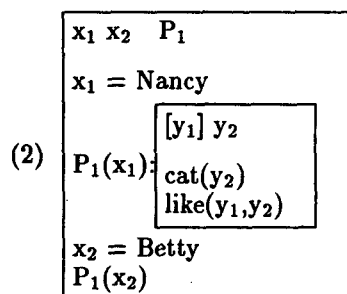# VP Ellipsis in a DRT-implementation

Johan Bos
Department of Computational Linguistics,
Faculty of Arts, University of Groningen,
P.O.Box 716, 9700 AS Groningen.
Email: s0615838@let.rug.nl

1 Introduction

Klein [Klein, 1986] introduced Predicate-DRSs for the resolution of VP Ellipsis. In that approach a Predicate-DRS (henceforth PDRS) serves as the representation of a verb phrase, as will be shown in an example now. Consider:

(1) Nancy likes a cat.
    Betty does too.

This discourse is interpreted as meaning that Nancy and Betty both like a cat (though not necessarily the same cat). The source clause, *Nancy likes a cat*, parallels the target clause *Betty does too*, where the subjects are parallel elements. The phrase *does too* represents a trace of the VP in the target clause. Klein's treatment of (1) is shown in (2).

(2)
$$
\boxed{
\begin{array}{l}
x_1 \ x_2 \quad P_1 \\
x_1 = \text{Nancy} \\
P_1(x_1) : \boxed{
\begin{array}{l}
[y_1] \ y_2 \\
\text{cat}(y_2) \\
\text{like}(y_1, y_2)
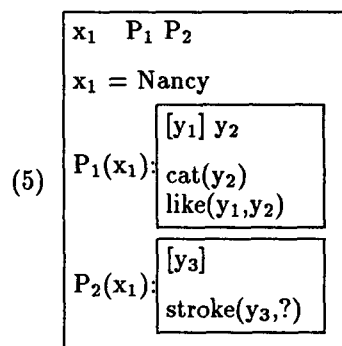\end{array}} \\
x_2 = \text{Betty} \\
P_1(x_2)
\end{array}}
$$

In the second sentence of (1), a do-anaphor appears that must be linked to a marker which has already been introduced into the universe of the DRS. The value of this marker, which is P_1, as we can see in (2), is constrained by the conditions associated with the previous VP in the discourse [Klein, 1986]. Following Klein, we call P_1 a *predicate marker*, and the Sub-DRS that is associated with P_1 a *Predicate-DRS*. To the domain of P_1, a distinguished reference marker y_1 (indicated by square brackets) is added, which plays the role of the individual, in this case x_1 which is applied to the predicate. This application can also be shown as a lambda expression:

(3) $\lambda \ y_1 \ (\text{cat}(y_2) \wedge \text{like} \ (y_1, y_2)) \ (x_1)$

In (2) the condition $P_2(x_2)$ in the main DRS will apply the object x_2 to the predicate and solve the do-anaphor in (1). The scope of marker y_2 is defined by the PDRS, instead of the main DRS, which allows that Nancy and Betty do not necessarily like the same cat.

But this same feature introduces a problem for pronoun resolution. This problem occurs when pronouns refer to indefinite NPs which are in the universe of a PDRS and therefore inaccessible. Let us give an example by considering the DRS (5) as the translation of (4).

(4) Nancy$_1$ likes a cat$_2$.
    She$_1$ strokes it$_2$.

(5)
$$
\boxed{
\begin{array}{l}
x_1 \quad P_1 \ P_2 \\
x_1 = \text{Nancy} \\
P_1(x_1) : \boxed{
\begin{array}{l}
[y_1] \ y_2 \\
\text{cat}(y_2) \\
\text{like}(y_1, y_2)
\end{array}} \\
P_2(x_1) : \boxed{
\begin{array}{l}
[y_3] \\
\text{stroke}(y_3, ?)
\end{array}}
\end{array}}
$$

Since, in DRT, an anaphor can only refer to antecedents from its own domain or from universes that its DRS subordinates, the pronoun *it* cannot be anaphorically linked to the indefinite description *a cat*. This means, in the situation of (5), candidate antecedents for *it* can only be found in the main DRS, since P_2 is subordinated to it. The desirable antecedent y_2 in P_1 is blocked.[1]
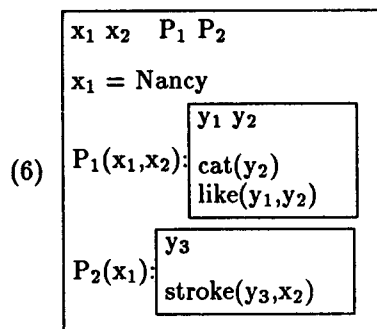
A solution to the problem of the indefinite descriptions appearing in PDRSs, is to make them accessible in the main DRS. This paper shows, by slightly modifying Klein's PDRSs, how that can be done, without losing their desirable characteristics.

Firstly, we outline informally how indefinite descriptions in PDRSs are made accessible. Then we show how this idea relates to aspects like negation, disjunction, quantification and the strict/sloppy identity of VP Ellipsis. Finally, we report about the implementation under development.

---

[1] Notice, that proper names and definite descriptions do not give rise to this problem. In DRT, these are usually added to the universe of the main DRS [Kamp and Reyle, 1990] or accommodated to it [Van der Sandt, 1992].

425

2 A new approach to Predicate-DRSs

By treating a PDRS just as an ordinary DRS, with the distinction that there is a correspondence between the arguments which are applied to the PDRS and the members of its domain, it is possible to extend the scope of reference markers in a PDRS to their superordinated DRS. The best way to show how this works is to look at a DRS for (4) in this new approach:

(6)

$$
\boxed{
\begin{array}{l}
x_1\ x_2\quad P_1\ P_2 \\[4pt]
x_1 = \text{Nancy} \\[4pt]
P_1(x_1,x_2)\!:\!\boxed{\begin{array}{l} y_1\ y_2 \\ \text{cat}(y_2) \\ \text{like}(y_1,y_2) \end{array}} \\[10pt]
P_2(x_1)\!:\!\boxed{\begin{array}{l} y_3 \\ \text{stroke}(y_3,x_2) \end{array}}
\end{array}}
$$

In (6), in the PDRS P_1, y_1 is linked to x_1, and y_2 to x_2. So, the difference here to Klein's approach is that, besides the referent for the individual which is applied to the PDRS, *all* indefinite descriptions in the universe of the PDRS are associated with corresponding arguments as well.[2] A lambda expression for P_1 in (6) is:

(7) $\lambda\, y_1\ \lambda\, y_2\ (\text{cat}(y_2) \wedge \text{like}(y_1,y_2))\ (x_1)\ (x_2)$

This treatment allows that we can refer to the indefinite *cat*, as is done in P_2 of (6). An added advantage is that we maintain the original properties of a PDRS outlined previously. Note, that the number of arguments applied to a PDRS directly depends on the number of indefinite descriptions in the VP. Consequently, a VP with a ditransitive verb could yield two indefinite descriptions, as in (8). Optional relative clauses can raise this number even higher (9).

(8) Nancy gives a man an iron.

(9) Nancy likes a man who has an iron that a woman gave him.

3 Negation

Concerning predicate negation, we will assume that the scope of negation does not embrace the subject (cf [Kamp and Reyle, 1990]). The approach we take
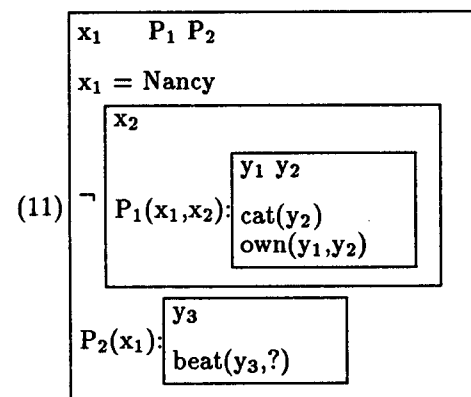
---

[2]Therefore, it is not necessary to distinguish them with square brackets any more. Note that the agent corresponds to the first referent in the PDRS.

here is similar to standard DRT, because a new subordinated DRS affixed with a negation symbol is introduced in case of negation. Let us consider (10):

(10) Nancy$_1$ doesn't own a cat$_2$.
* She$_1$ beats it$_2$.

Here we simply negate the predicate by constructing the PDRS in a negated DRS. In (10) the pronoun *it* does not permit a link to the NP *a cat*, and this seems to be the case in general as well, because negation blocks anaphoric links.[3]

Thus, in the case of a negated VP, the indefinites are raised to the superordinated DRS which is the DRS for negation. This construction is figured in DRS (11) and causes exactly the result we wish: *it* cannot be linked to *cat* because the referent for *cat*, x_2, is not accessible.

(11)

$$
\boxed{
\begin{array}{l}
x_1\quad P_1\ P_2 \\[4pt]
x_1 = \text{Nancy} \\[4pt]
\neg\ \boxed{\begin{array}{l} x_2 \\ P_1(x_1,x_2)\!:\!\boxed{\begin{array}{l} y_1\ y_2 \\ \text{cat}(y_2) \\ \text{own}(y_1,y_2) \end{array}} \end{array}} \\[14pt]
P_2(x_1)\!:\!\boxed{\begin{array}{l} y_3 \\ \text{beat}(y_3,?) \end{array}}
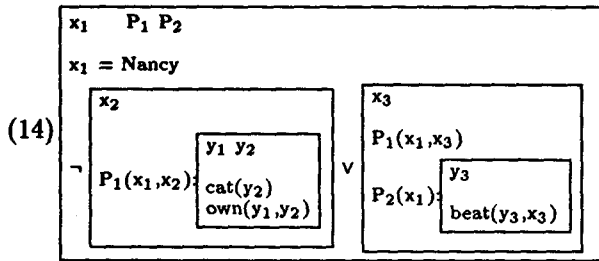\end{array}}
$$

Now consider (12), where an anaphoric link between *cat* and *it* is permitted. At first sight, this sentence would appear as a counterexample to our characterization of negation. But it is not, if we interpret the meaning of it as (13):

(12) Either Nancy doesn't own a cat, or she beats it.

(13) Either Nancy doesn't own a cat, or she does and she beats it.

An interpretation of (12) as (13) permits the accessibility of cat in (12). In our DRT-framework with PDRSs we easily can obtain a DRS for (12), as the disjunction of two SubDRSs. Then, in one disjunct predicate negation takes place, while in the other the

---

[3]However, [Kamp and Reyle, 1990] give as a possible counterexample to this generalization the discourse *Jones does not like a Porsche. He owns it*, interpreting it by saying that there is some Porsche that Jones both dislikes and owns. According to me, such an interpretation seems only permitted if that Porsche is already uttered in the processed discourse.

do-anaphor is resolved, resulting in a accessibility for the indefinite NP *a cat*.
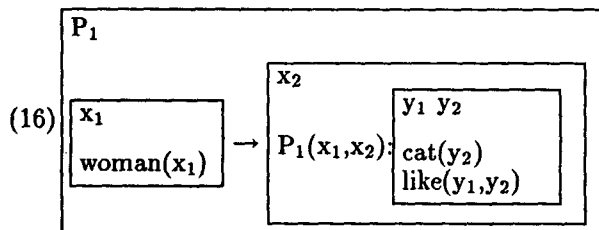
(14)
```
┌─────────────────────────────────────────────────┐
│ x₁  P₁ P₂                                         │
│ x₁ = Nancy                                        │
│   ┌──────────────────────┐   ┌─────────────────┐ │
│   │ x₂                    │   │ x₃              │ │
│   │           ┌────────┐  │   │ P₁(x₁,x₃)       │ │
│   │           │ y₁ y₂  │  │ V │        ┌──────┐ │ │
│ ¬ │ P₁(x₁,x₂):│ cat(y₂)│  │   │ P₂(x₁):│ y₃   │ │ │
│   │           │ own(y₁,y₂)│ │   │        │beat(y₃,x₃)│ │
│   │           └────────┘  │   │        └──────┘ │ │
│   └──────────────────────┘   └─────────────────┘ │
└─────────────────────────────────────────────────┘
```

4 Quantification

In this section we will see how the quantifiers *every* and *no* can be treated. We will demonstrate how quantification matches perfectly with our proposals about PDRSs and negation. Sentence (15)

(15) Every woman likes a cat.

involves applying the quantified NP *every woman* to the PDRS, visualized in DRS (16):

(16)
```
┌─────────────────────────────────────────────────┐
│ P₁                                                │
│   ┌──────────────┐     ┌──────────────────────┐  │
│   │ x₁           │     │ x₂                    │  │
│   │              │     │           ┌────────┐  │  │
│   │              │  →  │           │ y₁ y₂  │  │  │
│   │ woman(x₁)    │     │ P₁(x₁,x₂):│ cat(y₂)│  │  │
│   │              │     │           │like(y₁,y₂)│ │  │
│   └──────────────┘     │           └────────┘  │  │
│                        └──────────────────────┘  │
└─────────────────────────────────────────────────┘
```

Of interest here is that the argument of P_1 is the member of the antecedent DRS: x_1. Also worth noting is that the referent of the indefinite *a cat* in P_1 is not raised to the main DRS but to the DRS that holds the consequent of the implication relation. In this case the NP *a cat* has narrow scope within the quantified phrase *every woman*, and therefore not accessible in the main DRS (as in standard DRT).

In a similar way the quantifier *no* is interpreted, using the logical equivalence of the formulae (17) and (18):
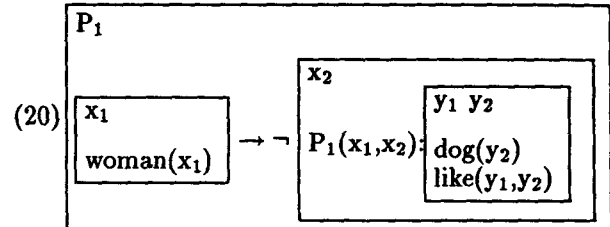
(17) $\neg \exists x \; P(x) \land Q(x)$

(18) $\forall x \; P(x) \rightarrow \neg Q(x)$

The traditional way to translate *no* in DRT is based on (17).[4] In this framework we use predicate nega-
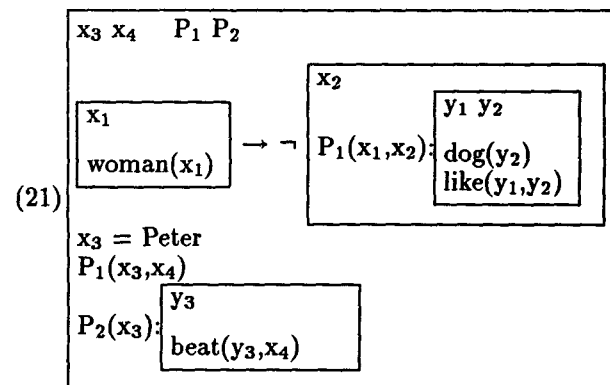
---

[4]Several proposals have been made to treat generalized quantifiers in DRT. Among them: [Klein, 1986; Kamp and Reyle, 1990; Zeevat, 1991].

---

tion combined with universal quantification, shown in (20), which is the translation for (19).

(19) No woman likes a dog.

(20)
```
┌─────────────────────────────────────────────────┐
│ P₁                                                │
│   ┌──────────────┐       ┌──────────────────────┐│
│   │ x₁           │       │ x₂                    ││
│   │              │       │           ┌────────┐  ││
│   │              │ → ¬   │           │ y₁ y₂  │  ││
│   │ woman(x₁)    │       │ P₁(x₁,x₂):│ dog(y₂)│  ││
│   │              │       │           │like(y₁,y₂)│││
│   └──────────────┘       │           └────────┘  ││
│                          └──────────────────────┘│
└─────────────────────────────────────────────────┘
```

This way of dealing with quantification is exactly what we need for VP Ellipsis resolution. A discourse as in (19) could proceed with a sentence like: *But Peter does, and he beats it*, which is an example of a 'missing antecedent' [Hankamer and Sag, 1976], since the pronoun *it* lacks an overt antecedent because the NP *a dog* is in the scope of negation and therefore not accessible. By generating a condition in DRS (20) applying *Peter* to the PDRS P_1, the 'missing' antecedent is found (21).

(21)
```
┌──────────────────────────────────────────────────┐
│ x₃ x₄   P₁ P₂                                      │
│   ┌──────────────┐       ┌──────────────────────┐ │
│   │ x₁           │       │ x₂                    │ │
│   │              │       │           ┌────────┐  │ │
│   │              │ → ¬   │           │ y₁ y₂  │  │ │
│   │ woman(x₁)    │       │ P₁(x₁,x₂):│ dog(y₂)│  │ │
│   │              │       │           │like(y₁,y₂)│ │ │
│   └──────────────┘       │           └────────┘  │ │
│                          └──────────────────────┘ │
│ x₃ = Peter                                         │
│ P₁(x₃,x₄)                                          │
│         ┌──────────────┐                           │
│ P₂(x₃): │ y₃           │                           │
│         │ beat(y₃,x₄)  │                           │
│         └──────────────┘                           │
└──────────────────────────────────────────────────┘
```

Summarizing so far, we have shown that PDRSs, with the ability to raise indefinite descriptions to its superordinated DRS, can be used quite effectively in our framework. Mainly, we distinguished two cases where referents of indefinite descriptions were not raised to the main DRS, but to a DRS subordinated to the top level. The first case concerns predicate negation, where a negated DRS is superordinated to the PDRS involved. The second case concerns quantification, where the PDRS is subordinated to the consequent-DRS of the implication relation.
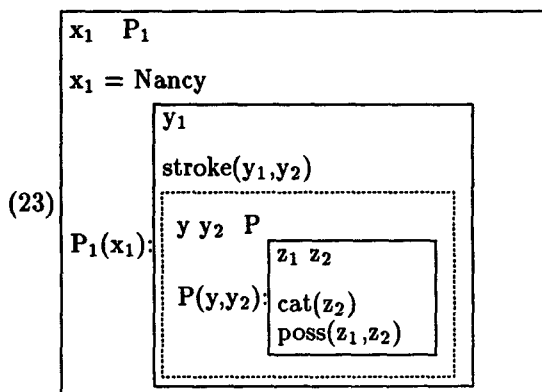
5 Strict and Sloppy Readings

This section shows how sloppy and strict readings arising in VP Ellipsis are obtained. Discourses like (22) are ambiguous as to whether Betty strokes
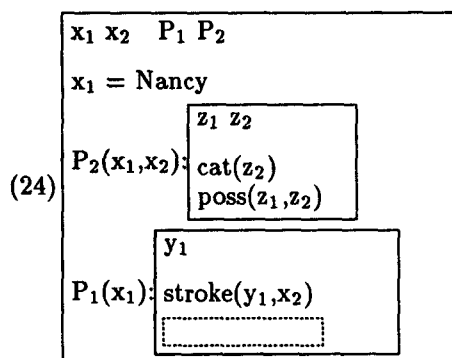
Nancy's cat (the strict reading) or Betty strokes Betty's cat (the sloppy reading).

(22) Nancy strokes her cat.
Betty does too.

Following [Van der Sandt, 1992], presuppositions are *accommodated* to the preceding discourse. That is, if discourse does not provide an antecedent, one will be created. In processing the first sentence of (22), DRS (23) is obtained, where the presuppositional possessive construction *her cat* is paraphrased in a dashed DRS to indicate information for accommodation.

$$
(23)\quad
\boxed{
\begin{array}{l}
x_1 \quad P_1 \\[2pt]
x_1 = \text{Nancy} \\[4pt]
P_1(x_1): \boxed{
\begin{array}{l}
y_1 \\[2pt]
\text{stroke}(y_1,y_2) \\[2pt]
\boxed{
\begin{array}{l}
y\ y_2\quad P \\[4pt]
P(y,y_2): \boxed{
\begin{array}{l}
z_1\ z_2 \\
\text{cat}(z_2) \\
\text{poss}(z_1,z_2)
\end{array}}
\end{array}}
\end{array}}
\end{array}}
$$

In the approach of [Van der Sandt, 1992] the anaphoric material in the dashed DRS is resolved after merging the DRS constructed for the sentence with the main DRS, resulting in a new DRS that contains no anaphoric material for accommodation still to be processed. This procedure is followed for (23) yielding DRS (24).

$$
(24)\quad
\boxed{
\begin{array}{l}
x_1\ x_2 \quad P_1\ P_2 \\[2pt]
x_1 = \text{Nancy} \\[4pt]
P_2(x_1,x_2): \boxed{
\begin{array}{l}
z_1\ z_2 \\
\text{cat}(z_2) \\
\text{poss}(z_1,z_2)
\end{array}} \\[8pt]
P_1(x_1): \boxed{
\begin{array}{l}
y_1 \\
\text{stroke}(y_1,x_2) \\
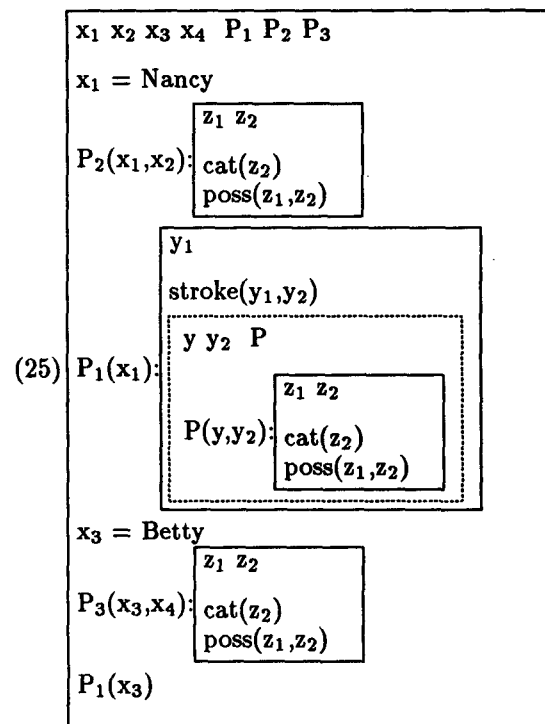\boxed{}
\end{array}}
\end{array}}
$$

Discourse (22) provides one suitable antecedent for the possessor, namely *Nancy*, and *Nancy possesses a cat* is established in the DRS. But this gives us only the strict reading when in case of an elliptical VP in the proceeding discourse is referred to P_1, which is the case in (22).

To represent the sloppy reading, the anaphoric material in (23) that holds the presupposition must not

be resolved at the stage of DRS-merging, but left there to provide accommodation another time (with other constraints, that depend on the antecedent of the possessor). In this way both the strict and sloppy reading are obtainable in case of VP Ellipsis.

We show this proposal with our example (22), corresponding with DRS (25). Similar to (24), the presupposition causes an antecedent to be created (i.e. *Nancy possesses a cat*), with this difference, that the anaphoric material is not resolved. The VP-anaphor finds as an antecedent P_1: *strokes her cat*. The presuppositional material in the dashed DRS can now be accommodated to two different antecedents: Firstly, *Nancy*, where no antecedent has to be created for the possessive construction, resulting in the strict reading. Secondly, the newly introduced *Betty*, where in that case the presupposition *Betty possesses a cat* is accommodated and the sloppy reading can be derived. The latter is shown in (25):

$$
(25)\quad
\boxed{
\begin{array}{l}
x_1\ x_2\ x_3\ x_4 \quad P_1\ P_2\ P_3 \\[2pt]
x_1 = \text{Nancy} \\[4pt]
P_2(x_1,x_2): \boxed{
\begin{array}{l}
z_1\ z_2 \\
\text{cat}(z_2) \\
\text{poss}(z_1,z_2)
\end{array}} \\[8pt]
P_1(x_1): \boxed{
\begin{array}{l}
y_1 \\
\text{stroke}(y_1,y_2) \\
\boxed{
\begin{array}{l}
y\ y_2\quad P \\[4pt]
P(y,y_2): \boxed{
\begin{array}{l}
z_1\ z_2 \\
\text{cat}(z_2) \\
\text{poss}(z_1,z_2)
\end{array}}
\end{array}}
\end{array}} \\[8pt]
x_3 = \text{Betty} \\[4pt]
P_3(x_3,x_4): \boxed{
\begin{array}{l}
z_1\ z_2 \\
\text{cat}(z_2) \\
\text{poss}(z_1,z_2)
\end{array}} \\[6pt]
P_1(x_3)
\end{array}}
$$

If we compare this approach to the higher-order unification approach to VP Ellipsis of [Dalrymple *et al.*, 1991], we can obtain all six readings of the complicated (26) generated by the equational analysis of [Dalrymple *et al.*, 1991].

(26) John revised his paper before the teacher did, and Bill did too.

The reading of (26) where John, the teacher, and Bill all revised John's paper, is translated in a DRS with the presupposition that *John possesses a paper*

accommodated to the main DRS. The reading where John and Bill revised their own papers before the teacher revised John's paper, causes accommodation twice, once for *John possesses a paper* and once for *Bill possesses a paper*. The other readings can be obtained analogously.

6 Implementation

The PROLOG-implementation is a natural language processing system which parses simple discourses. The way DRSs are constructed in this system will be discussed concisely.

The emphasis of the implementation lies on anaphora resolution (like do-anaphora and pronouns) in a domain of a small fragment of English. A parse of a typical discourse is:

```
> Mary likes a cat.
> She does not beat it.
> John does not either.

drs: [ x1 x3 x6 p2 p5 ]
      [ x1 = mary
        p2(x1,x3):[ y x4 ]
                  [ cat(x4)
                    like(y,x4) ]
        not [ ]
            [ p5(x1):[ y ]
                     [ beat(y,x3) ] ]
        x6 = john
        not [ ]
            [ p5(x6) ] ]
```

This implementation differs from other PROLOG-implementations of DRT (e.g. the threading account of [Johnson and Klein, 1986]) in the way it constructs DRSs. Following [Asher, 1990], DRSs are constructed in a bottom-up fashion, using λ-conversion.

Each lexical entry is associated with a SubDRS, representing the meaning of that entry. For instance, the lexical entries for *a*, *man*, and *runs* are:

```
lex(a,
    det:[agr=sing,
         def=ind,
         drs=(X^P)^(X^Q)^drs([],[P,Q])]).

lex(man,
    noun:[agr=sing,
          drs=X^drs([X],[man(X)]),
          gender=male,
          ref=X]).
```

```
lex(runs,
    iv:[agr=sing,
        drs=X^drs([],[do(P,X):drs([y],[run(y)])]),
        ref=P]).
```

As these entries make clear, a DRS is constructed of a PROLOG term containing two lists, where the first one contains the discourse markers (i.e. the domain) and the second one the constraints (these are represented as PROLOG terms). Furthermore, the lambda abstractor is constructed as the PROLOG operator '^' (this idea is taken from [Pereira and Shieber, 1987]).

While parsing a sentence, the DRS for that sentence is processed by λ-conversion and merging, using syntax rules of the following form[5] (as in [Alshawi, 1992]):

```
np:[drs=Drs,agr=Agr,...] --->
    [det:[drs=A2^Drs,...],
     noun:[drs=A1,agr=Agr,...],
     optrel:[drs=A1^A2,...]].
```

The output of a sentence parse is a constructed DRS for that sentence, but with referring expressions (if any) still unresolved. This sentence-DRS then is merged with the ingoing DRS, representing the computed discourse so far. During this merge, the following computing actions take place:

- Computing of arguments for PDRSs;
- Resolving of Pronouns and VP Ellipsis;
- Accommodation of Proper Names, Definite Descriptions, and Possessive Constructions.

An aid to these computations is a historylist computed during the sentence parse. This historylist contains all the items that are represented in the discourse, extended with information that is not purely semantic, such as type and gender of certain subjects, but necessary for the computations mentioned above.

This results in a new DRS, capturing the entire discourse, which will be the ingoing DRS for the merge after the next sentence is parsed.

7 Conclusion

By slightly changing Klein's treatment of Predicate-DRSs, that is making indefinite descriptions occurring in the scope of the VP accessible to the top level of the main DRS, we obtain a much better mechanism for handling VP Ellipsis in DRT without losing any old characteristics in the theory. Furthermore, we proposed to use Van der Sandt's theory on presuppositions in a different way in our framework to

---

[5] For reasons of clarity, some information in these rules is omitted.

429

derive both strict and sloppy readings where possible.

This presentation is informal. Formal definitions of this approach, and a comprehensive description of the PROLOG-implementation can be found in the author's Master thesis under preparation, to appear in August 1993.

Acknowledgments

References

[Alshawi, 1992] Hiyan Alshawi, editor. *The Core Language Engine*. The MIT Press, 1992.

[Asher, 1990] Nicholas Asher. Themes in Discourse Representation Theory. *Second European Summer School in Language, Logic and Information*, 1990.

[Dalrymple *et al.*, 1991] Mary Dalrymple, Stuart M. Shieber, and Fernando C.N. Pereira. Ellipsis and Higher-Order Unification. *Linguistics and Philosophy*, 14:339–452, 1991.

[Hankamer and Sag, 1976] Jorge Hankamer and Ivan Sag. Deep and surface anaphora. *Linguistic Inquiry*, 7(3):391–428, 1976.

[Johnson and Klein, 1986] Mark Johnson and Ewan Klein. Discourse, Anaphora, and Parsing. In *Coling*, 1986.

[Kamp and Reyle, 1990] Hans Kamp and Uwe Reyle. *From Discourse to Logic; An Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and DRT*. Kluwer, Dordrecht, 1990.

[Klein, 1986] Ewan Klein. VP Ellipsis in DR Theory. *Studies in Discourse Representation Theory and the Theory of Generalised Quantifiers*, 1986.

[Pereira and Shieber, 1987] Fernando C.N. Pereira and Stuart M. Shieber. *Prolog and Natural-Language Analysis*. CSLI, Stanford, 1987.

[Van der Sandt, 1992] Rob Van der Sandt. Presupposition Projection as Anaphora Resolution. *Journal of Semantics*, 9:333–377, 1992.

[Zeevat, 1991] Hendrik Willem Zeevat. *Aspects of Discourse Semantics and Unification Grammar*. PhD thesis, University of Amsterdam, 1991.