoun prompt: more subject references

Results: All predictions confirmed

Two Lessons for Computational Approaches

* In supervised approaches, the lack of annotated training data is an impediment to using anything beyond the most general features

* But the Bayesian model suggests that we don't need it:

  * The likelihood (production model) can be trained on (limited amounts of) annotated data

  * The prior (next-mention model) can be trained on cases of unambiguous reference in large, raw corpora

* The situation is entirely analogous to Bayesian approaches to other tasks (speech recognition, machine translation) that use a task-independent language model trained on raw data to estimate the prior

Two Lessons for Computational Approaches

* Language interpretation is not a collection of separable comprehension/disambiguation problems.

* It is a complex, interconnected dynamical system.

* Theoretically-grounded, linguistically-rich, graphical models may provide the path to capturing the multidirectional flow of information required to make progress on certain problems.

* The uphill battles are nonetheless substantial (e.g., the problem of identifying when a relative clause conveys a cause).

Thanks!