

## Stanford CoreNLP Coreference Resolution

Coreference resolution: What is it?.....	1
And why is it important?.....	1
Coreference Resolution with Stanford CoreNLP .....	2
Two types of coreference: nominal & pronominal .....	2
Manual coreference.....	2
References .....	4

### Coreference resolution: What is it?

Coreference resolution is the process of linking references to the same entity in a document. For example, in this sentence:

*"I voted for Nader because he was most aligned with my values," she said.*

(image © <https://nlp.stanford.edu/projects/coref.shtml>)

The red expressions (I, my, and she) refer to the same entity, and the blue expressions (Nader and he) refer to a different entity. It is necessary to determine which expressions refer to which entities for many different NLP tasks, such as summarization, question answering, and information extraction.

### And why is it important?

Coreference resolution is an important first step for the accuracy of other NLP tasks. For instance, it is one of the first data cleaning steps involved in the SVO extraction pipeline, in order to have consistent subjects and objects. Without coreference resolution, a frequency distribution of subjects or objects, for instance, may give you a list of "he" "she" "they" that may refer to completely different entities.

## Coreference Resolution with Stanford CoreNLP

We wrote a Java script – `Stanford_CoreNLP_coreference_resolution.jar` – that implements three different approaches to coreference resolution by Stanford CoreNLP:

Deterministic (fast rule based)

Statistical (machine learning requiring dependency parsing)

Neural network (most accurate and slowest)

The script replaces all expressions referring to the same entity in a text with one representative expression.

For example, the following sentence:

“**Bill Cato** attempted to assault **Mrs. Vickers**, but **her** **husband** stopped **him**.”

Would become:

“**Bill Cato** attempted to assault **Mrs. Vickers**, but **Mrs. Vickers’s** **husband** stopped **Bill Cato**.”

### *Two types of coreference: nominal & pronominal*

CoreNLP approaches implement both **pronominal** (i.e., pronouns referring to nouns, e.g., Barack Obama came to Boston; *he* said that...) and **nominal** (i.e., nouns referring to other nouns, e.g., Barack Obama came to Boston; *the President* said that...) coreference resolution. The algorithms do NOT resolve adverbial coreference (i.e., adverbs referring to nouns, e.g., Barack Obama came to Boston; *there* Obama said that...).


The NLP Suite SVO pipeline implementation of coreference filters out the nominal coreference and focuses on the pronominal coreference. Too many errors otherwise.

**Coreference resolution is still far from accurate, with perhaps 65% success.**

### Manual coreference

Due to the relatively low success rate of coreference resolution, the NLP Suite also implements a manual approach on the coreferenced output, displaying the original and coreferenced documents on two panels, side-by-side, **original on the left** and **coreferenced on the right** with the relevant pronouns and coreferences highlighted. Although the highlighting is not perfect, it does provide

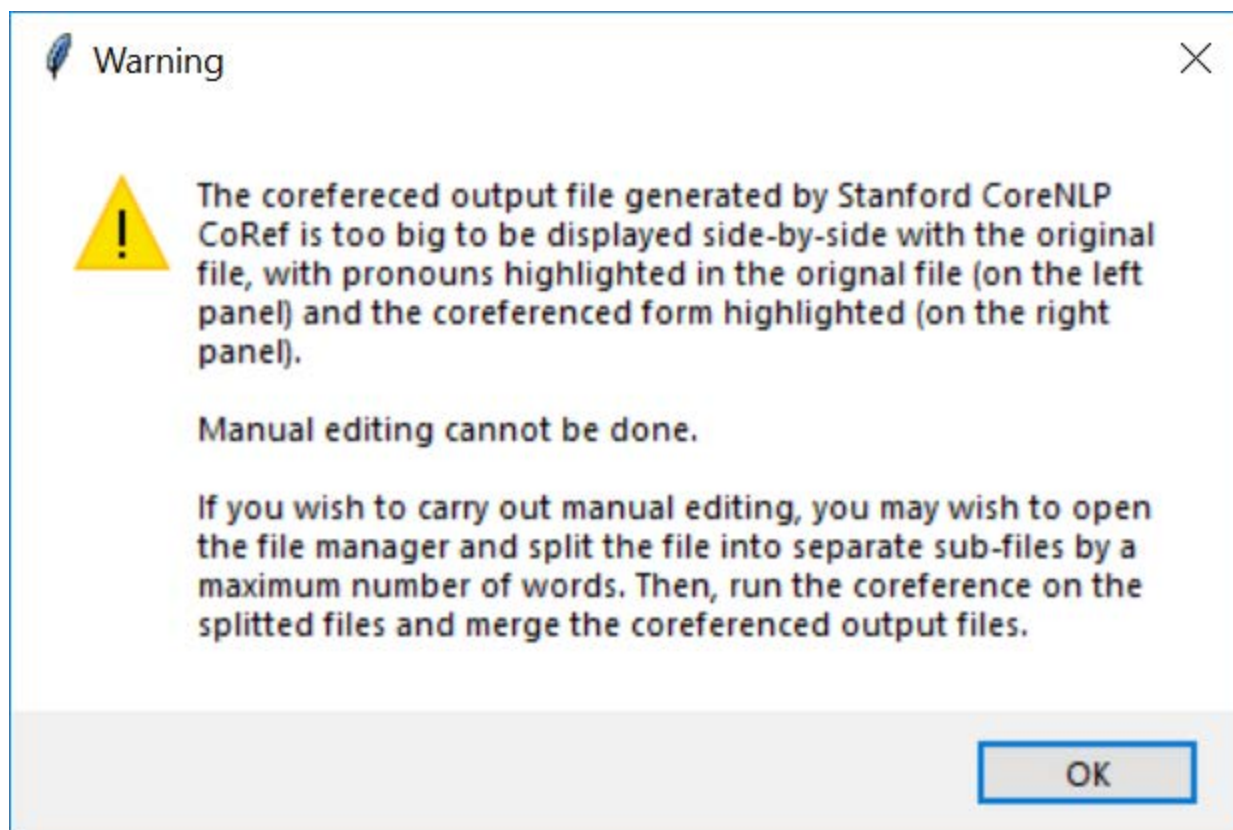
the user an immediate visual tool of comparison. The user can edit any unresolved or wrongly resolved pronominal cases.

 GUI for Stanford CoreNLP

Loading result from Neural Network (Edit text on the right hand side and Save)

<p>as an old sow with three little pigs, and as she had not enough p them, she sent them out to seek their fortune.</p> <p>st that went off met a man with a bundle of straw, and said to lease, man, give me that straw to build me a house".</p> <p>he man did, and the little pig built a house with it.</p> <p>ly came along a wolf, and knocked at the door, and said: "Little pig, let me come in".</p> <p>h the pig answered: "No, no, by the hair of my chiny chin chin".</p> <p>f then answered to that: "Then I'll huff, and I'll puff, and I'll blow your house in".</p> <p>uffed, and he puffed, and he blew his house in, and ate up the pig.</p> <p>ond little, pig met a man with a bundle of furze, and said: "Please, man, give me that furze to build a house".</p> <p>he man did, and the pig built his house.</p> <p>ong came the wolf, and said: "Little pig, little pig, let me come in, by the hair of my chiny chin chin".</p> <p>'ll puff, and I'll huff, and I'll blow your house in".</p> <p>uffed, and he puffed, and he puffed, and he huffed, and at last w the house down, and he ate up the little pig.</p> <p>rd little pig met a man with a load of bricks, and said: "Please, man, give me those bricks to build a house with".</p> <p>man gave him the bricks, and he built his house with them.</p> <p>wolf came, as he did to the other little pigs, and said: "Little pig, let me come in".</p> <p>, by the hair of my chiny chin chin".</p> <p>'ll huff, and I'll puff, and I'll blow your house in".</p> <p>e huffed, and he puffed, and he huffed and he puffed, and he puffed huffed; but he could not get the house down.</p> <p>found that he could not, with all his huffing and puffing, blow the house down, he said: "Little pig, I know where there is a nice f</p>	<p>THERE was an old sow with three little pigs, and a to keep their little pig's fortune.</p> <p>she sent their little pig's fortune.</p> <p>The first that went off met a man with a bundle of him: "Please, man, give me that straw to build m</p> <p>" Which the man did, and the little pig built a h little pig.</p> <p>Presently came along a wolf, and knocked at the do the pig, little pig, let me come in.</p> <p>" To which the pig answered: "No, no, by the ha n.</p> <p>" The wolf then answered to that: "Then the wol the wolf coming'll puff, and the wolf coming'll</p> <p>" So the wolf coming huffed, and the wolf coming coming blew the wolf coming's house in, and ate u</p> <p>The second little, pig met a man with a bundle of Please, man, give the second little, pig that furz</p> <p>" Which the man did, and the pig built a man's ho Then along came the wolf, and said: "Little pig, second little, pig come in.</p> <p>" " No, no, by the hair of the second little, pi</p> <p>" " Then the second little, pig'll puff, and the ll huff, and the second little, pig'll blow your h</p> <p>" So the wolf coming huffed, and the wolf coming coming puffed, and the wolf coming huffed, and at blew the house down, and the wolf coming ate up</p> <p>The third little pig met a man with a load of bric ase, man, give me those bricks to build a house wi</p> <p>" So the man gave the wolf coming the bricks, and house with those bricks to build a house with.</p> <p>So the wolf came, as a man did to the other little Little pig, little pig, let a man come in.</p> <p>" " No, no, by the hair of a man's chiny chin.</p>
---	--

Since the function works in memory, for large files memory this may not be an option. If that is the case, the script will warn the user.



The script similarly warns the user to deselect manual coreference when processing a directory.

## References

- Lee, Heeyoung, Angel Chang, Yves Peirsman, Nathanael Chambers, Mihai Surdeanu, and Dan Jurafsky. 2013. "Deterministic coreference resolution based on entity-centric, precision-ranked rules." *Computational Linguistics*, 39(4).
- Clark, Kevin and Christopher D. Manning. 2015. "Entity-Centric Coreference Resolution with Model Stacking." In *Proceedings of the ACL*.
- Clark, Kevin and Christopher D. Manning. 2016. "Deep Reinforcement Learning for Mention-Ranking Coreference Models." In *Proceedings of EMNLP*.
- Clark, Kevin and Christopher D. Manning. 2016. "Improving Coreference Resolution by Learning Entity-Level Distributed Representations." In *Proceedings of the ACL*.