

A Cost Metric for Pronoun Resolution: Uncertainty Increases Processing Cost

Olga Seminck & Pascal Amsili

- (1) Université catholique de Louvain, Institute of Neuroscience, Media innovation and intelligibility Lab, Centre de traitement automatique du langage
- (2) Université Sorbonne Nouvelle - Paris 3, Laboratoire Lattice, CNRS

olga.seminck@uclouvain.be

Saarbrücken, 25 October 2019

Presentation Outline

Introduction

Cost Metric

Experiment

Results

Conclusion

Introduction

Pronoun Resolution

Pronoun resolution is a form of anaphora resolution.

NP α_1 takes NP α_2 as its anaphoric antecedent if α_1 depends on α_2 for its interpretation.

(Van Deemter and Kibble 2000)

A secret's worth depends on the people from whom it must be kept.

The Shadow of the Wind, Carlos Ruiz Zafón

Pronoun resolution is the process of finding the antecedent of an anaphoric pronoun.

The Cognitive Cost of Pronoun Resolution

What leads to cognitive cost in pronoun resolution?

- ▶ Multiple Factors

- ▶ Grammatical role of the antecedent (Crawley et al. 1990)
- ▶ Parallel grammatical roles (Smyth 1994)
- ▶ Frequency of the antecedent (Shillcock 1982)
- ▶ Distance between the pronoun and the antecedent (Clark and Sengul 1979)
- ▶ ...

For modelisation of cognitive cost on corpus, difficult to take all factors into account

According to some theories, these factors reflect more abstract but broader processes

Cost Metric

Cost Metrics

Cost Metric: formula that predicts processing cost

- ▶ Translates hypothesis into prediction

Example: surprisal

- ▶ Hypothesis: unexpected events are harder to process
- ▶ Cost metric: $\text{Difficulty}(\text{event}) = -\log(P(\text{event}))$

A cost metric to predict the difficulty of pronouns

Prediction for pronouns resolution:

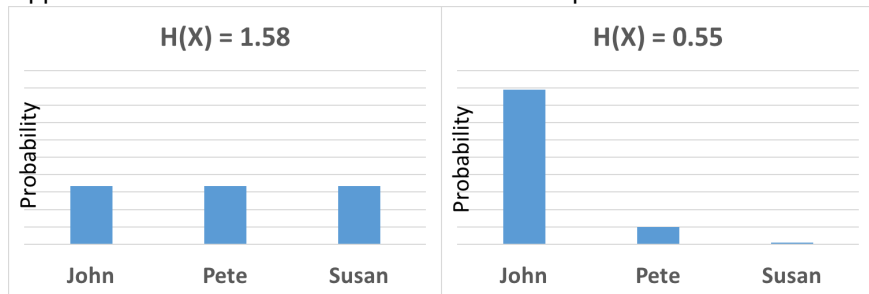
More uncertainty about the antecedent → more processing cost

Entropy: measure of uncertainty

$$H(X) = - \sum_{j \in X} p(X = j) \cdot \log_2(p(X = j))$$

Entropy

Applies to a random variable: antecedent of a pronoun



Relative Entropy

Entropy increases with the number of antecedent candidates.

- ▶ Keep scores comparable through the text
- ▶ 'Normalise' the entropy

Normalisation: relative entropy

'Distance' between actual probability distribution & flat distribution

$$H_{relative}(P||Q) = \sum_{i \in P \wedge i \in Q} P(i) \log \frac{P(i)}{Q(i)} \quad (1)$$

Larger distance \Rightarrow less uncertainty \Rightarrow less processing cost

Experiment

Test of the Entropy Cost Metric on Corpus

Does the cost metric make relevant predictions about reading behaviour on corpus?

- ▶ Dundee Corpus annotated for anaphoric pronouns
- ▶ Automatic coreference resolution system
 - ▶ Estimate relative entropy of pronouns

The Dundee Eye-Tracking Corpus (Kennedy et al. 2003)

Eye-movements of 10 native English speakers

Reading 65 texts

From the Independent (newspaper)

Total: 50 000 tokens

Annotated with part of speech (Frank 2010) and dependency relations (Barrett et al. 2015)

Annotation of the antecedent of all 1 109 anaphorical pronouns.
A data-set to study pronoun resolution in natural data.

Measuring reading time for pronouns: a hard problem

Pronouns are fixated only 20 - 30% of the time
(Ehrlich and Rayner 1983)

Pronouns are very short

Spill-over effects

Other experiments:

(Seminck and Amsili 2018; Von der Malsburg 2018)

Take a window of words around the pronoun.

- ▶ ... at a time [**when they are at greatest risk**], and then ...
- ▶ ... on it; [**but it would seriously degrade the**] quality ...

Problems:

- ▶ Need multiple models
- ▶ Few data-points per pronoun

Solution: binomial metric

A simpler reading metric was more suited:
Is the pronoun fixated or not?

Binomial outcome: yes/no answer.

Advantages:

- ▶ More data points
- ▶ There is only one point to measure

“a word is skipped because it has been identified on the previous fixation” (Brysbaert and Vitu 1998)

Hypothesis: a fixated pronoun indicates more processing difficulty.

An NLP-system gives the probability distribution

The Red House tells the story of **a mysterious, tormented individual** who breaks into **toy shops and museums** to steal **dolls and puppets**.
Once they are in his power...

1. Probability distribution from parameters of resolution system
2. Calculate relative entropy over this probability distribution

Antecedent of <u>they</u>	Probability	Relative entropy
The Red House	0.05	
a mysterious, tormented individual	0.04	
toy shops and museums	0.31	0.83
dolls and puppets	0.58	
\emptyset	0.02	

The resolution system

Obtain probabilities from a state of the art NLP-system
(Lee et al. 2017)

Lee et al.'s system:

- ▶ End to end: without pre-processing
- ▶ Neural-network architecture
- ▶ Ranking system

Result

Statistical Model

A statistical model predicted whether the pronoun is fixated or not.

Is the relative entropy of importance to this prediction?

Mixed effects model:

$$\text{fixated} \sim \text{length} + \text{frequency} + \text{comma} + \text{punctuation} + \text{rel_ent} \\ + (1 + \text{rel_ent} \mid \text{participant}) + (1 \mid \text{item})$$

The entropy cost metric predicts reading behaviour

The relative entropy was a predictor in reading behaviour.

A lower distance between the entropy and the maximal entropy
⇒ more participants fixating the pronoun

Estimate: -0.07 (95% Credible interval = [-0.01, -0.13])

Conclusion:

More uncertainty about the antecedent of the pronoun leads to more people fixating it.

Conclusion





Conclusion

Uncertainty about the antecedent leads to more cognitive cost in pronoun resolution





Notions from Information Theory are also relevant for pronoun resolution

NLP-systems can be used to estimate probabilities that are relevant to human language processing





Bibliography I

-  Barrett, Maria, Željko Agić, and Anders Sjøgaard (2015). “The Dundee Treebank”. In: *The 14th International Workshop on Treebanks and Linguistic Theories (TLT 14)*.
-  Brysbaert, Marc and Françoise Vitu (1998). “Word skipping: Implications for theories of eye movement control in reading”. In: *Eye guidance in reading and scene perception*. Elsevier, pp. 125–147.
-  Clark, Herbert H and CJ Sengul (1979). “In search of referents for nouns and pronouns”. In: *Memory & Cognition* 7.1, pp. 35–41.
-  Crawley, Rosalind A., Rosemary J. Stevenson, and David Kleinman (1990). “The use of heuristic strategies in the interpretation of pronouns”. In: *Journal of Psycholinguistic Research* 19.4, pp. 245–264.

Bibliography II

-  Ehrlich, Kate and Keith Rayner (1983). “Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing”. In: *Journal of Verbal Learning and Verbal Behavior* 22.1, pp. 75–87.
-  Frank, Stefan L (2010). “Uncertainty reduction as a measure of cognitive processing effort”. In: *Proceedings of the 2010 workshop on cognitive modeling and computational linguistics*. Association for Computational Linguistics, pp. 81–89.
-  Kennedy, Alan, Robin Hill, and Joël Pynte (2003). “The dundee corpus”. In: *Proceedings of the 12th European conference on eye movement*.
-  Lee, Kenton, Luheng He, Mike Lewis, and Luke Zettlemoyer (2017). “End-to-end Neural Coreference Resolution”. In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 188–197.

Bibliography III

-  Seminck, Olga and Pascal Amsili (2018). “A Gold Anaphora Annotation Layer on an Eye Movement Corpus”. In: *11th edition of the Language Resources and Evaluation Conference*.
-  Shillcock, Richard (1982). “The on-line resolution of pronominal anaphora”. In: *Language and Speech* 25.4, pp. 385–401.
-  Smyth, Ron (1994). “Grammatical determinants of ambiguous pronoun resolution”. In: *Journal of Psycholinguistic Research* 23.3, pp. 197–229.
-  Van Deemter, Kees and Rodger Kibble (2000). “On coreferring: Coreference in MUC and related annotation schemes”. In: *Computational linguistics* 26.4, pp. 629–637.



Von der Malsburg, Titus (Mar. 6, 2018). "The president will give her inauguration speech: Explicit belief and implicit expectations in language production and comprehension". In: *Forum Entwicklung und Anwendung von Sprach-Technologien (Oral Presentation)*. Universität des Saarlandes.