Coref and negation

*Some jotted down notes by ØE improved by critical questions from AC.*

* The introduction of negative statements into CRM or CRM extension. What are the consequences? Is it necessary?
* The relationship between positive and negative coref statements. Are they similar or do they just look similar in the same way that other pairs look similar but have different ontological status, e.g., existing and fictitious persons; past and future events?

# Open and closed world

CRM represents an open world. Co-reference assignment is not to be used for expressing co-reference as such, it is to express an explicit claim made by someone that a co-reference exists (or not, given a negative statement). In this context a closed world would be strange indeed. Would the lack of a co-reference statement mean that nobody ever made the claim that there is a co-reference? This is obviously unknowable.

If a negative co-reference statement is made it is not the latter kind of a negation. It is rather a claim that someone made another statement of the type non-co-reference assignment. In the following I will argue that this latter claim is too different from co-reference assignment to make it meaningful to lump them together in the same class.

# Explicit and implicit

It is quite clear that the communication about the relationship between implicit and explicit coref was partly misunderstood before the last meeting. Thus, not even within our group people understood what we tried to say. I had much “clearer” messages sent to me off-list than what made it to the SIG list -- I don’t mind but it is an interesting sign that this must be explained better.

For implicit co-reference, negation is not an issue (at least not in an open world system?) What is discussed here are the consequences for explicit coref.

# The coref target

This may be the core of the problem. Let me use one example to illustrate: When I discussed negative statements in my text-map experiments it was clear that in

(1) “there are farms in area A”

and

(2) “there are no farms in area A”

the ontological status of farms are different. In (1) the discussion is about a number of farms. We can point at the farms in the real world or refer to others doing so in documents and say, these farms are in area A. If the statement is false it is still a statement about farms claimed to exist.

In (2) we do not talk about farms at all. There are no farms referred to in the sentence. Rather the sentence is about that status of area A. It has the type of a “farm-free-area”. While this is in line with area A in sentence (1) being a “farm-containing-area”, there is still a difference in the use of the word “farms”.

If we see the farms as sets this may be slightly different. The difference is in the denotation (or referent?) of the word. In case 1 it denotes a set with a number of farms (objects), in 2 it denotes a null set i.e. the absence of farms (no farm objects but there could be other sets of objects). If the target (referent) is understood as a set, this difference might be less problematic. The set can have objects or be null yet still a valid target.

I would claim that a set based understanding of statements such as (1) and (2) is strange seen in light of a historical meaning, at least for the 18th century material I worked with. To me the reply to the question “which farms” is enlightening: for (1) the reply would be “Well, it is the one farmed by… and then the other one who used to be owned by…” whereas for (2) it would rather be “Are you deaf, stupid, or mad?”

# The existence of the target

If we claim that A and B co-refer we also claim that there exist an object X which is the common target of the reference functions of A and B.

If we claim that A and B does not co-refer we claim that X does not exist. Thus, using the same class (E91 Co-Reference Assignment) for both positive and negative coref statements is wrong as the properties are different for the two statement types. The former has a P155 coref target, the latter has none. This is always the case.

To be more precise: We have two propositional objects A, and B. A refer to X and B refer to Y. Claiming that they co-refer is to claim that B refer to X and A refer to Y, or, with other words: X is the same as Y. Because a coref target is an E1 entity and thus not necessarily a propositional object, X and Y are not pointers. They are (representation of) things. So X and Y are not different names of the same thing, they are the same thing and calling them X and Y is misleading.

Claiming that they do not co-refer is to claim that A does not refer to Y and B does not refer to X. In this case the referents of A and B are different. So even if there are reference targets there is no coref target. That A and B both have targets does not mean that there exist a coref target.

The semantics of a co-reference target must be “the Entity that two or more propositional objects refer to”. This target exist in the case of positive co-reference, but it does not exist in the case of negative co-reference.

The conclusion to this is:

1. we may add a class E91 coref assignment to CRM or to an extension
2. we may add a class Exx non-coref assignment to CRM or to an extension
3. we cannot do 2) without doing 1) but we can do 1) without doing 2)

To me it is clear that 1) is needed. I think 2) would be useful but I am uncertain about the consequences for the healthiness of CRM if we introduce it.

# Can positive and negative assignments be kept together?

I believe that removing property P155 has co-reference target from E91 will not be enough to make **E91 Co-reference Assignment** and **Exx Non-Co-reference Assignment** the same. Even if the attribute creating the explicit problem is removed, this discussion above has shown that the two classes have fundamentally different ontological status and cannot (that is, should not) be combined into one class.