Increasing the Genericity of the Mate Annotation Famework : The Case of Reference

Susanne Salmon-Alt, Laurent Romary, André Schaaff

Increasing the Genericity of the Mate Annotation Famework :

The Case of Reference

Susanne Salmon-Alt, Laurent Romary, André Schaaff

Loria – UMR 7503
Campus Scientifique, B.P. 239, 54506 Vandoeuvre-lès-Nancy Cedex, France
{Susanne.Alt, Laurent.Romary, André.Schaaff}@loria.fr

Abstract

This paper addresses the need of a meta-model for corpus annotation schemes. Through the analysis of a specific annotation level – the reference level defined in the MATE project – and the experience we gained in annoting multimodal corpora for , we show that the current MATE annotation scheme can be extended to a general framework.The available tags are integrated into an abstract model – a meta-scheme – which can be instantiated in different ways depending on the context of the actual corpus to be annotated (e.g. coreference annotation, reference annotation with or without gestures) oron the theory to be tested.

1. Introduction

This paper argues, through the analysis of a specific annotation level, that some degree of genericity has to be added to annotation schemes such as defined in the MLIS/MATE project (Mengel & al., 2000). Our aim is indeed to show that the scheme which has been proposed for coreference annotation in MATE (Poesio, 2000), although extremely valuable, may nevertheless face cases or theoretical domains which will demand specific extensions. Still, it does seem sensible to consider that one should not multiply the number of annotation schemes and thus that a general framework for extensions is becoming mandatory.

Our analysis will start from the coreference annotation scheme in Mate. We will show that there exist some problematic cases where the scheme has to be extended to deal with referential phenomena in an multimodal environment, in particular when we have to take into account perception and gestures. But since these particular cases do not hold for every annotation situation, we propose to use the MATE framework as a base for designing an abstract meta-model for referential annotation which can be instantiated in different ways depending on the goals of the annotator. The meta-model and possible instances are represented using XML schemata (Thompson & al, 2000). This ensures more flexibility than a classical description in terms of a DTD and make it easy to extend this approach to other annotation levels.

2. The MATE project

2.1. General Objectives

The aim of the MATE project (Multilevel Annotation Tools Engineering) is to facilitate re-use of language resources. This includes creating, acquiring, and maintaining language corpora. The problems are addressed through the development of a standard for annotating resources and through the provision of tools which will make the processes of knowledge acquisition and extraction more efficient. Specifically, MATE treats spoken dialogue corpora at multiple levels like prosody, syntax, coreference and dialogue acts.

2.2. Coreference Annotation

Within the framework of MATE, the coreference coding module is intended to cover the annotation of anaphoric relations. Anaphoric relations are considered to hold between two textual elements, where the first one is the antecedent and the second one the subsequent mention. The relation between these two elements may be an identity relation, (1) or a more complex relation like bridging (2)[1].

(1) maintenant il faut prendre <u>un grand triangle</u> et <u>le</u> mettre à gauche de l'écran
now you have to take <u>a big triangle</u> and to put <u>it</u> on the left of the screen

(2) maintenant je vais faire <u>la maison</u>
faut mettre <u>un toit</u> dessus
now I'm going to make <u>the house</u>
you have to put <u>a roof</u> on it

To take into account multimedia applications, it is also possible to code direct reference relations which can hold between a text span and an object that has not been mentioned before, but is accessible because it is part of the visible situation (3).

(3) [a range of two triangles visible on the screen….] donc tu vas commencer par prendre une petite barre que tu vas mettre à gauche de la pointe <u>du premier triangle</u>
so you will start by taking a small line and put it on the left of the top of <u>the first triangle</u>

[1] All examples are taken from a corpus of human dialogues (Ozkan, 1994) where two persons, A and B, have to reconstruct together simple pictures composed of geometrical figures.

The coding elements proposed by the MATE framework are the following :

Element	Meaning	Attributes
<coref:de>	discourse entity (antecedent or subsequent mention)	ID, HREF
<coref:link>	anaphorical or referential relation	HREF, TYPE
<coref:anchor>	discursive or visual antecedent	HREF
<coref:universe>	visual situation	ID
<coref:ue>	visible entity	ID

Table 1: MATE Coding scheme for coreference

It is important to note that the <coref:link> element has an obligatory attribute "type" in order to type the linking relation between the antecedent and the subsequent mention. Possible values are given in a list and include "ident", "member", "subset", "poss" (possessive) or "genrel" (general relation).

Applying the coding propositions to our examples (1) and (2) looks respectively like (4) and (5):

(4) maintenant il faut prendre <coref:de id="de1"> un grand triangle </coref:de> et <coref:de id="de2"> le </coref:de> mettre à gauche de l'écran
<coref:link href="de2" type="IDENT">
<coref:anchor href="de1"/>
</coref:link>

(5) maintenant je vais faire <coref:de id="de1"> la maison </coref:de>
faut mettre <coref:de id="de2"> un toit </coref:de> dessus
<coref:link href="de2" type="POSS">
<coref:anchor href="de1"/>
</coref:link>

3. The case of reference

In the following, we propose to distinguish different coding situations – or levels –, depending on the corpora to annotate and on the aim of the annotation.

Starting with a purely coreferential point of vue (section 3.1.1), we show that the MATE coreference coding scheme is valuable for coding most of anaphorical and associative relations.

Some cases for which a purely coreferential point of view cannot completely account underline the need of introducing a few supplementary elements in order to treat multimodal corpora, especially those including perceptual elements and gestures (sections 3.1.2 and 3.1.3).

For three situations – coreference coding, coding of situations including the perceptual environment and coding of situations including perception and gestures, we propose special coding schemes which are in fact instantiations of a meta-scheme derived from the MATE propositions (section 3.2). We show thus that the MATE framework could be easily extended to generate meta-models.

Another group of problematic cases on the coreferential level is linked to the theoretical notion of reference domains: without going into the coding details, we will show how the MATE notion of "universe" could be extended to deal with discourse phenomena where a discourse entity is not directly linked to another one, but actually in relation with a discourse domain comprising it (section 3.3).

3.1. Different coding levels for reference

3.1.1. The coreference level

One way to use the MATE scheme is to limit the annotation level to a purely coreferential level, i.e. to mark only links between two discourse entities (and not between a discourse entity and a universe entity). The first entity is considered as the anchor and the second one as a subsequent mention of the entity introduced by the anchor. 'Coreference' can be used here in both a narrow and a broader sense.

In the narrow sense, "coreference" means the relationship between two discourse entities referring to the same object and coincides with the idea of Hawkins 'anaphoric' use of definite descriptions (1978). Restricting the set of possible relations between two entities to the identity relation, like in MUCCS (Hirschmann, 1997), is sufficient for certain tasks and presents the advantage that it seems to be the most reliable annotation scheme (Poesio & Vieira, 1998). An example for coding coreference is given in (4).

In a broader sense, "coreference" includes both "anaphoric" and "associative" uses. In the later case, the entity introduced by the subsequent mention has to be inferred in some way from the entity introduced by the anchor. The manner of inferring can be coded using the different values for the type attribute of the link entity, like "POSS" in (5).

The necessary elements for coding coreference in this way are shown in Table 2. In particular, we do not need the universe elements, because only intra-discursive relationships are considered.

Element	Meaning	Attributes
<coref:de>	discourse entity (antecedent or subsequent mention)	ID, HREF
<coref:link>	coreferential relation	HREF, TYPE
<coref:anchor>	antecedent : discourse entity	HREF

Table 2 : Coding elements for coreference

However, even on this restricted coding level, we have to deal with some difficulties: one of them is the choice of the link type between two entities. (6) and (7) – (7) is taken from Dale (1992) – are examples for which the present state of the MATE framework is insufficient for coding the relation between the underlined element and its antecedent. In (6), none of the given link types can be used : the discourse entity "une deuxième" is not directly linked to "une grande barre verticale", but actually in relation with a discourse domain (or a class) introduced by this antecedent. In (7), the antecedent is not a single discourse entity, but a dynamically constructed set of previous discourse entities. For the moment, it is not possible to account for topic construction references which actually gather up a set of previously mentioned (but not explicitely coordinated) entities. The solution proposed for conjoined NPs – introduce a complex discourse entity – cannot be applied here, because the complex entity is dynamically modified during the discourse.

(6) alors tu prends une grande barre verticale et
tu en mets <u>une deuxième</u> à coté
now take a big vertical line and put <u>another one</u> beside it

(7) Peel and chop the onions and potatoes.
Crape and chop the carrots.
Slice the celery.
Melt the butter.
Add <u>the vegetables</u>.

These examples show that some work is still needed in order to account for more complex relations between discourse entities. Without going into the details for coding, we will see how these cases could be integrated into a general approach of reference (section 3.3).

3.1.2. The reference level

In addition to some tricky examples like (6) or (7) just discussed, there is another case which requires more complex coding schemes: in fact, dialogues present often not only referential links between two discourse units, but also between a discourse unit and an accessible object, for example in the perceptual environnement, like in (3). Coding such dialogues demands to introduce discourse external resources, for example for visible objects, and to allow links between these objects and discourse entities. For this reason, MATE follows Bruneseaux et Romary (1997) who propose to extend the basic coding scheme by additional elements for coding universes containing universe entities. These entities can then function as antecedents for a referring expression. Applying this scheme to (3), the coded example looks like (8):

(8)
```
<coref :universe  ID= "u1">
<coref:ue  ID ="ue1"> T1 </coref:ue>
<coref:ue  ID ="ue2"> T2 </coref:ue>
</coref :universe>
```

donc tu vas commencer par prendre une petite barre que tu vas mettre à gauche de la pointe <coref:de ID="de1"> <u>du premier triangle</u> </coref:de>

```
<coref:link href="de1" type="IDENT">
  <coref:anchor href="ue1"/>
</coref:link>
```

This scheme allows links between universe entities and discourse entities, but there is also a problem: it concerns the possible confusion of the coreferential and the referential level. In cases where visible objects are available, it is often possible to link a referring expression to both a textual antecedent and a visible object on the screen (9).

(9) [different objects visible on the screen…]
maintenant je voudrais prendre <u>le grand carré</u>
et <u>le</u> mettre un peu plus vers la gauche
now I would like to take <u>the big square</u> and to move <u>it</u> a little bit to the left

With the aim to be coherent, we propose to decide in favour of the referential level: 'reference' means here – in opposition to 'coreference' – a systematical link between a discourse entity (referring expressions) and a universe entity (the referent). The coding elements for this new scheme are presented in Table 3.

Element	Meaning	Attributes
<coref:de>	discourse entity	ID, HREF
<coref:link>	referential relation	HREF, TYPE
<coref:anchor>	universe entity	HREF
<coref:universe>	perceptual situation	ID
<coref:ue>	perceptual entities	ID

Table 3: Coding elements for reference

3.1.3. Adding gestures

For certain tasks of referential annotation, especially in multimodal dialogues, the coder may need to take into account another kind of markable entities: gestures. In our corpus, we found indeed some interesting examples of interaction between gestures and referring. (10) for example shows that without interpreting the gesture in B1, the referring expression in A2 cannot be understood. In (11), a referreing expression co-occurs with a gesture.

(10) A1 alors il va falloir que tu prennes deux grandes barres
now you have to take two big lines
B1 [takes a first line]
A2 et <u>l'autre</u> tu vas la mettre parallèle à la première
put <u>the other one</u> parallel to the first one

(11) [two circles visibles on the screen…]
et tu vas reprendre encore un petit triangle
et le placer au-dessus de <u>ce rond</u>
and now take still another small triangle
and put it over <u>this circle</u> [+gesture]

In order to represent these phenomena we need a framework that allows us to mark, additionnally to discourse elements, gestures and to link them to universe elements. Therefore, we introduce a new element – <kinesic> – to the scheme in Table 3. This element corresponds to gestures of pointing or moving visible objects. It has at least two attributes: an ID and a description. It functions like a discourse entity: it can be linked to an entity in the universe. The new scheme is shown in Table 4.

Element	Meaning	Attributes
<coref:de>	discourse entity	ID, HREF
<coref:link>	referential relation	HREF, TYPE
<coref:anchor>	universe entity	HREF
<coref:universe>	perceptual situation	ID
<coref:kinesic>	pointing or moving	ID DESC
<coref:ue>	perceptual entities	ID

Table 4: Coding elements for reference

Coding (11) by applying this scheme leads to (12) :

```
(12) <coref :universe  ID= "u1">
     <coref:ue ID ="ue1"> C1 <coref:ue>
     <coref:ue  ID ="ue2"> C2</coref:ue>
     </coref :universe>

     et tu vas reprendre encore un petit triangle
     et le placer au-dessus de <coref:de
     ID="de1"> ce rond </coref:de >

     <coref:kinesic ID="ki1"
       desc="pointing"> </coref:kinesic>

     <coref:link href="de1" type="IDENT">
       <coref:anchor href="ue1"/>
     </coref:link>
     <coref:link href="ki1" type="IDENT">
       <coref:anchor href="ue1"/>
     </coref:link>
```

3.2. A meta-model for a reference coding scheme

In 3.1., we presented different coding levels for reference annotation. At each level – excepting the reference level including gestures, where we introduced a new element, <kinesic> – we used a subset of elements proposed by the MATE framework. For this reason, we think that it should be possible to consider this framework as a starting point for constructing an abstract coding scheme for reference, from which it would be possible to derive the schemes for our three coding levels.

3.2.1. The meta-model : abstract elements

Comparing the elements used at the three coding levels leads to the construction of the following structure of an abstract coding scheme: At the top, we find the document, composed of an optional head, an optional element for external resources and a body. The body contains two kinds of elements : markable segments and links between these segments and an anchor. This general structure is illustrated in the head of Table 5. It is instantiated in different ways depending on the aimed coding levels.

3.2.2. Instantiation of abstract elements

Table 5 gives an overview about the instantiation of the abstract elements for reference coding. It sums up the coding schemes presented in section 3.1. The most important features are:

- no external resources at the coreference level
- universe elements as anchors at the reference level
- gestures as markable elements at the gesture level

Coding level	Document			
	Head	External Resources	Body	
			Markable segments	Links Anchors
Co-reference	optional	no	discourse elements	discourse elements
Reference	optional	universe & universe elements	discourse elements	universe elements
Reference Gestures	optional	universe & universe elements	discourse elements & gestures	universe elements

Table 5: Instantiation of abstract elements for reference

3.2.3. Modelling schemata using XML: an example

Both the meta-model and the instantiations can be modelled using XML schemata (Thompson & al., 2000). The main idea is to use the possibility of defining abstract elements and types in order to represent the structure of the meta-scheme and to instantiate these elements depending on the actual coding level. Figure 1 shows an extract of the meta-scheme concerning the definition of the "body" element. Note that the element "segment" is declared as abstract. Figure 2 provides the instantiation schema for the reference/gesture level. In this case, a segment is replaced by either a discourse element or a kinesic element.

```
<xsd:element name="body">
 <xsd:complexType>
      <xsd:element name="segment" minOccurs="1"
        maxOccurs="unbounded"/>
      <xsd:element name="link" minOccurs="0"
        maxOccurs="unbounded"/>
 </xsd:complexType>
</xsd:element>

<xsd:complexType name="Segment" abstract="true">
 <xsd:attribute name="id" type="xsd:ID" use="required"/>
</xsd:complexType>
```

```
<xsd:element name="segment" type="Segment" abstract="true"/>
```

Figure 1 : Extract of the XML meta-scheme

```
<xsd:complexType name="Segment_de" base="target:Segment"/>

<xsd:complexType name="Segment_ki" base="target:Segment">
      <xsd:attribute name="desc"/>
</xsd:complexType>

<xsd:element name="de" type="Segment_de" equivClass=
"target:segment"/>

<xsd:element name="kinesic" type="Segment_ki"
equivClass="target:segment"/>
```

Figure 2 : Extract of the XML scheme for reference/gestures

3.3. Looking forward

After having shown that the MATE framework for coding coreference can be used as a base for a meta-model of different annotation situations, we will give some ideas about possible further investigation which can lead to an extended use of the basic schemes.

3.3.1. Reference Domains

In section 3.1.2., we introduced the elements <universe> and <universe entity> for taking into account extradiscursive elements in the perceptual environment as antecedents for referring expressions. But one can consider that the use of these universes may not be limited to perceptual entities. In fact, every referring expression has an extradiscursive counterpart, which is itself part of a larger universe or reference domain. This is for example the approach of Dale (1992) or that of Corblin (1987) who consider that any referring expression isolates an entity in a larger domain. The theoretical advantage of this point of view is a unified approach to reference, covering at the same time classical coreferential cases (1), bridging (2), perceptual contexts (3) and even "larger situation uses" (Hawkins, 1978).

The domain may not be limited to physically accessible elements like a perceptual group of triangles on a screen (3). Rather, domains can be introduced and activated by referring expressions themselves. Note that this is the case in (10), where " deux grandes barres" in A1 introduces explicitely a domain from which are extracted both the referent of the gesture in B1 and of the referring expression "l'autre" in A2. But domains are also introduced implicitely: in (6), "une grande barre verticale" activates a generic class *big_vertical_line* from which is extracted the referent of "une deuxième". Without this notion of activated domains, the link between these two expressions is not easy to explain.

Using the universe elements for coding reference domains in general, rather than only for coding perceptual environments, could be one possibility to extend the use of the proposed annotation schemes.

3.3.2. Dynamic universes

A second possibility consists of extending the use of the universe elements for coding dynamic evolutions of reference domains: in fact, it is well known that the set of available referents changes during the discourse processing (Grosz & Sidner (1986), Kamp & Reyle (1993)). This can be the case for visual environments, for example a set of objects on a screen, but also for cognitive constructs like domains composed during the discourse processing. Without taking into account these facts, it is not possible to explain what happens in the example (7).

4. Conclusion and Prospectives

From the experience we gained in coding corpora for reference in the context of testing various theoretical hypotheses, we conclude that the current MATE annotation scheme for reference should be seen as a general framework where the five available tags (i.e. *<de>, <link>, <anchor>, <universe>* and *<ue>*) should be integrated into an abstract framework – a meta-scheme – which can be instantiated in different ways depending on the context of the actual corpus to be annotated (e.g. when gestures are also available) or on the theory to be tested. This instantiation process then relies on the selection of a set of categories to be used to typify discourse entities, links and universes.

We think that this approach can be extended to any annotation scheme as defined in MATE (e.g. POS annotation, dialogue acts, etc.). We suggest that any annotation module should be based upon a two-tiered description comprising :

- a meta-model describing the abstract structure needed for the corresponding module (markable source segments, link structure, external resources) and
- a set of data-categories from which a given user can select a subset or even, when needed, to which he can add specific categories.

From the meta-model, it is then possible to generate automatically an XML-schema incorporating the corresponding constraints resulting from the data-category description.

Using this representation instead of DTDs leads to the unification of the syntax for describing the DTD and coding the documents. It increases the modularity of the description, allows for the control of data types and facilitates the maintenance and the design of interfaces. It ensures, while providing more flexibility, an even better context to validate annotation files relevant to a given module.

5. References

Bruneseaux F. et Romary L., 1997. Codage des références et coréférences dans les DHM *ACH-ALLC '97*, Kingston Ont. CAN.

Corblin F., 1987. *Indéfini, Défini et Démonstratif*, Droz, Genève.

Dale R., 1992. *Generating Referring Expressions. Constructing Descriptions in a Domain of Objects and Processes.* The MIT Press, Cambridge, London.

Grosz B.J. and Sidner, C. ,1986. Attention, Intention and the Structure of Discourse. *Computational Linguistics,* 12, 175-204.

Hawkins, J. A., 1978. *Definiteness and Indefiniteness*. Croom Helm, London.

Hirschman, L., 1997. *MUC-7 Coreference Task Definition, Version 3.0.* In Proc. MUC-7.

Kamp H. and Reyle U., 1993. *From Discourse to Logic.* Kluwer Academic Publishers. Dordrecht, Boston, London.

Mengel, A., Dybkjaer, L., Garrido, J.M., Heid, U., Klein, M., Pirrelli, V., Poesio, M., Quazza, S., Schiffrin, A., and Soria, C., 2000. *MATE Dialogue Annotation Guidelines*, January 2000. (http://www.ims.uni stuttgart.de/projekte/mate/mdag/).

Ozkan N.,1994. Vers un modèle dynamique du dialogue : analyse de dialogues finalisés dans une perspective communicationnelle. Ph.D.Thesis, Cognitive Science, INP Grenoble.

Poesio, M., Vieira R., 1998. A corpus-based investigation of definite description use. Computational Linguistics, 24 : 2.

Poesio M., 2000. Coreference. MATE Dialogue Annotation Guidelines-Deliverable D2.1, January 2000, 126-182.(http://www.ims.uni stuttgart.de/projekte/mate/ mdag/).

XML Schema Part 1 and 2. W3C Working Draft 7 April 2000. Edited by Thompson H. S., Beech D., Maloney M., Mendelsohn N. (http://www.w3.org/TR/2000/WD-xmlschema-1-20000407/).