Differential Representation of Predicates for Extensional Reference Resolution
Guillaume Pitel, Jean-Paul Sansonnet

To cite this version:
Guillaume Pitel, Jean-Paul Sansonnet. Differential Representation of Predicates for Extensional Reference Resolution. 2003. hal-00000911

HAL Id: hal-00000911
https://hal.science/hal-00000911
Preprint submitted on 2 Dec 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A Differential Representation of Predicates for Extensional Reference Resolution

Guillaume Pitel
LIMSI-CNRS
BP 133
F-91403 Orsay Cedex
pitel@limsi.fr

Jean-Paul Sansonnet
LIMSI-CNRS
BP 133
F-91403 Orsay Cedex
jps@limsi.fr

Abstract
In this paper, we focus on a method in practical dialogue for resolving extensional descriptions containing vague or relational predicates as well as predicates on intrinsic properties. It is shown that different kinds of predicates can be handled within a unified approach. This method is built upon the work of Salmon-Alt and is intended to be included into her general resolution model, extending her account for differentiation criteria. We indeed argue about the lack of expressiveness of logical predicates when dealing with vague or context-sensitive linguistic predicates. Our solution is a differential, function-oriented approach to reference resolution, based on the idea that predicates’ meaning may be represented with a comparison function and two partitioning functions. In order to address these problems, we make two propositions. The first is that reference resolution process should primarily rely on the use of a comparison function, not on predicates indicating absolute properties of entities. The second is that functions we use to represent the semantics of referential extractors should be able to take context as an argument.

1 Extensional Reference Resolution in Practical Dialogue
In the framework of practical dialogue, extensional reference resolution is defined as the process of searching for entities referred to by a linguistic description (Byron and Allen, 2001). We have particularly examined the case of descriptions containing so-called vague predicates such as tall or heavy, but also relational predicates such as on the left to and purely intrinsic predicates such as red or square. While there is a large literature about vagueness in the field of linguistics, its application to reference resolution in dialogue systems is almost non-existent. The main reason is that literature about vague predicates focuses on a descriptive in statements, while reference resolution rather requires an extensional interpretative account. Because of this, most extensional reference resolution methods in dialogue systems use a basic logical approach. Within a logical predicative approach, reference resolution is handled by considering all the entities of a model $M$ satisfying the predicates derived from the description. For instance, when a user says “the big red square”, a predicative approach would result in searching for entities satisfying the formula: $\text{big}(x) \land \text{red}(x) \land \text{square}(x)$. When one wants to cope with more complex properties, such as spatial relations, it becomes necessary to make use of specific heuristics (Winograd, 1973).

This paper presents an extension of the reference resolution frame defined in (Salmon-Alt, 2001b). Contrary to Salmon-Alt’s position about differentiation criteria, we adopt a function-
oriented view about the meaning of referential extractors, and examine the idea of using differentiation procedure in place of differentiation criteria. This approach allows suppressing the need for absolute valuation of entities’ properties, and thus can be used for strong qualitative predicates, such as true/false or masculine/feminine as well as quantitative predicates of vague predicates. Splitting the semantic representation of vague predicates into several functions moreover allows dispatching the account for context and helps explaining the processing order of the predicates.

2 Mental Representations and Referential Extractors

Salmon-Alt (2001) examines a model for recruiting contextual entities referred to by referential expressions. Her model is based on the idea of mental representations (Reboul, 1999) used to stand for the dialogue’s context and more specifically reference domains. Reference domains are local contextual sets of entities. They are structured in order to reflect the way they were built, and consequently can be used to predict the distribution of different reference markers. In this model, mental representations introduced by a referential expression can be recursively used as reference domains.

This model however fails in representing relational constraints on entities, such as ‘big’. Indeed, in our opinion, Salmon-Alt fails to distinguish between referential extractors and entities properties. For instance, when processing “the big circle”, she proposes to partition the initial reference domain between big and ¬big. In other words, it separates entities for which property size is big from others.

Referential extractors are the parts of a referential description that constrain properties of the expression’s extension, i.e. the objects1 of the world referred to by the expression. Typically, in a sentence like “Take the blue box”, the word ‘blue’ serves as an extractor for objects having a color that can be characterized as blue. In this sentence, ‘box’ is also a referential extractor, but ‘the’ plays another role. The role of determiners is best described by the term meta-extractor, since they control the way extractors are used in the resolution process. It is also the case of degree words and modifiers. Salmon-Alt’s model has been primarily designed to account for the behavior of different kinds of referential expressions for French, such as demonstrative, definite, indefinite and pronominal forms. In our model, referential extractors are explicitly distinguished from other roles. For instance, in the following sentences, ‘blue’ is not a referential extractor: “Is your car blue?”, “Blue squares are beautiful”, “Add a blue square on the grid”.

2.1 On the vagueness of properties

In practical dialogue systems, semantics of size adjectives like big or small are often reduced to predicates. However as (Dale and Reiter, 1995) notice it seems obvious that the entity referred to by an expression containing an adjective such as ‘big’ or ‘small’ is not supposed to hold a property asserting whether it is objectively big or small.

While it is almost not explored in practical dialogue, vagueness of properties is thoroughly studied in linguistics domain. Various approaches have been proposed, either qualitative (Kamei and Muraki, 1994) or quantitative (Simmons, 1993) or hybrid (Staab and Hahn, 1997). These studies primarily focus on the problems raised by vague predicates in positive, comparative, equative or superlative assertions (e.g. “Paul is slightly taller than Susan”). An issue that seems more important to us for practical dialogue is the problem of resolving reference containing vague predicates.

Whereas literature about vagueness proposes interesting solutions for the logical representation of vague predicates, and for the effect of degree words used in conjunction, it doesn’t help to find applicable solutions for the specific problem of extensional reference resolution.

Pateras et al. (1995) have proposed to deal with reference resolution containing weight predicates with a method based on fuzzy logic. The authors propose to use functions based on fuzzy sets to choose the right referent. However, these functions are defined relatively to a particular task, and are not designed to take ontological or operating context into account.

---

1 Objects are internal representations of the software with which the user want to talk, while entities denote underspecified elements manipulated by the natural language understanding system.
because they have no influence on the task presented in their article. According to us this does not help much for solving our problem, since dynamic context is still not considered. The same problem can be found in (Lammens and Shapiro, 1993) where the authors construct color categorization functions with a learning algorithm, but do not address influences from the context.

An interesting approach is drawn in (Staab and Hahn, 1997), which serves to compute the comparison class with which the vague predicate is applicable. The proposed algorithm finds the concept (a subpart of the ontological context) within which the predicate must be applied. The problem with this approach is that the comparison class can not be defined with the dynamic knowledge from the operating context but only with the ontological knowledge.

In order to address these problems, we make two propositions. The first is that reference resolution process should primarily rely on the use of a comparison function, not on predicates indicating absolute properties of entities. The second is that functions we use to represent the semantics of referential extractors should be able to take context as an argument.

2.2 Properties, Relations and Types

Dale and Reiter (1995) distinguish three kinds of referential extractors: those involving typological properties (e.g. being a figure, being a square…), those involving intrinsic properties (e.g. size, color) and those involving relational properties (e.g. spatial position, temporal relations).

Most dialogue system’s models assume access to an adequate representation of world entities, coherent with the way the user refers to them in natural language. This can be observed even in recent work such as (Salmon-Alt, 2001a:149), (Byron et al., 2001) or (Dale et Reiter, 1995) with usage of predicates like (color x RED) or (size x SMALL) for selecting red and small objects. Such models assume that referential extractors purely rely on intrinsic properties of objects to designate entities. Consequently, dialogue systems cannot propose a unified account for relational constraints, leading to ad hoc approaches to relational references resolution.

We make the hypothesis that these three kinds of extractor can be represented in a unified way using a comparative or differential approach. Indeed, many intrinsic properties are actually relational, which is shown by their use in comparative clauses. For instance, one can say “Your shirt is redder than mine”, or “Your boyfriend is more masculine than mine”. The main hypothesis sustaining our proposal is that even strong predicates are actually relational predicates applied on the target entity against a default, prototypical entity. For instance a prototypical male for masculine, or a prototypical color for red.

3 Differential Representation

We propose to represent referential extractors with a structure of functions that is to be used by a generic resolution algorithm. For the purpose of illustrating our view, we define a situation described in FIG. 1. In order to focus the discussion on the issue of resolving intrinsic references, we will consider that elements of the mediated software are objects (in the object-oriented algorithmic sense). These objects are defined by a type and two attributes: color (a 3-uple of values in (0, 1) and a size (a value in (0,∞)).

FIG. 1: Experiment case for procedural representation of size and color.

3.1 Functional Representation of Referential extractors

We represent referential extractors as a structure of three functions. These functions are detailed in FIG. 2. The fSimil comparison function is used to sort a set of objects among a given property (say, blueness or bigness). For instance, the fSimil function of the referential extractor ‘blue’ give the similarity ratio between the col-
ors of two objects projected along the blue axis\(^2\). If \(a\) is bluer than \(b\), then the returned value is above 1, while the contrary returns a value between 0 and 1 (exclusive), and equality returns 1. If one of the two colors is outside the area considered to be the limit of blue colors, returned value is \(\perp\).

The \(f_{\text{Excl}}\) function serves to select in the sorted list (produced by the resolution algorithm using \(f_{\text{Simil}}\) function) objects that will be excluded from the candidates list that is going to be submitted to the next extractor. All extractors do not select the same way, for instance extractor \(\text{big}\) will select objects of the same size as the littlest one (that is, the last one in the list). For any extractor for \(\text{color}\), this function selects all objects with similarity ratio \(\perp\).

The \(f_{\text{Pref}}\) function selects the preferred objects for a given extractor. Here again, there is no general rule for the function. The extractor \(\text{big}\) could select it based on any heuristic (either the 30% upper sizes, or all the objects from the beginning to the first point where the secondary differential coefficient is over a given threshold).

The resolution algorithm is triggered during the natural language analysis process when an extensional referring expression has been isolated. The rules leading the algorithm follow the guidelines defined in (Salmon-Alt, 2001b), so that the type (definite, indefinite, demonstrative, pronominal) of the referring expression determines the way the initial reference domain is built, and how the extractors are used to return the final result of the reference resolution in the restructuring phase.

We consider each referential extractor to be a sub-process of the following form\(^3\):

1. Transform each entity of the original reference domain through\(^4\) a point of view that would make the extractor able to process the domain. The initial reference domain may be sorted, partitioned and focused by a preliminary step.
2. Use the \(f_{\text{Simil}}\) comparison function of the extractor in order to sort the domain by the adequate criteria.
3. If the produced reference domain is to be passed to another extractor, or the referring expression is in plural form, use the \(f_{\text{Excl}}\) function to partition the domain among possible and impossible candidates.
4. Use the \(f_{\text{Pref}}\) function to extract best candidates if the referring expression is in plural form.

Compared to Salmon-Alt’s differentiation criteria, our algorithm makes use of two different partition functions (\(f_{\text{Excl}}\) and \(f_{\text{Pref}}\)) in order to produce a ternary partition instead of a binary one. Indeed, as we introduce a total order relation in the differentiation process, we are faced with an issue that was hidden when using simple predicates: when referring to several entities, like in “remove blue squares”, the user refers to several objects at the top of the list sorted by blueness; however, squares less blue but still blue must be distinguished from non-blue squares, because if the user asks for “remove big blue squares”, and big squares are not at the top of the list sorted by blueness, they must be passed from ‘blue’ extractor to ‘big’ extractor in the list of possible candidates for the referring expression. So it is necessary to distinguish between best candidates, needed to answer to group references, and possible and impossible candidates in order to respond “nothing matches” to ambiguous referential expressions.

\(^2\) Do not mistake it for the blue component of a red, green and blue decomposition of color. The blue axis or dimension represents the distance (in a cognitive view) of any color compared to blue. For instance, sky blue or indigo are farer from blue than marine blue. Note that there is no measurable distance between green and yellow on the blue axis, and thus not all colors can be projected on this axis.

\(^3\) The full algorithm is not detailed here, because of its length. Moreover its details still have to be checked against results of not yet finished psychological experiments.

\(^4\) The transformation through a point of view changes all objects of the domain into objects of a class suitable for the \(f_{\text{Simil}}\) comparison function. This point is crucial, but out the scope of this paper, and will not be discussed further.
3.3 A practical situation

From the situation presented in FIG. 1, we illustrate how to use differentiation functions in the case of resolving the referring sub-expression “big blue squares”.

We take as a starting point for reference resolution the referential domain \( RD_{\text{initial}} \) of all the squares of the scene. The referential domain contains labels of objects. In order to shorten the example, we will not detail the ‘squares’ extractor, since all objects of the scene are squares. We first apply the referential extractor for ‘blue’.5 Theoretically, this function will produce when applied to a set of \( n \) objects, a matrix of \( n \times n \) elements. For all practical purpose, one can consider that in many cases, \( f_{\text{Simil}}(x, z) = f_{\text{Simil}}(x, y) \times f_{\text{Simil}}(y, z) \), providing a full order relation. This point could be subject to discussion, but we will consider that it is a reasonable assumption. Algorithms for sorting sets with a full order relation have complexity of order \( O(n \log n) \). Using the \( f_{\text{Simil}} \) function, we construct an ordered list (\( \text{Ord}_{\text{blue}} \)) whose links between objects are annotated with similarity ratios (\( \text{Sim}_{\text{blue}} \)), and partition it with the two functions \( f_{\text{Excl}_{\text{blue}}} \) and \( f_{\text{Pref}_{\text{blue}}} \).

\[
\begin{align*}
\text{RD}_{\text{initial}} &= \{1, 2, 3, 4, 5, 6, 7, 8\} \\
\text{Sim}_{\text{blue}}(\text{RD}_{\text{initial}}) &= \{(5, \text{protoblue}), 0.98\}, \{(1, \text{protoblue}), 0.9\}, \{(3, \text{protoblue}), 0.6\}, \{(4, \text{protoblue}), 0.55\}, \{(6, \text{protoblue}), \perp\}, \{(7, \text{protoblue}), \perp\}, \{(8, \text{protoblue}), \perp\} \\
\text{Ord}_{\text{blue}}(\text{RD}_{\text{initial}}) &= 5 > 1 \times 3 > 4 \times x \forall x \in \{2, 6, 7, 8\} \\
\text{Pref}_{\text{blue}}(\text{RD}_{\text{initial}}) &= \{5 > 1\} \\
\text{Poss}_{\text{blue}}(\text{RD}_{\text{initial}}) &= \{5 > 1 \times 3 \times 4\} \\
\text{Excl}_{\text{blue}}(\text{RD}_{\text{initial}}) &= \{2, 6, 7, 8\}
\end{align*}
\]

As \( \text{card}(\text{Pref}_{\text{blue}}(\text{RD}_{\text{initial}})) \) is only 2, and the referring expression is in plural form, big must at least take the context from \( \text{Poss}_{\text{blue}}(\text{RD}_{\text{initial}}) \) for argument (because it must manage to make a partition).

\[
\begin{align*}
\text{Sim}_{\text{big}}(\text{Poss}_{\text{blue}}(\text{RD}_{\text{initial}})) &= \{(3, 4), 1\}, \{(4, 1), 1\}, \{(1, 5), 0.75\}
\end{align*}
\]

Since the similarity ratio is too high between square 1 and square 5, the partitioning is impossible. The context must be enlarged. The extraction is then applied on the whole original domain.

\[
\begin{align*}
\text{Sim}_{\text{big}}(\text{RD}_{\text{initial}}) &= \{(3, 4), 1\}, \{(4, 1), 1\}, \{(1, 5), 0.75\}, \{(2, 3), 0.9\}, \{(5, 6), 0.25\}, \{(6, 7), 1\}, \{(7, 8), 1\}\} \\
\text{Poss}_{\text{big}}(\text{RD}_{\text{initial}}) &= \{2 \times 1 \times 3 \times 4 \times 5\} \\
\text{Excl}_{\text{big}}(\text{RD}_{\text{initial}}) &= \{6, 7, 8\}
\end{align*}
\]

As the partitioning succeeded, one can compute the intersection between \( \text{Pref}_{\text{blue}}(\text{RD}_{\text{initial}}) \) and \( \text{Poss}_{\text{big}}(\text{RD}_{\text{initial}}) \). The result of reference resolution for “big blue squares”, in the situation of FIG. 1 produces squares number 1 and 5 as the best candidates. Same method applied to “small blue square” produces the square number 5. In these two cases, the result seems to corroborate the basic intuition and the human validation6.

3.4 Arranging extraction operations

The processing order of referential extractors is of great importance in extensional resolution. Salmon-Alt’s model rely on the propositions that (Dale and Reiter, 1995) have made about the relative importance of objects’ properties in the framework of referential expressions generation. The authors mainly base their research on the implication of Grice’s maxims (Grice, 1975) on reference generation, following the principle of maximum economy in language generation. Authors propose to take properties into account in the following order: type > intrinsic property > relational property. That means that, when a user refers to an entity, he first uses the type of the entity to describe it, then if it is not sufficient to distinguish the entity from others, he use an intrinsic property, and finally a relational property. For instance, if (1) “The red square” and (2) “The square on the right” refers to the same object, (1) will be used, since it contains only an intrinsic property.

Adapting this approach from generation to interpretation leads to consider that in a referring expression, extractors (mainly nouns and adjectives) must be processed in the same order, first the nouns, then intrinsic adjectives, then relational properties. As a consequence, the interpretation of “the big blue square” begins with the selection of squares giving a set of entities \( S_1 \), then blue objects from the set \( S_1 \) giving the set

---

5 We address in section 3.4 the issue of extractors’ processing order.

6 We have made a first experiment with a dozen of volunteers. Participants have been faced with several situations like the one described in FIG. 1. They were then asked to select on paper the objects denoted by a given sentence, for instance “the big squares” or “the blue squares”.
S_2, and finally the big objects from the set S_3, giving the set S_3 used to find the best candidate for the reference. It is obvious that choosing another order will lead to different results, incompatible with the natural interpretation.

Our model of context handling provides a rationale for the empirical observation of resolution order. Indeed, while the calculation of the subset referred to by a type is computationally cheap, the calculation of the subset of entities referred to by a relational adjective is very expensive. Moreover, the more the predicate’s meaning is potentially influenced by context, the more probable is the fact that context must be included in the computation in order to ensure that the reference domain is correct.

The cost of processing an extractor is related to the quantity of contextual information that is necessary to take into consideration. As the context is restricted by each extractor, low-cost extractors must be processed first. We propose the following principle:

*The more the operating context is important to process an extractor, the later this extractor should be put in the resolution chain.*

As the perception of colors is only a little dependent on the situation, whereas the perception of the size is highly related to the other objects present in the scene, the arrangement of extractor execution is *color* then *size*. This approach offer the advantage to explain also why the type (the noun) is always the most important of the properties and is always considered before other ones. So the expression “*the big blue square*” is not disambiguated as: “*the squares among the blue objects among the big objects*”, but as: “*the big ones among the blue ones among the squares*”.

4 Conclusion

We have made two propositions:

- That all kinds of referential extractors should be accounted within a differential approach.
- That functions used to represent referential extractors take the context as arguments.

We also have proposed the principle that extractors processing order in a referential chain should be computed from the amount of contextual information needed to process each extractor. Those propositions are made to be integrated into a resolution model which follows Salmon-Alt’s guidelines. This approach aims to cover relational and typological extractors as well, but there is still some important work to be done in order to reach this goal. One of the issues raised by this extension of the preliminary model is the need for a mechanism to measure the distance between two objects in any dimension. For typological properties, this implies to perform a metaphorical transformation of reference domains from one dimension to another.

Our team is currently developing this mechanism inside the InterViews project (Sansonnet et al., 2002). This project is built around the concept of conversational agents for assistance to ordinary people. We aim to provide a platform for helping design of natural language interface to common software. The formalism presented in this paper, as well as the overall model of natural language analysis integrating it, is under development in this framework. There is consequently no evaluation of this system so far, but we plan to reach a working platform in order to be able to compare results from our reference resolution method with the results of psychological experimentations.

References


