Towards Humorous Anaphora Resolution

Anton Nijholt & Matthijs Mulder

An important concept of natural language processing is anaphoric reference. Anaphoric used words are words that are referring back to something that was earlier mentioned or that is known because of the discourse situation and/or the text as it is read or heard. The anaphorically used word is called 'the anaphor', the text to which it refers 'the antecedent'. The extralingual entity they corefer to is called the referent.

Deictic used words are words that refer to something directly or indirectly present in the situation. The word is then used instead of a gesture, or utterance of the word is accompanied by gesturing. Cataphorically reference is reference to something that follows in the text or that will be specified later by the text.

Anaphora resolution is the process of determining the antecedent of an anaphor. The antecedent can be in the same sentence as the anaphor, or in another sentence. The first case is called intra-sentential referencing, the second case inter-sentential referencing.

Incorrect resolution of anaphoric references can be used in order to create a humorous remark in a text or in a dialogue situation. Consider for example the text used in a Dilbert cartoon where a new "Strategic Diversification Fund" is explaned:

Adviser: "Our lawyers put your money in little bags, then we have trained dogs bury them around town."

Dilbert: "Do they bury the bags or the lawyers?"

Adviser: "We've tried it both ways."

In this presentation we discuss the possibility to use algorithms for anaphora resolution in order to obtain incorrect solutions and how we can 'guarantee' that this incorrect solution leads to a funny remark. Comparisons with existing computational humor research are made and it will be discussed what results can be used of that research.

Our general aim is to have embodied conversational agents that are able, depending on the dialogue or conversation, to make funny remarks in order to get a more natural (human-like) conversation. For that raeson we also investigated how the utterances in two of our corpora can give rise to funny remarks because of possibly incorrect reference resolution. One of the corpora contains dialgue where references are made to the dialogue only. The other corpus is about dialogues that contain also deictic references.