Representation and Information in Dynamic Semantics

Paul Dekker[*]
University of Amsterdam

Among the variety of phenomena that have been the subject of study within dynamic semantics, the phenomenon of inter-sentential anaphora probably has received most attention. Traditionally, pronouns have been associated with (syntactically free, but semantically somehow bound) variables. Put in a nutshell, the semantic relationships between pronouns and their antecedents are established in a compositional way by associating both with variables, and defining the interpretation algorithm as a function updating information about the possible values of these variables. Thus, information about the values of antecedent terms is available when co-indexed pronouns are encountered.

In Dekker 1994 I have shown that the same empirical results can be obtained without labeling the subjects introduced by candidate antecedents with variables, that is, by *not* conflating natural language pronouns with a logic's variables. In the system of predicate logic with anaphora which is presented in that paper, anaphoric relationships are accounted for by keeping track of the possible values of potential antecedent terms, not of the variables they have been associated with.

In this paper I want to discuss the impact of this distinction between pronouns and variables upon the notions of representation and information involved in a dynamic semantics dealing with the interpretation of anaphoric relationships.

1 Dynamic Semantics

In addition to the generally acknowledged context dependence of expressions, systems of dynamic semantics set out to account also for the context change potential of expressions. The leading idea is that the meaning of a sentence does not (solely) lies in its truth conditions, but rather "in the way it changes (the representation of) the information of the interpreter" (Groenendijk and Stokhof 1991). In the area of natural language semantics, a dynamic notion of meaning has been applied with considerable success in accounts of a whole array of phenomena, including presupposi-

[*]The research for this paper was supported by the ESPRIT Basic Research Project 6852, DYANA, Dynamic Interpretation of Natural Language. The paper is part of a more substantial contribution to DYANA deliverable R2.1.B, *Integrating Semantic theories II*, edited by Robin Cooper and Jeroen Groenendijk.

tion, epistemic modality, the temporal structure of texts, discourse focus, defaults and, last but not least, inter-sentential anaphora.

The semantics of inter-sentential anaphoric relationships perhaps most clearly reveals the basic need of a dynamic notion of meaning. For instance, consider the following very simple sequence of sentences:

(1) A man is walking through the park. He is whistling.

These two sentences together are generally used to claim that a man who is walking through the park is whistling. That is, the pronoun *he* is interpreted as coreferential with its antecedent *a man*, even though there need not be one unique man who is walking through the park. Examples like this one pose a problem for classical, static theories of interpretation. As long as the two sentences are assigned independent denotations, there appears to be no non-ad hoc way of relating the interpretation of the pronoun with that of its antecedent.

A dynamic semantics appears to be well suited to deal with such anaphoric relationships. In a dynamic semantics the interpretation of a piece of discourse involves a constant update and adjustment of the information which is passed on for the processing of subsequent discourse. This dynamic perspective upon meaning gives us a handle to deal with an example like 1. If only, after processing the clause *a man is walking through the park*, we keep track of the possible men who walk through the park, then we are able to interpret subsequent pronouns as referring back to them.

Discourse representation theory (*DRT*, Kamp 1981, Kamp and Reyle 1993) and dynamic predicate logic (*DPL*, Groenendijk and Stokhof 1991) are two well-known examples of a dynamic semantics which deal, among others, with anaphoric relationships. In *DRT*, interpretation is in the first place defined in terms of update of (discourse) representations. These representations consist of a set of discourse referents, a kind of pegs which stand in for candidate antecedents of anaphoric pronouns, and a set of conditions on discourse referents. Anaphoric relationships are established at this level of representation by the association of pronouns with discourse referents which have been introduced by antecedent noun phrases in the current discourse representation. Discourse representations, in their turn, are interpreted in terms of the possible assignments of objects to discourse markers under which the conditions on discourse markers are satisfied. Thus, via the mediating discourse representations, pronouns and their antecedents are co-valuated.

With their system of *DPL*, Groenendijk and Stokhof show that an intermediary level of representation is not essential for an account of anaphoric relationships. It appears that the kind of information underlying the *DRT* notion of a discourse representation is in fact that of information about the values of discourse referents, or, more simply, of variables. In terms of this notion of information Groenendijk and Stokhof give a dynamic, compositional reformulation of (basic aspects of) the *DRT* account of anaphoric relationships. The *DPL* system models the interpretation of example 1 by defining interpretation as a function updating information about the values of variables. The result of interpreting the first sentence *a man is walking through the park* is an information state which contains the information that the value of a variable, say x, is a man who is walking through the park. By matching the pronoun *he* in the subsequent sentence *he whistles* with this variable x, the two terms are co-valuated.

It may be noticed that the kind of information employed in *DPL* (and implicit in *DRT*) is of a mixed syntactic/semantic nature. Information states are sets of assignments, which are functions from syntactic entities like variables to the individuals of a (predicate logic) model. (And something essentially similar goes for related approaches like those presented in Heim 1982, Barwise 1987, Rooth 1987, Zeevat 1989, Chierchia 1992, Dekker 1993, Pagin and Westerståhl 1993.) In the next two sections I will show that the *DPL* and the *DRT* interpretation procedure can be reformulated using ordinary notions of information and representation, respectively. For this, it is essential to make an explicit distinction between pronouns and variables, and to introduce anaphoric pronouns as a category of terms of its own.

2 Predicate Logic with Anaphora

In the *DPL* account of natural language anaphora, the information that is passed on in the process of interpretation is information about the possible values of variables that have been associated with possible antecedents of anaphoric pronouns. In Dekker 1994 I have shown that the semantics of anaphoric relationships can as well be accounted for without this intermediary use of variables. What has to be accounted for with regard to inter-sentential anaphora, is the correlation between the interpretation of pronouns and that of their antecedents. On a dynamic account of such anaphoric relationships, the possible values of (possible) antecedents must be passed on in the process of interpretation. There is no need to encode this information in terms of the values of *variables* which are associated with potential antecedents.

In Dekker 1994's Predicate Logic with Anaphora (*PLA*), the formulas of a language of predicate logic are used to represent the meanings of simple sentences or sentential clauses of natural language. Like in Groenendijk and Stokhof 1991, existentially quantified formulas are used to represent the context change potential of natural language indefinites. The difference with *DPL* is that the semantic contribution of these indefinites is not hooked up to the specific variables quantified over.

The language of this system of *PLA* is constructed like that of predicate logic from sets of relation constants R^n of arity n, from a set C of individual constants, and infinite sets V and $A = \{\mathrm{p}_i \mid i \in \mathcal{N}\}$ of variables and pronouns, respectively. A *PLA* model is an ordinary *PL* model $M = \langle D, F\rangle$, where D is a non-empty domain of individuals, and F an interpretation function which assigns individuals in D to individual constants and sets of n-tuples of individuals to n-place relation expression.

In order to deal with anaphoric relationships, *PLA* interpretation is defined as an update function on information states. These information states encode the possible values of possible antecedents of anaphoric pronouns. They consist of the 'cases' (tuples of objects) 'verifying' preceding discourse (cf., Lewis 1975). Like in Dekker 1994, I use the term 'subjects' to indicate the possible denotations of such antecedents:

Definition 1 (Information states)

- $S^n = \mathcal{P}(D^n)$ *is the set of information states about* n *subjects*
- $S = \bigcup_{n\in\mathcal{N}} S^n$ *is the set of information states*

A state of information about n subjects can be pictured as follows:

$$\left\{ \begin{array}{c} \langle d_1,\ldots,d_n\rangle \\ \\ \vdots \\ \\ \langle d'_1,\ldots,d'_n\rangle \end{array} \right\} \quad \begin{array}{l} \nwarrow \\ \leftarrow \;\; \text{Possible cases} \\ \swarrow \end{array}$$

$$\nwarrow \uparrow \nearrow$$

Possible values
of subjects

For a state $s \in S^n$ and $0 < j \leq n$, and for any case case $e = \langle d_1, \ldots, d_n\rangle \in s$, d_j is a possible value of the j-th subject of s, and this value will also be indicated as e_j. An information state s contains the information that the i-th subject is a farmer who owns a donkey which is the j-th subject, iff, for every case e in s, e_i is a farmer, e_j a donkey and e_i owns e_j.

Before we turn to the interpretation of *PLA* formulas, I first give the *PLA* definition of information growth. In what follows, I will use the following notation conventions:

- if $e \in D^n$ and $e' \in D^m$, then $e \cdot e' = \langle e_1, \ldots, e_n, e'_1, \ldots, e'_m\rangle \in D^{n+m}$
- e' is an extension of e, $e \leq e'$, iff $\exists e''$: $e' = e \cdot e''$
- for $s \in S^n$, $N_s = n$, the number of subjects of s

Information growth consists in getting more informed about more subjects. The first aspect of information growth boils down to the reduction of the number of alternatives and the second to the extension of possible cases. Putting things together, an update of a state s is a state that consists, only, of extensions of possibilities in s:

Definition 2 (Information update) *State s' is an update of state s, $s \leq s'$, iff $N_s \leq N_{s'}$, and $\forall e' \in s'\ \exists e \in s$: $e \leq e'$*

As may be expected, information growth is transitive, reflexive and antisymmetric:

Observation 1 *$\langle S, \leq \rangle$ is a partial order*

We can now turn to the interpretation of terms. Constants and variables are evaluated in the usual way with respect to the model and an assignment function. The interpretation of a pronoun depends both on an information state s, and a case $e \in s$. A pronoun is associated with a subject of s, say the i-th subject, and relative to some case $e \in s$ its value is the i-th individual of e:

Definition 3 (Interpretation of terms)

- $[c]_{M,s,e,g} \ = F(c)$ *if* $c \in C$
- $[x]_{M,s,e,g} \ = g(x)$ *if* $x \in V$
- $[\mathtt{p}_i]_{M,s,e,g} = e_{N_s - i}$ *if* $\mathtt{p}_i \in A$, $e \in s$ *and* $N_s > i$

A pronoun $\mathtt{p}_i$ is associated with the $i+1$-th last subject of an information state (if it exists, that is; otherwise, the interpretation of the pronoun is undefined). So, the pronoun with the index 0 picks out the subject introduced last, and its value is the last individual of any case with respect to which it is evaluated; the pronoun with index 1 picks out the forelast mentioned subject, etc., etc.

The interpretation $s[\![\phi]\!]_{M,g}$ of a *PLA* formula ϕ in an information state s is defined with respect to a model M and an assignment g. The parameters M and g are the ordinary ones from predicate logic, and they behave in the same, static, way. The state parameter s is, as it were, the dynamic one. The interpretation of a formula ϕ in a state s yields an update of s. The interpretation of *PLA* formulas is defined as follows (here, if X is a set of terms, then I_X is the smallest number greater then or equal to the index of every pronoun in X):

Definition 4 (Semantics of PLA)

- $s[\![Rt_1 \ldots t_n]\!]_{M,g} = \{e \in s \mid \langle [t_1]_{M,s,e,g}, \ldots, [t_n]_{M,s,e,g} \rangle \in F(R)\}$
  (*if* $N_s > I_{\{t_1,\ldots,t_n\}}$)
- $s[\![t_1 = t_2]\!]_{M,g} \ = \{e \in s \mid [t_1]_{M,s,e,g} = [t_2]_{M,s,e,g}\}$ (*if* $N_s > I_{\{t_1,t_2\}}$)
- $s[\![\neg\phi]\!]_{M,g} \ = \{e \in s \mid \neg\exists e'$: $e \leq e'$ & $e' \in s[\![\phi]\!]_{M,g}\}$
- $s[\![\exists x\phi]\!]_{M,g} \ = \{e' \cdot d \mid d \in D$ & $e' \in s[\![\phi]\!]_{M,g[x/d]}\}$
- $s[\![\phi \wedge \psi]\!]_{M,g} \ = s[\![\phi]\!]_{M,g}[\![\psi]\!]_{M,g}$
- $s \models_{M,g} \phi$ *iff* $\forall e \in s$: $\exists e'$: $e \leq e'$ & $e' \in s[\![\phi]\!]_{M,g}$

An atomic formula At is evaluated with respect to any case e in the state of information s in which interpretation takes place. If such a formula only contains variables and individual constants as terms, its evaluation is in fact independent from these cases (and from s). In such a case, $s[\![At]\!]$ either equals s, iff At is classically true with respect to M and g, or $\emptyset$, iff At is classically false with respect to M and g. Only when pronouns come into play the differences between the various cases in s may matter. If, for instance, the formula is $W\mathtt{p}_i$ ("the $i+1$-th last subject walks"), then its interpretation in a state s will involve the elimination of those cases e in s of which the $i+1$-th last element does not walk (in M).

Existentially quantified formulas are taken to 'set up' discourse referents: they introduce subjects to information states. The interpretation of an existentially quantified formula $\exists x\phi$ in s involves the extension of the cases in the update of s with ϕ with x's 'witnesses'. Thus, they remain accessible for future anaphoric (co-)reference. The interpretation of a formula $\neg\phi$ in s preserves the cases in s that don't survive the update of s with ϕ, i.e., the cases that don't support ϕ. In keeping with the dynamic view on interpretation, sentence sequencing, or conjunction, involves the composition of the two update functions associated with the conjuncts. A state s supports a formula ϕ, $s \models \phi$, iff all cases in s survive the update with ϕ. In other words, ϕ is supported by s if s already contains the information conveyed by ϕ about s's subjects. (As usual, $\forall x\phi$ and $\phi \to \psi$ are taken to abbreviate $\neg\exists x\neg\phi$ and $\neg(\phi \wedge \neg\psi)$, respectively.)

Since pronouns may fail a denotation in a state s, the interpretation of a *PLA* formula can be partial. However, notice that the interpretation of a formula ψ may be partial, since it presupposes the presence of a certain number of subjects, while the interpretation of a conjunction $\phi \wedge \psi$ is total, that is, when ϕ involves the introduction of at least that number of subjects. I will say that a pronoun in a formula is resolved if it refers to a subject introduced in the very same formula, and if all the pronouns in a formula are resolved then the formula itself is called resolved.

In order to illustrate the above definitions, let us take a look at a couple of examples. First, consider the interpretation of the sentence *There is a man*, translated as $\exists xMx$, in some state s:

(2) There is a man.

$$\begin{aligned} s[\![\exists xMx]\!]_g &= \{e \cdot d \mid e \in s[\![Mx]\!]_{g[x/d]}\} \\ &= \{e \cdot d \mid e \in s \;\&\; d \text{ is a man}\} \;(= s') \end{aligned}$$

The interpretation of this example yields a state consisting of all the cases $e \in s$ extended with an individual d which is a man. The last subject in the resulting state (which I will refer to as s' in the following examples) simply is an arbitrary man.

With a pronoun we can refer back to this man. Above I have stipulated that the i-th last subject of a state is referred back to by the pronoun with index $i-1$. So, pronoun $\mathtt{p}_0$ refers to the subject introduced last, as in the following continuation of example 2:

(3) (There is a man.) He walks.

$$\begin{aligned} s'[\![W\mathrm{p}_0]\!] &= \{e' \in s' \mid \text{the last element of } e' \text{ walks}\} \\ &= \{e \cdot d \mid e \in s \;\&\; d \text{ is a man } \&\; d \text{ walks}\} \end{aligned}$$

The interpretation of this example involves the elimination of all those cases $e' \in s'$ the last element of which does not walk. The last subject in the resulting state is an arbitrary walking man.

As an illustration of the clause dealing with negation, consider a continuation of example 2 with the sentence *Nobody knows him* with translation $\neg\exists x K x \mathrm{p}_0$:

(4) (There is a man.) Nobody knows him.

$$\begin{aligned} s'[\![\neg\exists x K x \mathrm{p}_0]\!] &= \{e' \in s' \mid \neg\exists e''\text{: } e' \leq e'' \;\&\; e'' \in s'[\![\exists x K x \mathrm{p}_0]\!]_{M,g}\} \\ &= \{e' \in s' \mid \neg\exists d'\text{: } e' \in s'[\![K x \mathrm{p}_0]\!]_{M,g[x/d']}\} \\ &= \{e' \in s' \mid \neg\exists d'\text{: } d' \text{ knows the last element of } e'\} \\ &= \{e \cdot d \mid e \in s \;\&\; d \text{ is a man } \&\; \neg\exists d'\text{: } d' \text{ knows } d\} \end{aligned}$$

The last subject in this state is an arbitrary, unknown man.

To conclude this section, I point at some characteristic properties of *PLA*. Quite unlike *DPL*, *PLA* obeys the following ordinary substitution law:

Observation 2 (α-conversion) $\exists x\phi \Leftrightarrow \exists y[y/x]\phi$ *if y is free for x in ϕ and y does not occur free in ϕ*

In short, in *PLA* the ordinary notions of scope and binding apply. In fact, the subsystem of *PLA* without pronouns is fully equivalent with classical predicate logic:

Observation 3 (PL and PLA) *For any PL formula ϕ: PL $\models_{M,g} \phi$ iff $s \models_{M,g} \phi$ (for any state s)*

These two observations may go to show that the *PLA*-system is a proper extension, not modification, of ordinary predicate logic.

It is also readily established that interpretation in *PLA* always produces information update:

Observation 4 (Update) $\forall s$*: $s \leq s[\![\phi]\!]$ (if defined)*

The system of *PLA* simply models the introduction of subjects (by existentially quantified formulas) and the update of information about these subjects (by means of pronouns). It is fairly obvious that this reflects the natural language practice of indefinitely setting up and anaphorically referring back to subjects. It may be noticed that most bound variable approaches to anaphora do not have this update property. In such approaches, the introduction of a subject as the value of a variable x involves the elimination of a subject introduced earlier as the value of x. Put crudely, in a bound variable subjects may get 'dumped' as a consequence of unfortunately indexing natural language terms.

The last observation can be strengthened as follows:

Observation 5 (Registration) *For all s, $e \in D^{N_s}$: $e \in s$ & $\{e\} \models \phi$ iff $\exists e'$: $e \leq e'$ & $e' \in s[\![\phi]\!]$*

The update of a state s with ϕ contains (only) cases that register, i.e., extend, the cases in s that all by themselves support ϕ. For this reason, it is appropriate to define the notion of support in the way we did, i.e., in terms of a state s and the update of s with ϕ. I just note that such a way of defining support is crucial for an extension of the system with an account of epistemic modalities along the lines of Groenendijk et al. 1994.

3 Representation Theory for Anaphora

The notion of information employed in *PLA* is that of information about subjects, modeled in terms of sets of sequences of individuals. *PLA* information states are ordinary model-theoretic objects (relations), and they are the denotations of ordinary relation expressions. For this reason, the *PLA* updates of information states can also be represented as updates of relation expressions. This section presents such a representational formulation of *PLA*. The basic difference between the ensueing representation theory for anaphora (*RTA*) and *DRT* is that ordinary representations of existing formalisms are employed.

For perspicuity's sake, and in order to simplify comparison, I present a representational semantics for the *PLA*-translations of natural language sentences. The *RTA* interpretation of a *PLA* formula is defined to be a function 'updating' type theoretical relation expressions. In order to keep the correspondence with the *PLA* semantics as close as possible, I will assume Orey's relational models for the type theory, together with Muskens 1989's analysis of abstraction and application. To be precise, I use a basic type e, and relational types $\langle a_1, \ldots, a_n \rangle$, where, $a_1, \ldots, a_n$ are types, for $0 \leq n$. For each type a there is a domain D_a, defined such that $D_e \neq \emptyset$, and $D_{\langle a_1, \ldots, a_n \rangle} = \mathcal{P}(D_{a_1} \times \ldots \times D_{a_n})$. The distinguished clauses of the interpretation function are the following (for β of some type $\langle a_0, a_1, \ldots, a_n \rangle$, and γ of type $\langle a_1, \ldots, a_n \rangle$, and α and x of type a_0):

- $[\beta(\alpha)]_{M,g} = \{\langle d_1, \ldots, d_n \rangle \mid \langle [\alpha]_{M,g}, d_1, \ldots, d_n \rangle \in [\beta]_{M,g}\}$
- $[\lambda x \gamma]_{M,g} = \{\langle d_0, d_1, \ldots, d_n \rangle \mid \langle d_1, \ldots, d_n \rangle \in [\gamma]_{M,g[x/d_0]} \& d_0 \in D_{a_0}\}$

(Notice that these are harmless assumptions.) Moreover, I will assume that *PLA*'s variables are variables (of type e) of the type-theoretical language $\mathcal{L}$ and that the constants of $\mathcal{L}$ are those of *PLA*. Thus, we can assume models $M = \langle D, F \rangle$ for $\mathcal{L}$ which are also *PLA* models.

I use the following abbreviations for relational types:

- $\sigma^n \quad = \langle e_1, \ldots, e_n \rangle$
- $(a, b) = \langle a, a_1, \ldots a_n \rangle$ if $b = \langle a_1, \ldots a_n \rangle$
- $\tau^{n,m} = (\sigma^n, \sigma^m)$

The type σ^n is that of n-place relations between individuals, i.e., of states of information about n subjects. Types (a, b) mimic functional types in the relational system, and in terms of that, the type $\tau^{n,m}$ is defined to be that of functions from states of information about n subjects to states of information about m subjects. Finally, the following notation conventions will be employed. If $\vec{x}^n$ is a sequence of variables $x_1, \ldots, x_n$ (all of type e), then:

- $s(\vec{x}^n) = s(x_1) \ldots (x_n)$ (of type σ^m, for s of type σ^{n+m})
- $\lambda\vec{x}^n\gamma = \lambda x_1 \ldots \lambda x_n \gamma$ (of type σ^{n+m}, for γ of type σ^m)
- $\exists\vec{x}^n\phi = \exists x_1 \ldots \exists x_n \phi$ (of type σ^0, for ϕ of type σ^0)

For s of type σ^n, I write $\downarrow s$ for the closure $\exists\vec{x}^n s(\vec{x}^n)$ of s of type σ^0.

We may now turn to the definition of $s^n(\!(\phi)\!)$, the representational update of a relation term s^n of type σ^n by ϕ. The result of this, if defined, is always a relation term of some type σ^{n+m}:

Definition 5 (RTA)

- $[c]_{\vec{x}^n} = c \quad [y]_{\vec{x}^n} = x \quad [\mathtt{p}_i]_{\vec{x}^n} = x_{n-i}$ *(if it exists)*
- $s^n(\!(Rt_1 \ldots t_m)\!) = \lambda\vec{x}^n\ s(\vec{x}^n) \wedge R([t_1]_{\vec{x}^n}, \ldots, [t_m]_{\vec{x}^n})$
  $s^n(\!(t_1 = t_2)\!) = \lambda\vec{x}^n\ s(\vec{x}^n) \wedge [t_1]_{\vec{x}^n} = [t_2]_{\vec{x}^n}$
  $s^n(\!(\neg\phi)\!) = \lambda\vec{x}^n\ s(\vec{x}^n) \wedge \neg{\downarrow}s(\!(\phi)\!)(\vec{x}^n)$
  $s^n(\!(\exists y\phi)\!) = \lambda\vec{x}^{n+m}\lambda y\ s(\!(\phi)\!)(\vec{x}^{n+m})$ *(for* $(\!(\phi)\!)$ *of type* $\tau^{n,n+m}$*)*
  $s^n(\!(\phi \wedge \psi)\!) = s(\!(\phi)\!)(\!(\psi)\!)$

observing appropriate variable conventions (viz., that the variables in x^n *and* x^{n+m} *are different and not free in the updated representation or in the formula inducing the update).*

Let us quickly go through one example to illustrate these definitions. Consider the update of a representation s^n induced by the sequence *There is a man. He walks*, $\exists yMy \wedge W\mathtt{p}_0$. First we determine the update with the first conjunct:

$$\begin{aligned} s^n(\!(\exists yMy)\!) &= \lambda\vec{x}^n\lambda y\ s(\!(My)\!)(\vec{x}^n) \\ &\Leftrightarrow \lambda\vec{x}^n\lambda y\ s(\vec{x}^n) \wedge My \end{aligned}$$

This representation, abbreviated as t^{n+1}, is next updated with the second conjunct $W\mathtt{p}_0$:

$$\begin{aligned} t^{n+1}(\!(W\mathtt{p}_0)\!) &= \lambda\vec{x}^{n+1}\ t(\vec{x}^{n+1}) \wedge Wx_{n+1} \\ &\Leftrightarrow \lambda\vec{x}^n\lambda x_{n+1}\ s(\vec{x}^n) \wedge Mx_{n+1} \wedge Wx_{n+1} \end{aligned}$$

Here we see that the predicates M and W are co-instantiated. The resulting relation term denotes an information state the last subject of which is an arbitrary man who walks.

It is relatively easily established that the *RTA* update of a representation s denotes the *PLA* update of the denotation of s:

Observation 6 (RTA and PLA) $[s(\!(\phi)\!)]_{M,g[s/\mathbf{s}]} = \mathbf{s}[\![\phi]\!]_{M,g}$

So, as appears from this observation, *RTA* really is the representational correlate of *PLA*.

It may be clear that there is a close connection between the representations generated by *RTA*, and the *DRSs* of *DRT*. The relation terms that result from the *RTA* update of a minimal relation $\top$ (e.g., $x = x$) are easily transformed into equivalent *DRSs*. In order to be more precise, first observe that the β-normal forms of all the *RTA* representations are of the form $\lambda\vec{x}^n\ c_0 \wedge c_1 \wedge \ldots \wedge c_m$, where c_0 is $\top$, and every formula c_i, for $0 < i \leq m$, is a *condition*, viz., an atomic formula $Rt_1 \ldots t_m$ or $t_1 = t_2$, or a negation $\neg\exists\vec{y}^k\phi$ (where ϕ, in its turn, is a conjunction $c_0 \wedge c'_1 \wedge \ldots \wedge c'_l$ of conditions). These representations can be turned into *DRS*s by the following translation $()^\clubsuit$ (I assume the *DRS* syntax as presented in Groenendijk and Stokhof 1991):

- $(Rt_1 \ldots t_m)^\clubsuit = Rt_1 \ldots t_m$
  $(t_1 = t_2)^\clubsuit \quad = t_1 = t_2$
- $(\lambda\vec{x}^n\ c_0 \wedge c_1 \wedge \ldots c_m)^\clubsuit = [\vec{x}^n][(c_1)^\clubsuit \ldots (c_m)^\clubsuit]$
  $(\neg\exists\vec{y}^k\ c_0 \wedge c'_1 \wedge \ldots c'_l)^\clubsuit = \neg[\vec{y}^k][(c'_1)^\clubsuit \ldots (c'_l)^\clubsuit]$

This gives us a translation of *PLA* into the *DRS* language. The composition of i) the *RTA* update of minimal representation $\top$, ii) the β-normalization of the resulting state $()^\beta$, and, iii) its subsequent translation $()^\clubsuit$ produces a truth-conditionally equivalent *DRS*:

Observation 7 (PLA and DRT) *if ϕ is a resolved formula without free variables, then $\top^0 \models_M \phi$ in* PLA *iff $((\top(\!(\phi)\!))^\beta)^\clubsuit$ is true in M in* DRT

The discussion sofar shows that, simply by not conflating pronouns with variables, *DRT* and *DPL* results regarding the semantics of anaphoric relationships in natural language can be obtained using orthodox notions of representation and information. Here, one might feel a temptation to object. With the development of an *RTA*- or *PLA*-style semantics, what we have been doing is pushing the theory of interpretation to one of the two extremes of representationalism and non-representationalism. The objection then might be that this is the wrong way to go about, since, in the end, a most viable theory of interpretation should have to be located somewhere in the middle of these two extremes. Now I can agree that a comprehensive theory of natural language interpretation in the end probably is in need of informational structures in which all relevant aspects of representation and information are integrated. However, as the preceding may have shown, the phenomena we have been concerned with here appear to have no immediate bearing on *that* issue. For, as we have seen just now, the syntactic aspects of the *DPL* notion of information, or the semantic idiosyncrasies of *DRT*'s *DRSs*, simply are no inalienable ingredient of an account of the semantics of anaphoric links.

4 Representation and Translation

The *DPL* reformulation of *DRT* was, by and large, given in by considerations concerning the principle of compositionality, a principle which probably most would agree should be secured if it is not really too expensive. With *DPL* Groenendijk and Stokhof have developed a dynamic notion of the meaning of sentences which enables an account of local semantic dependencies between independently interpreted sentences. However, a fully general dynamic formulation of a Montagovian semantics of natural language which is also compositional at the sub-sentential level, requires a generalization of such a dynamic notion of meaning which applies to all the types that sub-sentential expressions of natural language may have. It appears that this task is complicated by the analysis of pronouns as variables. For, for a fully compositional analysis of anaphoric relationships along such lines either variables or variable assignments are to be counted in some or other way among the (model-theoretic) domains of interpretation. For this reason, in compositional elaborations additional types or sorts have been used which, by means of postulates, are made to behave like variables or variable assignments (cf., among others, Groenendijk and Stokhof 1990, Muskens 1994).

In this final section I want to show that no further syntactico-semantic assumptions really are in order if pronouns and variables are properly kept distinct. A fully compositional interpretation of a fragment of natural language, which includes anaphora, can be given in terms of an unconstrained extensional type theory, with only one basic type: that of individuals.

I now turn to an example of a fully compositional treatment of a minuscule fragment of natural language covering anaphora. The sentences of this fragment will be interpreted as *PLA* update functions on information states. *RTA* 'interpretations' are used here as the sentences' translations which denote these functions. Here we must take serious one aspect of *PLA* interpretation (and that of *RTA*, for that matter) that may have remained implicit thusfar: its polymorphism. Strictly speaking, any two *PLA* information states about a different number of subjects belong to different types. When we extend a *PLA* semantics to the sub-sentential level, this kind of type-polymorphism must be explicitly dealt with. For this reason, it is convenient to adopt a notation convention to state all the (different) translations of different types in a uniform way. I will use the following format:

$$A \rightsquigarrow B \quad [C]$$

which must be read as saying that an expression A is associated with a representation B under the condition(s) C. The condition(s) C are polymorphic type declarations. For example:

$$A \rightsquigarrow \lambda s \lambda \vec{x}^n \; s(\vec{x}^n) \wedge \phi \;\; [s : \sigma^n]$$

must be taken to say that A receives all of the following translations:

1. $\lambda s\ s \wedge \phi$ (for $s : \sigma^0$)
2. $\lambda s \lambda x\ s(x) \wedge \phi$ (for $s : \sigma^1$)
3. $\lambda s \lambda x \lambda y\ s(x)(y) \wedge \phi$ (for $s : \sigma^2$)

$\vdots$

It may be noticed that, although every expression of our fragment will be associated with an (infinite) set of translations, these denote singular (polymorphic) functions.

Using this notation convention we can give the following sample translations of the basic expressions and operations of a fragment of natural language (which is left fully implicit itself):

Definition 6 (Examples of basic translations)

- man $\rightsquigarrow \lambda x \lambda s \lambda \vec{z}^n\ s(\vec{z}^n) \wedge \mathtt{man}(x) \quad [s : \sigma^n]$
- walk $\rightsquigarrow \lambda x \lambda s \lambda \vec{z}^n\ s(\vec{z}^n) \wedge \mathtt{walk}(x) \quad [s : \sigma^n]$
- own $\rightsquigarrow \lambda T \lambda x\ T(\lambda y \lambda s \lambda \vec{z}^k\ s(\vec{z}^k) \wedge \mathtt{own}(y)(x))$
  $[s : \sigma^k \quad T : ((e, \tau^{k,k}), \tau^{n,m})]$
- a $\rightsquigarrow \lambda P \lambda Q \lambda s \lambda \vec{z}^m \lambda y\ (Q(y)(P(y)(s)))(\vec{z}^m)$
  $[s : \sigma^n \quad P : (e, \tau^{n,k}) \quad Q : (e, \tau^{k,m})]$
- he$_i$ $\rightsquigarrow \lambda Q \lambda s \lambda \vec{z}^{n+m}\ Q(z_{n-i})(s)(\vec{z}^{n+m})$
  $[s : \sigma^n \quad Q : (e, \tau^{n,n+m})]$

Definition 7 (Construction rules)

- $(\beta_{B/A}\ \alpha_A)_B \rightsquigarrow \beta'(\alpha') \quad [\beta' : (a, b) \quad \alpha' : a]$
- $(\text{Not}\ \pi_S)_S \rightsquigarrow \lambda s \lambda \vec{z}^n\ s(\vec{z}^n) \wedge \neg{\downarrow}\pi'(s)(\vec{x}^n)$
  $s : \sigma^n \quad \pi' : \tau^{n,n+m}$
- $(\pi_S\ .\ \rho_S)_S \rightsquigarrow \lambda s\ \rho'(\pi'(s))$
  $s : \sigma^n \quad \pi' : \tau^{n,k} \quad \rho' : \tau^{k,m}$
- $(\alpha_{CN}\ \text{who}\ \beta_{IV})_{CN} \rightsquigarrow \lambda x \lambda s\ \beta'(x)(\alpha'(x)(s))$
  $s : \sigma^n \quad \alpha' : (e, \tau^{n,k}) \quad \beta' : (e, \tau^{k,m})$

In order to show how things work out, let us consider the interpretation of two examples. The first example is *A man walks. He talks*. The translation of the first sentence of this example is obtained by applying translations of *a* to fitting translations of *man* and *walk* respectively. The results of the first application can be specified as follows (all reductions here and in what follows are obtained by λ-conversion):

$$A\ man \rightsquigarrow \qquad [s : \sigma^n \quad P : (e, \tau^{n,n}) \quad Q : (e, \tau^{n,m})]$$
$$(\lambda P \lambda Q \lambda s \lambda \vec{z}^m \lambda y\ Q(y)(P(y)(s))(\vec{z}^m))(\lambda x \lambda s \lambda \vec{z}^n\ s(\vec{z}^n) \wedge \mathtt{man}(x))$$
$$\Leftrightarrow \quad \lambda Q \lambda s \lambda \vec{z}^m \lambda y\ Q(y)(\lambda \vec{z}^n\ s(\vec{z}^n) \wedge \mathtt{man}(y))(\vec{z}^m)$$

Application to *walk* yields:

$$A\ man\ walks \rightsquigarrow \qquad [s : \sigma^n \quad Q : (e, \tau^{n,n})]$$

$$(\lambda Q \lambda s \lambda \vec{z}^n \lambda y Q(y)(\lambda \vec{z}^n s(\vec{z}^n) \wedge \texttt{man}(yt))(\vec{z}^n))(\lambda x \lambda s \lambda \vec{z}^n s(\vec{z}^n) \wedge \texttt{walk}(x))$$
$$\Leftrightarrow \quad \lambda s \lambda \vec{z}^n \lambda y \ s(\vec{z}^n) \wedge \texttt{man}(y) \wedge \texttt{walk}(y)$$

In a similar way we construct the translation(s) of the second sentence:

$$\textit{He}_0 \ \textit{talks} \rightsquigarrow \qquad [s' : \sigma^m \quad Q : (e, \tau^{m,m})]$$
$$(\lambda Q \lambda s' \lambda \vec{z}^m \ Q(z_m)(s')(\vec{z}^m))(\lambda x \lambda s' \lambda \vec{z}^m \ s'(\vec{z}^m) \wedge \texttt{talk}(x))$$
$$\Leftrightarrow \quad \lambda s' \lambda \vec{z}^m \ s'(\vec{z}^m) \wedge \texttt{talk}(z_m)$$

Equating m with $n + 1$, we can combine the two (sets of) translations according to the construction rule for sentence conjunction. The following translation(s) result:

$$\textit{A man walks. He}_0 \ \textit{talks} \rightsquigarrow \qquad [s : \sigma^n \quad s' : \sigma^m \quad \mathrm{m = n{+}1}]$$
$$\lambda s(\lambda s' \lambda \vec{z}^m s'(\vec{z}^m) \wedge \texttt{talk}(z_m))((\lambda s \lambda \vec{z}^n \lambda y \ s(\vec{z}^n) \wedge \texttt{man}(y) \wedge \texttt{walk}(y))(s))$$
$$\Leftrightarrow \quad \lambda s \lambda \vec{z}^n \lambda y \ s(\vec{z}^n) \wedge \texttt{man}(y) \wedge \texttt{walk}(y) \wedge \texttt{talk}(y)$$

This expression denotes the function π such that for any information state s, $\pi(s)$ is the information state that consists of all the extensions of tuples in s with a man who walks and talks.

The second example is *If a man walks, he talks*. As before, the interpretation of a conditional is stated in terms of negation and conjunction. So, *If* π*, then* ρ is rendered as *Not* (π *. not* ρ*)*. Spelling out the construction rules, we find that:

$$(\text{If } \pi_S, \text{ then } \rho_S)_S \rightsquigarrow \qquad [s : \sigma^n \quad \pi' : \tau^{n,n+k} \quad \rho' : \tau^{n+k,n+k+m}]$$
$$\lambda s \lambda \vec{u}^n \ s(\vec{u}^n) \wedge \forall \vec{v}^k (\pi'(s)(\vec{u}^n)(\vec{v}^k) \rightarrow \exists \vec{w}^m \rho'(\pi'(s))(\vec{u}^n)(\vec{v}^k)(\vec{w}^m))$$

In this translation scheme π' can be replaced by the translation of *A man walks* given above, and ρ' by that of *He talks*. Thus, for $s : \sigma^n$, we get $\pi' : \tau^{n,n+1}$, and $\rho' : \tau^{n+1,n+1}$, and we arrive at the following specification of the translation of the conditional sentence:

$$\lambda s \lambda \vec{z}^n \ s(\vec{z}^n) \wedge \forall y((s(\vec{z}^n) \wedge \texttt{man}(y) \wedge \texttt{walk}(y)) \rightarrow$$
$$(s(\vec{z}^n) \wedge \texttt{man}(y) \wedge \texttt{walk}(y) \wedge \texttt{talk}(y))) \Leftrightarrow$$
$$\lambda s \lambda \vec{z}^n \ s(\vec{z}^n) \wedge \forall y((\texttt{man}(y) \wedge \texttt{walk}(y)) \rightarrow \texttt{talk}(y))$$

With the preceding exposition I hope to have shown that a dynamic compositional interpretation of an extensional fragment of natural language, can be appropriately stated in terms of an ordinary extensional type theory. In fact, this is something one might have expected in the first place.

5 Concluding remarks

In *PLA*, the semantic connections between pronouns and their antecedents are accounted for in terms of update of information, not about the possible values of the variables associated with these antecedent terms, but about the possible values of these antecedent terms themselves. This simplified way of doing things enables us to account for the dynamics of establishing anaphoric relationships by means of a proper extension, not modification, of ordinary logical systems. Moreover, it avoids certain complications which pertain to a *DPL*-style approach to natural language anaphora. In particular, it does not automatically lead to arbitrary downdate of information.

As we have seen in this paper, the system of *PLA* also shows that a compositional treatment of the semantics of anaphoric relationships does not enforce upon us an idiosyncratic representation language, or some specialized intensional or many-sorted logic. The system of *PLA* and its representational correlate *RTA* remain well within the bounds of ordinary extensional type theory.

References

Barwise, Jon. 1987. Noun Phrases, Generalized Quantifiers and Anaphora. In *Generalized Quantifiers: Linguistic and Logical Approaches*, ed. Peter Gärdenfors. 1–29. Dordrecht: Reidel.

Chierchia, Gennaro. 1992. Anaphora and Dynamic Binding. *Linguistics and Philosophy* 15:111–83.

Dekker, Paul. 1993. *Transsentential Meditations. Ups and downs in dynamic semantics.* Doctoral dissertation, ILLC, Amsterdam.

Dekker, Paul. 1994. Predicate Logic with Anaphora (seven inch version). In *Proceedings of SALT IV*, ed. Mandy Harvey and Lynn Santelmann. Ithaca, NY. Cornell University.

Groenendijk, Jeroen, and Martin Stokhof. 1990. Dynamic Montague Grammar. In *Papers from The Second Symposium on Logic and Language*, ed. László Kálmán and László Pólos. 3–48. Budapest: Akadémiai Kiadó.

Groenendijk, Jeroen, and Martin Stokhof. 1991. Dynamic Predicate Logic. *Linguistics and Philosophy* 14:39–100.

Groenendijk, Jeroen, Martin Stokhof, and Frank Veltman. 1994. Coreference and Modality. In *Integrating Semantic Theories II*, ed. Robin Cooper and Jeroen Groenendijk. Amsterdam: ILLC. Dyana-2 deliverable R2.1.B.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases.* Doctoral dissertation, University of Massachusetts, Amherst.

Kamp, Hans. 1981. A Theory of Truth and Semantic Representation. In *Formal Methods in the Study of Language*, ed. Jeroen Groenendijk, Theo Janssen, and Martin Stokhof. 277–322. Amsterdam: Mathematical Centre.

Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic.* Dordrecht: Kluwer.

Lewis, David. 1975. Adverbs of Quantification. In *Formal Semantics of Natural Language*, ed. Edward L. Keenan. 3–15. Cambridge, U.K.: Cambridge University Press.

Muskens, Reinhard. 1989. A Relational Formulation of the Theory of Types. *Linguistics and Philosophy* 12:325–46.

Muskens, Reinhard. 1994. Compositional Discourse Representation Theory. In *Proceedings of the ninth Amsterdam colloquium*, ed. Paul Dekker and Martin Stokhof, 467–86. Amsterdam. ILLC.

Pagin, Peter, and Dag Westerståhl. 1993. Predicate Logic with Flexibly Binding Operators and Natural Language Semantics. *Journal of Logic, Language and Information* 2:89–128.

Rooth, Mats. 1987. NP Interpretation in Montague Grammar, File Change Semantics and Situation Semantics. In *Generalized Quantifiers*, ed. Peter Gärdenfors. 237–69. Dordrecht: Kluwer.

Zeevat, Henk. 1989. A Compositional Approach to Discourse Representation Theory. *Linguistics and Philosophy* 12:95–131.