

A Model of Grouping for Plural and Ordinal References

Alexandre Denis, Guillaume Pitel, Matthieu Quignard

LORIA, BP239 F-54206 VANDOEUVRE-LES-NANCY

Contact author: [Guillaume.Pitel \[at\] gmail.com](mailto:Guillaume.Pitel@gmail.com)

We present a model for the resolution of plural references on groupings based on Reference Domains Theory. While the original theory does not take plural reference into account, this paper shows how several entities can be grouped together by building a new domain and how they can be accessed later on. We introduce the notion of super-domain representing the access structure to all the plural referents of a given type.

Introduction

In the course of a discourse or a dialogue, referents introduced separately could be referenced with a single plural expression (pronoun, demonstrative, etc.). The grouping of these referents may depend on many factors: it may be explicit if they were syntactically coordinated or juxtaposed, or implicit if they only share common semantic features (Eschenbach *et al.*, 89). Time is also an important factor because it may be difficult to group old mentioned referents with new ones. Because of this multiplicity of factors, choosing the right discursive grouping for a referential plural expression is ambiguous, and this ambiguity needs to be explicitly described.

We present a model of grouping based on Reference Domains Theory (Salmon-Alt, 01) which considers that a reference operation consists of extracting a referent in a domain. However the original theory barely takes into account plural reference. This paper shows how several entities can be grouped together by building a new domain and how they can be accessed later on. It introduces also the notion of super-domain D^+ that represents the access structure to all the plural referents of type D . This work is being implemented and evaluated in the MEDIA/EVALDA project (Devillers, 04).

The goal of this research is both to find a practical solution to deal with the kind of situations we met in the corpus of the MEDIA/EVALDA project, and to improve the coverage of the Reference Domain Theory, which is a representational theory of reference that focuses on describing the selection preferences from several ambiguous candidates.

1. Groupings and plural anaphora

Several kinds of clues can specify that referents should be grouped together, or at least could be grouped together. These clues may occur at several language levels, from the noun phrase level to the rhetorical structure level. We have not explored in detail the different ways of groupings entities together in a discourse or dialogue. We just describe here some of the phenomena we were confronted with while developing a reference resolution module for a dialogue understanding system.

- **Explicit Coordination** - The most basic way to explicitly express the grouping of two or more referents is using a connector such as *and*, *or*, *as well as*, etc.
“*Good afternoon, I would like to book a single room **and** a double room*”
- **Implicit Coordination** - An implicit coordination occurs when two or more referents of the same kind are present in one sentence, without explicit connector between them.
“*Does the hotel de la gare have a restaurant, like the Holiday Inn?*”
- **Repetitions/Specifications** – In some particular cases, groupings are explicitly described by the enumeration of their referents. For instance “*Two rooms. A single room, a double room*”.
- **Inter-Sentential** – In the course of a dialogue, referential expressions can be grouped together, depending on several factors (common type, common predicate, semantic link).

Most of these different situations have already been thoroughly investigated in previous work. However, these methods are, from our point of view, unable to fulfill the needs we met in the particular task of the MEDIA/EVALDA project, especially with plurals, otherness, and ordinals being very frequent in our corpus.

In the standard model of plurals in the DRT (Kamp & Reyle, 93), discourse referents are grouped and assigned to a plural discourse referent (this is represented using the \oplus operator), but no information can be assigned to the relative role of the individual referents within the group (which is necessary for the resolution of ordinal anaphora). Moreover, without the presence of specific markers or constructions, it seems difficult to allow the emergence of several groupings from a single list of referring expressions (for instance in the case of a co-occurrence of several referents - X,Y,Z - in the same predicate, while several others - Y,Z,W - share a common type). Other approaches deal with referring expressions sharing the same type for making a group (Eschenbach *et al.*, 89), which is not sufficient for our problems, since sharing a common type is only one of the enablers of grouping.

2. Reference Domains Theory

The Reference Domains Theory (Salmon-Alt, 01) supposes that every act of reference is related to a certain domain of interpretation, in that it both describes how to extract a referent and which set of elements to extract it from. In the reference domains theory, an act of reference also modifies the structure of the reference domains of the discourse, in term of focus and partitions.

A reference domain is composed of any group of entities in the hearer’s memory (discursive referents, visual objects, or concepts) and describes how each entity could be addressed through a referential expression. The theory has been developed in order to represent the diversity of access modes to the referents. The claim is that every referential expression has a different behavior which depends on the vericonditional description its referent must satisfy and on its conditions of use (the actual structure of context).

The theory considers the referring process as a dynamic extraction of a referent in a domain instead of a binding between two entities (Salmon-Alt, 00). Hence doing a reference act consists of isolating a particular entity from other rejected candidates (Olson, 70), amongst all the accessible entities composing the domain. This dynamic discrimination relies on projecting an access structure focusing the referent in the domain and facilitates further access: any extraction in a domain increases its salience, thus it is preferred for the next

interpretations.

The preferences for choosing a suitable domain are inspired from the Relevance theory (Sperber & Wilson, 86) taking into account such focalization and salience. F. Landragin & L. Romary (03) have also studied the usage of reference domains in order to model a visual scene.

2.1. Basic type

A reference domain is a structure which can reference entities by differentiating them. It is modelled as a set of **entities** (ground) and a set of **partitions** of these entities (equivalency classes or alternatively differentiation functions as in Pitel, 03). Each class groups the elements which are accessible by the same referential expression *i.e.* which can be viewed as the same object from a certain point of view. Identifying completely a referent requires to find a domain where a partition gives an unambiguous access to this entity. Using a referential expression presupposes then (optimistically) that this condition is respected for all the speakers.

2.2. Access structures

We suppose that any distinction between the referents from the excluded alternatives requires highlighting the discrimination criterion opposing them. This criterion, given by the referential expression and its context of use behaves like a partition of the accessible entities, grouping them together according to their similarities and their differences. A partition may have one of its parts **focused** (or profiled). There are, at least, three kinds of discrimination criteria:

- **discrimination on description.** Entities can be discriminated by their type, their properties, or by the relations they have with other entities. For example the name of the hotels is a discrimination criterion in “*the Ibis hotel and the hotel Lafayette*”.
- **discrimination on focus.** Entities can also be discriminated by the focus they have when they are mentioned in the discourse or designated by a gesture. For example, “*these rooms*” would select focused referents in a domain, whereas “*the other room*” would select a non-focused one.
- **discrimination on time of occurrence.** Entities can also be discriminated by their occurrence in the discourse. For example “*the second hotel*” would discriminate this hotel by its rank in the domain.

Every referential expression aims to distinguish referents and exhibit a differentiation criterion; however the referents are sometimes not distinguished intentionally. For example, the indefinite plural “*two hotels*” will introduce two hotels that cannot be differentiated. But even in this case it is possible to project a differentiation structure *a posteriori* for example by next saying “*the first one*”. Figure 1 shows the state of a domain after the sequence “*two hotels*”, “*the first one*” (the ground is represented here as a concept in description logics, each subdomain is then a lower concept in the hierarchy). The domain H of the two hotels contains a partition which focalizes the first of its elements, the subdomain H₁.

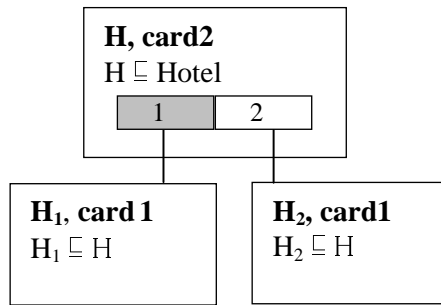


Figure 1: A domain containing two subdomains

2.3. Classical resolution algorithm

Each activated domain belongs to a list of domains ordered according to their recentness (the referential space). The resolution algorithm consists of two phases:

1. It searches a suitable, preferred domain in the referential space when interpreting a referring expression. The suitability is defined by the minimal conditions the domain has to conform to in order to be the base of an interpretation (particular description, or presence of a particular access structure). The general preference factor is the minimization of the access cost (recentness, salience or focalization).
2. It extracts a referent and restructures the referential space, taking into account this extraction. It not only focuses the referent in its domain, but also increases the salience of the domain itself which will be preferred for further extractions.

According to the determination and description of the referential expression or to the gesture made to access to the referent, this generic scheme will be instantiated in different ways. For example a definite “*the N*” will search for a domain in which a particular entity can be solely discriminated by its type N, and the restructuring consists of focalizing the found referent in this domain. A demonstrative “*this N*” behaves differently in that it tries directly to access to the referent without imposing a strong discrimination criterion on the type, *i.e.* it finds a focalized referent in a domain which could be cast into a “N” during the restructuring phase. See (Landragin & Romary, 2003) for a classification of the different access modes.

The algorithm highlights two types of ambiguities, domain ambiguities when there are many preferred domains of interpretation with no mean to choose during the first phase, and the referent ambiguities when many referents are found without preferences. Of course a domain ambiguity implies a referent ambiguity. In a dialogue system, it is not the role of the reference module to disambiguate completely referents, but instead to propagate the ambiguity to next modules which, for instance, will solve it by asking clarifications to the user.

3. Super-domains

In order to take groupings into account in the Reference Domains Theory, we introduce two constructs in our formal toolbox. Indeed, having only one kind of domain construct doesn’t allow for a correct distinction between different referent statuses.

First we distinguish plural and simple domains. The simple domains D serve as bases for profiling a **subpart**, or **related part** of a simple referent. For instance, if $D = Room$, then

one can profile a *Price* from D . The plural domains D^* serve either as a **generic base** or as a **plural representative** for profiling a simple domain D . A generic base is mandatory in our model to support the insertion of new extra-linguistic referents evoked with an indefinite construct (for instance “*I saw a black bird on the roof*”), while plural representatives are used for explicit groupings. A domain D^*_1 can also be profiled from a D^*_0 , provided D^*_1 profiles a subset of the elements of D^*_0 .

Second, we introduce the notion of **super-domain** D^+ , from which a D^* can be profiled. The relations allowed between domains are represented on figure 2. A super-domain D^+ is the domain of all groupings D^* , including a special D^*_{all} grouping which is the representative of all evoked instances of a given category. This configuration is not intended to deal with long dialogues where several, trans-sentential groupings occur, and where older groupings may become out of access. Doing this would require a rhetorically driven structuring of the D^*_{all} .

As Reference Domain Theory is primarily targeted toward extra-linguistic referents occurring in practical dialogue, the construction of domain trees representing the supposed structuring of referent accessibility is based on an ontology. As a consequence, for each “natural” type and each subtype (for instance $Room \wedge Single$), a domain tree is potentially created (actually, one can easily imagine how this creation may be driven ‘on-demand’).

Another evolution from the initial Reference Domain Theory is the possibility to focalize several items of a partition. Indeed, since the resolution algorithm can focalize a whole plural domain, all elements of this domain must be focalized in all the plural domains they occur in.

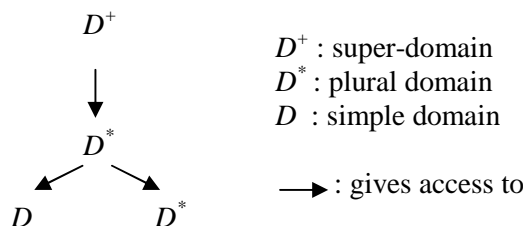


Figure 2: Access tree of Reference Domains

When new extra-linguistic referents are evoked, they are individually profiled under the D^*_{all} corresponding to their types (that is, their “natural” type, and all the subtypes they are eligible to). When some sentence-level grouping occurs or when a plural extra-linguistic referent is evoked, a D^* is created, with each of its components as children, when possible (that is, when each component is described). Figure 3 illustrates the state of the *Hotel* domain tree after a scenario with at least two dialogue acts, the first one introducing $Hotel_1$, the second one inserting a grouping of $Hotel_2$ and $Hotel_3$ (due to their co-occurrence in the same utterance). One can see that all referents introduced are accessible through the special $Hotel^*_{all}$ domain.

In short:

- All new referents (singular or plural) become subdomains of D^*_{all}
- All new plural referents build up a subdomain of D^+

When a referring expression occurs, one performs the resolution through the following algorithm:

- If the referring expression is singular, performs the classical resolution algorithm in the plural domains D^* (including D^*_{all})
- If the referring expression is plural, performs the classical resolution algorithm with D^+ as the base

U: The Ibis Hotel ($Hotel_1$) is too expensive
 S: How about the Hotel Lafayette ($Hotel_2$) or the Hotel de la cloche ($Hotel_3$)

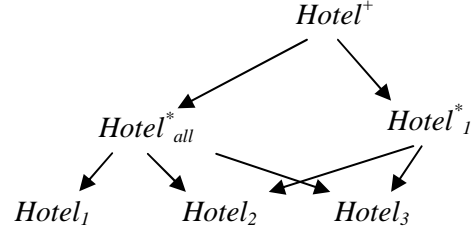


Figure 3: A domain tree built from a the scenario on the left, containing a grouping

4. Examples

A sample dialogue (figure 4) is analyzed through the algorithm presented above. This example shows how the referents introduced in an explicit coordination could be referenced as a whole “the two hotels”, or extracted discriminately by an ordinal “the second one” or by an otherness expression “the other one”. All the subdomains of H^+ (i.e. the plural domains of hotels) are indicated *after* each interpretation using a simplified notation. Only the ordered list of accessible entities and their focalization (bold) are noted for each subdomain (only one access structure is represented for each domain). For instance $H^*_{all} = (h_1, h_2, h_3)$ means that the domain H^*_{all} is focalized in H^+ , and that h_3 is focalized in H^*_{all} .

Dialogue	H^+
U: <i>Is there a bathroom at the Ibis hotel (h_1) and the hotel Lafayette (h_2)?</i>	$H^*_0 = (\mathbf{h_1}, \mathbf{h_2})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2})$
S: <i>No they don't have bathrooms</i>	$H^*_0 = (\mathbf{h_1}, \mathbf{h_2})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2})$
S: <i>But I propose you the Campanile hotel (h_3)</i>	$H^*_0 = (h_1, h_2)$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$
U: <i>Hmm no, how much were the two hotels?</i>	$H^*_0 = (\mathbf{h_1}, \mathbf{h_2})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$
S: <i>The hotel Lafayette is 100 euros, the Ibis hotel is 75 euros</i>	$H^*_1 = (\mathbf{h_2}, \mathbf{h_1})$ $H^*_0 = (\mathbf{h_1}, \mathbf{h_2})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$
U ₁ : <i>Ok, I'll take the second one</i>	$H^*_1 = (\mathbf{h_2}, \mathbf{h_1})$ $H^*_0 = (\mathbf{h_1}, \mathbf{h_2})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$
U ₂ : <i>Ok, I'll take the third one</i> U ₃ : <i>OK, I'll take the other one</i>	$H^*_1 = (\mathbf{h_2}, \mathbf{h_1})$ $H^*_0 = (\mathbf{h_1}, \mathbf{h_2})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$

Figure 4: Example of dialogue (focused domains and referents are in bold)

In order to interpret U_1 , U_2 and U_3 one needs to rely on the previous structuring of H^+ . In U_1 , the previously focalized domain H^*_1 is preferred to be the base for interpreting “*the second one*” because of the order discrimination. This leads to extracting h_1 hence focalizing it in H^*_1 but also in H^*_0 and in H^*_{all} . In U_2 , H^*_1 cannot be the base for interpreting “*the third one*” because no entity could be discriminated this way. Therefore the only suitable domain is H^*_{all} . It is also impossible to interpret “*the other one*” in H^*_1 because of the lack of a focus discrimination between h_1 and h_2 . It is however possible to choose H^*_{all} for the domain of interpretation: the excluded referents h_1 and h_2 are unfocused while h_3 gains focus.

Another example (figure 5) shows that keeping the way the referents are accessed is important in order to have a reliable state of the referential space. Compare the sequences $S_0U_0S_1U_1$ and $S_0U_0S_2U_2$. In the first one the system does not distinguish the referents from each other, and the referential expression “*the second one*” address the hotel Lafayette. In the second one the system answers the question by mentioning the prices of each hotel separately and “*the second one*” address the Campanile hotel. A reason for such phenomenon seems that it is difficult to corefer to the same referent by two different ordinal expressions successively: the extraction of h_3 instead of h_2 in U_1 would sound strange. On the contrary in U_2 , “*the second one*” could refer to h_3 because of the new domain which differentiates the hotels by their prices. Actually the model could predict such behavior by the access structure of H^*_1 introduced in U_0 specifying an ordinal discrimination criterion (noted by a “o:”): if the structure does not change each hotel h_1 or h_2 could be accessed by the ordinal expression they were introduced with. The pronoun “*They*” in S_1 does not change this structure in the same way as S_2 does, that is by increasing the salience of the referents accessed by their names. This way we can constrain the interpretation of ordinals.

Dialogue	H^+
S_0 : I propose you the Ibis hotel (h_1), the hotel Lafayette (h_2) and the Campanile hotel (h_3).	$H^*_0 = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$
U_0 : How many are the first and the third hotel ?	$H^*_1 = \mathbf{o:}(\mathbf{h_1}, \mathbf{h_3})$ $H^*_0 = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$
S_1 : They are expensive.	$H^*_1 = \mathbf{o:}(\mathbf{h_1}, \mathbf{h_3})$ $H^*_0 = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$
U_1 : OK, I'll take the second one.	$H^*_1 = \mathbf{o:}(\mathbf{h_1}, \mathbf{h_3})$ $H^*_0 = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$
S_2 : The Ibis hotel is 100 euros and the Campanile hotel is 50 euros.	$H^*_2 = (\mathbf{h_1}, \mathbf{h_3})$ $H^*_1 = \mathbf{o:}(\mathbf{h_1}, \mathbf{h_3})$ $H^*_0 = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$
U_2 : OK, I'll take the second one.	$H^*_2 = (\mathbf{h_1}, \mathbf{h_3})$ $H^*_1 = \mathbf{o:}(\mathbf{h_1}, \mathbf{h_3})$ $H^*_0 = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$ $H^*_{all} = (\mathbf{h_1}, \mathbf{h_2}, \mathbf{h_3})$

Figure 5: Example of cascading ordinals

5. Discussion

The scope of the groupings considered by this extension to the Reference Domains Theory is still limited. First the trans-sentential groupings are not fully studied yet. We guess that such groupings would need a rhetorical description of the discourse à la SDRT (Asher, 93). Second it considers only extra-linguistic referents, *i.e.* those having an existence outside discourse. When trying to solve references in a dialogue, one should also take into account the domains of interpretation of each speaker. Consider “U: I want an hotel in Paris”, “S: I propose you two hotels”. The first hotel is interpreted in the domain of mental representations of U while the two hotels proposed by S are assumed to exist outside the discourse. They can hardly be grouped together by a referential expression because they belong to different levels of reality. We guess that this kind of phenomenon could be rendered by defining accurate differentiation criteria.

Conclusion

We presented a model of grouping in the Reference Domain Theory. This theory considers that reference resolution is a matter of extracting the referent in an accurate reference domain. It suits well our needs: the groups are considered as reference domains, where any further reference (ordinals or other) can be interpreted. We introduced a particular type of domain, the superdomain, which references all the plural domains constructed at a certain time. Given this domain, the model can render dynamic effects like ordinals or otherness in plural contexts. The conditions for grouping are not examined in detail, however such domain architectures could be the backbone for modelling more complex reference effects using more precise differentiation criteria: time between two evocations of referent, rhetorical or dialogical structure or different mental spaces. This would be future work but in the meantime the model and algorithms are currently evaluated in the MEDIA/EVALDA project which aims to compare the semantic and pragmatic understanding of dialogue systems.

References

- Devillers, L., Maynard, H., Rosset, S., Paroubek, P., McTait, K., Mostefa, D., Choukri, K., Bousquet, C., Charnay, L., Vigouroux, N., Béchet, F., Romary, L., Antoine, J.-Y., Villaneau, J., Vergnes, M., and Goulian, J. (2004). The French MEDIA/EVALDA Project: the Evaluation of the Understanding Capability of Spoken Language Dialog System. In *Proceedings of LREC 2004*, Lisbon, Portugal.
- Eschenbach, C., Habel, C., Herweg, M., Rehkämper, K., (1989). Remarks on plural anaphora. In *Proc. Fourth Conference of the European Chapter of the Association for Computational Linguistics*.
- Kamp, H. and Reyle, U. (1993). *From Discourse to Logic: Introduction to Model-theoretic Semantics of Natural Language. Formal Logic and Discourse Representation Theory*. Kluwer Academic Publisher.
- Landragin, F. and Romary, L. (2003) Referring to Objects Through Sub-Contexts in

Multimodal Human-Computer Interaction. In *Proc. Seventh Workshop on the Semantics and Pragmatics of Dialogue* (DiaBruck'03), Saarbrücken, Germany, 2003, pp. 67-74.

Olson D. (1970). Language and Thought: Aspects of a Cognitive Theory of Semantics. *Psychological Review*, 77/4, 257-273.

Pitel G., Sansonnet J-P. (2003) A Differential Representation of Predicates for Extensional Reference Resolution, In *Proc. of the 2003 International Symposium on Reference Resolution and its Application to Question Answering and Summarization*, Venice, Italia, june 2003

Salmon-alt, S. (2000) Interpreting referring expressions by restructuring context. *Proc. ESSLLI 2000, Student Session*, Birmingham, UK, August 2000.

Salmon-Alt, S. (2001) Reference Resolution within the Framework of Cognitive Grammar. *Proc. International Colloquium on Cognitive Science*, San Sebastian, Spain

Sperber, D. and Wilson, D. (1986) *Relevance, Communication and Cognition*. Basil Blackwell, Oxford.