Coreference and Representationalism*

Paul Dekker
University of Amsterdam

`http://turing.wins.uva.nl/~pdekker`

Abstract

The compositional interpretation of structures in discourse has raised the question whether some level of representation is indispensable in the analysis of the semantics of natural language. In this paper we formulate and motivate three notions of representationalism (a strong, a medium or midweak, and a weak one) and discuss to what extent existing formal semantic frameworks qualify as (strongly, midweakly or weakly) representational.

Research into the semantics of anaphorical relationships in the eighties has led to a debate about the question whether some level of representation is indispensable in the analysis of the semantics of natural language. Within the logically oriented or model-theoretic tradition of natural language semantics, Hans Kamp is a prominent proponent of the thesis that it is; Jeroen Groenendijk and Martin Stokhof have explicitly argued against it. Interestingly, the two nowadays agree on this issue: that some level of representation is indispensable in modeling the *interpretation* of natural language. But what about the *semantics* of natural language?

1 Introduction

In traditional theories of meaning, the question what is the meaning of a sentence or discourse has been kept separate from the question of how people actually assign meanings to texts when they interpret them. With the growing attention for the (inter-sentential) structure of discourse, this distinction has turned out to be not unobjectionable. Standard conceptions of what is the meaning of a sentence have been seen to be inadequate if the aim is to assign meanings to (structured) discourses in a compositional way. For instance, truth-conditions, or sets of possible worlds, have been seen to be insufficiently fine-grained means to model the semantics of inter-sentential anaphoric relationships, and therefore other, richer, notions of meaning have been developed for use in more elaborate frameworks of discourse interpretation.

*The research for this paper was supported by the Royal Netherlands Academy of Arts and Sciences (KNAW), and by the ESPRIT basic research project DYANA. For their stimulating ideas I am glad I can thank Irene Heim, Hans Kamp, Fred Landman and Ede Zimmermann, besides, of course, Jeroen Groenendijk, Martin Stokhof and Frank Veltman.

Of course, if additional means are called for to account for certain phenomena, they can, in principle, be assigned to the semantics proper, or to interpretation proper. Thus, e.g., some versions of situation semantics can be conceived of as locating the additional structure which is required to deal with anaphoric relationships in the semantics (Gawron and Peters 1990; Milward 1996). Alternatively, Cooper's E-type pronoun theory (cf. Evans 1977; Cooper 1979) assigns the additionally required means or processes to the area of interpretation (or to pragmatics, if you want). Heim 1990 presents a combined E-type and situation theoretic account, one in which the distinction between meaning and interpretation can be left intact.

Yet, in quite an important branch of nowadays semantic theory meaning and interpretation have grown towards each other. Discourse representation theory (Kamp 1981; Kamp and Reyle 1993), file change semantics (Heim 1982; Heim 1983), and various versions of dynamic semantics (Groenendijk and Stokhof 1991; Groenendijk et al. 1996), adopt some sort of *computational* meaning assignment, according to which the interpretation of a sentence is conceived of as a *process*. In discourse representation theory (*DRT*, henceforth), interpretation involves operations on discourse representation structures (*DRS*s), in file change semantics (*FCS*) operations on files, and in systems of dynamic semantics originating from dynamic predicate logic (*DPL*) operations on information states. In the three kinds of semantic architecture, meaning and interpretation seem to merge.

The dynamic taken with these last-mentioned systems of interpretation raises a number of theoretical questions. In the first place there is the question to what extent old-fashioned concepts of static semantic systems can be retained, or may have to be re-assessed when interpretation and meaning merge. For instance, although *DRT* does employ familiar concepts of satisfaction and truth for its language of *DRS*s, it is not clear from the outset how these concepts apply to the *represented sentences*, the sentences of the language which *DRT* aims to give an interpretation for. Similarly one may wonder whether concepts like rigid or demonstrative reference fit in in a system where sentences are taken to update information states (if these concepts have to fit in at all).

An apparently related question is the one we started out with. If interpretation involves some level of representation, and if interpretation and meaning constitute an inseparable whole, does this mean we are committed to a representational concept of meaning? Or might it be possible to delineate a part of the more general system of interpretation which could qualify as 'genuinely semantic'? In this paper we delineate such a, not overtly representational, part of a theory of the interpretation of singular pronouns. Matters of resolution are put aside. We take some resolution mechanism for pronouns (and demonstratives) for granted, and assume that the input of semantic interpretation consists of indexed syntactic structures, in which anaphoric pronouns are disambiguated, in a sense too be made precise below. We then argue that the delineated theory is not strongly representational (although it is seen to be 'weakly' representational).

We proceed as follows. We start sketching and motivating a system of epistemic predicate logic (*EPL*) which deals with anaphora and demonstratives. Then we present a variety of notions of representationalism, and discuss to what extent *EPL* and kindred dynamic systems of interpretation qualify as representational. In this discussion the question comes up to what extent independent motivation can be given for the extended notion of meaning which seem to be required to deal with anaphora. Our, provisional, answer is that it involves the organization of information around 'subjects', which are understood to stand in for potential sources of information.

In this paper we do not discuss 'second order' information, information which agents may have about the information of others. The reason is that we do not want to introduce representations in our object models from the outset. Throughout we assume familiarity with the basic concepts of *DRT*, *DPL*, and E-type pronoun theories.

2 Representation and Information

The general picture which we want to capture in our models is the following. In the first place, there is the real world, with individuals living there. There are also other possible worlds, with other possible individuals, but these are assumed not to interact with the real one. We also take it that certain individuals in the real world carry around information about it, and that they represent the ways things are and that it is useful for them to do so.

The representations that individuals make up can be pictured in the following simple (and artificial) way. They contain a sequence of subjects (in a completely pre-theoretical sense) which these representations are about (equally pre-theoretical) and a list of conditions associated with these subjects. We can simply picture them as *DRT*'s *DRS*s, with discourse markers standing in for the subjects which an epistemic agent is concerned with, and conditions describing what they look like to him or what he knows of them. In this way epistemic agents can be taken to represent (subsets of) the individuals with whom they inhabit the actual world as things having certain properties and standing in certain relations. A *DRS* can be pictured as follows:

$$(1)\quad \begin{array}{|c|}\hline x_1,\ldots,x_m \\ \hline \phi_1 \\ \vdots \\ \phi_n \\ \hline \end{array}$$

A *DRS* of this form represents the information that m subjects satisfy the conditions imposed on them by ϕ_1 through ϕ_n. These conditions may be simple property ascriptions, for instance, that x_i is a gift, that x_i is green, or that x_j is called 'Fred'; but they can also express more complicated relations, for instance, that x_j has shown x_i to everybody who has ever shown x_k to x_l. All this is just standard discourse representation theory, with the term 'discourse' broadly conceived.

Situated representations and represented situations The above picture of a representation can be extended somewhat. Like we said, individuals (epistemic agents) represent (information about) individuals in the actual world, *including themselves*. (A prime subject for the self-supporting being which an epistemic agent is, is he himself, of course.) For this reason we will assume that representations always include one distinguished discourse marker which represents themselves (or their 'carrier', the individual, whoever it is, who has that representation).[1]

Furthermore, in many ordinary cases agents are also concerned with individuals who are demonstratively present. They may gain knowledge about these individuals by perception. One additional assumption that therefore has to be made is that demonstratively present individuals are represented as subjects, too, and that epistemic agents in general are capable of connecting up the markers for these individuals with targets of perception and action, for a proper coordination of their representational apparatus with their senso-motorical activities. We assume that it is an empirical matter of fact which demonstrable individuals are represented by which markers. More generally we assume it is also an empirical matter of fact that certain markers represent individuals which an epistemic agent has been acquainted with in the past.[2] A situated representation can now be pictured as follows:

$$\text{(2)}\quad \boxed{\begin{array}{c} d_1\ ,\ \ldots\ ,\ d_k \\ \mathcal{R}_i = \begin{array}{|c|}\hline x_1,\ldots,x_n \\ \hline \phi_1 \\ \vdots \\ \phi_m \\ \hline\end{array} \qquad l_i(x_1)=d_i \quad l_i(x_j)=d_l \quad \ldots \end{array}}$$

Here, d_1 through d_k are the individuals which are assumed to be demonstratively present. $\mathcal{R}_i$ is the representation of epistemic agent d_i. Clearly, $i \leq k$, for the agent d_i is present himself. Furthermore, there is a linking (or 'anchoring') relation l which determines which markers of $\mathcal{R}_i$ are related to which individuals. In particular, $l_i(x_1) = d_i$, assuming that the first marker represents the Self (the first person). Notice that we have drawn a frame around the situation above. This is not to say that these situations themselves are representations, and that their individuals are markers. Obviously, they are not. What the frame does show is that the aspects of situations which are relevant here can be depicted by representations.

Sofar we have talked about the ways in which epistemic agents might represent the information they have, but our interest does not lie in the specific ways

[1] We assume that one cannot see from a representation who is the carrier of that representation, because different agents may have (or carry) identical representations. More on this below.

[2] These assumptions seem to imply some psychologically realistic conception of these markers. Although we don't want to make any claim about the psychological or cognitive plausibility of *DRS*s, we think this is compatible with our assumption that epistemic agents somehow represent (or 'model', or whatever term the reader thinks appropriate) both the world in which they live *and* the individuals which do so likewise.

in which they do so, but in the actual contents of these representations. As our interest in linguistics lies in spelling out the meanings of sentences, and in meaningful relationships like that of entailment and meaning inclusion, so we are interested, here, in the meanings of representations, and in meaningful relationships like that of information containment.

So what do these representations mean? As a possible meaning of a representation we can think of a sequence of individuals which the representation can be said to be about. If the representation contains n markers $x_1 \ldots x_n$, it can be associated with a sequence of individuals $d_1 \ldots d_n$ such that the individuals $d_1 \ldots d_n$ satisfy the conditions imposed on $x_1 \ldots x_n$. However, since there may be more than one such sequence which satisfies the conditions, it is more appropriate to think of the meaning of a representation generally as the set of all such verifying sequences (or 'cases'). Moreover, since representations also represent general information about the world, we may equate the meaning of a representation with all combinations of a possible world w and a sequence of individuals $d_1 \ldots d_n$, such that, in w the individuals $d_1 \ldots d_n$, in that order, satisfy the conditions imposed on $x_1 \ldots x_n$ by the representation at issue. Such sets of 'witnessing cases' can be pictured as follows:

(3)
$$\left\{\begin{array}{c}\langle w, d_1, \ldots, d_n\rangle \\ \\ \vdots \\ \\ \langle w', d'_1, \ldots, d'_n\rangle\end{array}\right\} \qquad \begin{array}{c}\nwarrow \\ \leftarrow \\ \swarrow\end{array} \begin{array}{l}\text{Possible} \\ \text{cases}\end{array}$$

$$\nwarrow \uparrow \nearrow$$

Possible values
of n subjects

Notice that instead of taking sets of worlds plus sequences of individuals as the interpretations of representations, we might also have chosen to take just the sets of worlds in which there are such witnesses.[3] So, one may wonder, what exactly is the added value of having these subjects? For the moment we can say that subjects are used to model an epistemic notion of aboutness. We have already said above that information states can be said to contain information about real individuals in virtue of certain links which are assumed to exist between an information state's markers and real individuals, but these links are *external* to these states. Now we are dealing with a notion of aboutness *internal* to these states.

Epistemically speaking, individuals are only given to us under their guises. They enter information states as the carriers of the properties we ascribe to them and the relations in which we think or know they stand with others (including ourselves). Now, although we assume that there are real individuals which instantiate all these properties, and which can be said to be the causal root

[3]So, if a *DRS* consists of the markers $x_1 \ldots x_n$ and the conditions $\phi_1 \ldots \phi_m$ one may interpret it, in Montague's *Intensional Logic*, as the intension of $\exists x_1 \ldots \exists x_n(\phi_1 \wedge \ldots \wedge \phi_m)$, instead of that of $\lambda x_1 \ldots \lambda x_n(\phi_1 \wedge \ldots \wedge \phi_m)$, which is our way of interpreting it.

of our information, our information by itself needs not single out any specific individuals at all. Still we can say we have information about the subjects as they are known to us. For instance, suppose you have been told about an African country, about its population, economy, history and whatever, without having been told which country. Then you can say you have information about that country, without exactly knowing which country. In this epistemic sense you may say your information is about an object the identity of which is only partially known to you.

Like partial information about the world, which is modeled in terms of sets of worlds—viz., the set of possible worlds which might be the actual one for as far as you know—we model partial information about (the identity of) individuals in terms of sets of individuals. A partially known individual, a subject, is characterized by the set of individuals each one of which has all the properties you ascribe to him. Here we may notice that a sequence of n subjects is not characterized by a sequence of n sets of individuals, but by a set of sequences of n individuals. For some subjects can be ascribed the property of standing in certain relations to other subjects, and therefore they hang together, like the links of a chain. If I have been told about an African state and about the husband of the private doctor of the king, then by getting more information about which country this is about, I get more information about which husband this is about, and vice versa. Notice that this point would be hard to understand if subjects were taken to be discrete individuals.[4]

Combinatorics of meaning It has been shown at various places that any notion of information about numbers of subjects similar to the one sketched here enables a semantic treatment of the interpretation of singular pronouns and demonstratives. The main idea is that the referents of demonstratives and the antecedents of pronouns are represented by subjects. Since subjects can be conceived of as argument holes of relations, sentences with demonstratives and pronouns can be taken to further constrain the entities which fit in these holes.

Let us consider one example, first from a more representational perspective, and then from a more semantical one. Consider the following sequence of two sentences:

(4) A bear has escaped from the circus. It is chasing *him*.

The interpretation of this sequence can be taken to be built up from three representations. An initial representation $\mathcal{R}_0$ representing the interpreter's initial state of information, and two representations $\mathcal{R}_1$ and $\mathcal{R}_2$, corresponding to the two sentences of our example:

(5) $$\mathcal{R}_0 = \begin{array}{|c|}\hline x_1,\ldots,x_n \\ \hline \phi_1 \\ \vdots \\ \phi_m \\ \hline \end{array} \qquad \mathcal{R}_1 = \begin{array}{|c|}\hline x \\ \hline x \text{ is a bear} \\ x \text{ has escaped} \\ \hline \end{array} \qquad \mathcal{R}_2 = \begin{array}{|c|}\hline \mathtt{p}?, \mathtt{d}? \\ \hline \mathtt{p} \text{ chases } \mathtt{d} \\ \hline \end{array}$$

[4]For a more detailed explanation of the notion of a partial object, cf. Landman 1986, Ch. 2; Dekker 1996.

Here it is assumed that $\mathcal{R}_1$ comes with a side-condition that its marker x be new, because the indefinite noun phrase *a bear* is taken to introduce a new discourse referent. The question marks on the markers in $\mathcal{R}_2$ indicate that the markers have to be resolved. The marker $\mathtt{p}$ translates the anaphoric pronoun *it* and has to be resolved by a marker introduced by an antecedent noun phrase. The marker $\mathtt{d}$ translates the demonstrative *him* and has to be resolved by a marker representing a demonstratively present individual.

Properly merging $\mathcal{R}_1$ with $\mathcal{R}_0$, the marker x is replaced by a new one y and the two representations are put together. Next $\mathcal{R}_2$ can be merged with the result. The marker $\mathtt{p}$ then has to be resolved, into y, for instance, which represents the subject introduced last, and $\mathtt{d}$ has to be resolved, for instance, as the third subject (assuming that this marker x_3 represents a demonstratively present individual pointed at). The net result can be depicted as follows:

$$(6)\quad \begin{array}{|cc|}\hline \multicolumn{2}{|c|}{x_1,\ldots,x_n,y} \\ \hline \phi_1 & y \text{ is a bear} \\ \vdots & y \text{ has escaped} \\ \phi_m & y \text{ chases } x_3 \\ \hline \end{array}$$

Let us now see how the same results can be obtained in a more semantic fashion. Since, as we have said in the introduction, we have side-stepped the process of resolution in this paper, we can think of the meanings of sentences as having the results of such a resolution process built-in. There are several ways in which the resolution of these context dependent items can be formulated. Like in *DPL* (Groenendijk and Stokhof 1991), *EDPL* (Dekker 1996), and *MDPL* (Groenendijk et al. 1996), we can think of subjects as being named by variables, and equate resolved pronouns (and demonstratives) with the variables naming the subjects they refer to. Alternatively, in the system of *EPL* we think of resolved demonstratives and pronouns as referring to a determinate position in the rows of subjects (Dekker 1994; Dekker 1997, cf. also the appendix of this paper). In any case, in order to state the semantic impact of a resolved sentence, its meaning or interpretation has to be specified relative to the result of the interpretation of previous sentences. Thus, the interpretation of a sentence involves injecting the sentence's contents to the information state it is interpreted in.

Let us assume, again, that the subjects which represent demonstrably present individuals occupy the first positions of an information state. A (resolved) demonstrative $\mathtt{d}_i$ then is taken to refer to the i-th subject of that information state (if i is smaller than the total number of demonstrable subjects). Likewise, a pronoun $\mathtt{p}_j$ can be taken to refer to the $n-j$-th subject, if n is the total number of subjects of the information state, and the $n-j$-th one is a subject introduced by preceding discourse. Now, let σ_0 be the interpretation of $\mathcal{R}_0$. The interpretation of the above example then can be spelled out as follows. First σ_0 is updated with our rendering $\exists x(BEARx \wedge ESCAPEx)$ of the first sentence above, which yields σ_1:

$$(7)\quad [\![\exists x(BEARx \wedge ESCAPEx)]\!](\sigma_0) =$$

$$\{\langle w, d_1, \ldots, d_n, d'\rangle \mid \langle w, d_1, \ldots, d_n\rangle \in \sigma_0\ \&$$

$$
\begin{aligned}
&d' \in F_w(BEAR) \ \& \\
&d' \in F_w(ESCAPE)\} \quad (= \sigma_1)
\end{aligned}
$$

($F_w(P)$ here indicates the extension of a predicate P in a world w.) The first update leads to an extension of any case e in σ with a bear in e's world w_e which has escaped in w_e. (That is, if there are escaped bears in such a world; otherwise, the whole case is dismissed.) Thus, an $n+1$-th subject is created that is an arbitrary bear that escaped. Our rendering of the second sentence reads, after resolution, $CHASE\mathtt{p}_0\mathtt{d}_3$. Thus, it adds the condition that this last subject chases the demonstratively present individual which is represented as the third subject:

(8) $[\![CHASE\mathtt{p}_0\mathtt{d}_3]\!](\sigma_1) =$

$$
\{\langle w, d_1, \ldots, d_n, d'\rangle \mid \langle w, d_1, \ldots, d_n, d'\rangle \in \sigma_1 \ \& \ \langle d', d_3\rangle \in F_w(CHASE)\} =
$$

$$
\begin{aligned}
\{\langle w, d_1, \ldots, d_n, d'\rangle \mid\ &\langle w, d_1, \ldots, d_n\rangle \in \sigma_0 \ \& \\
&d' \in F_w(BEAR) \ \& \\
&d' \in F_w(ESCAPE) \ \& \\
&\langle d', d_3\rangle \in F_w(CHASE)\} \quad (= \sigma_2)
\end{aligned}
$$

In the overall result of this tiny interpretation process a subject has emerged which is an arbitrary bear which escaped and which is chasing number three.

To conclude, we have a relatively simple theory here which deals with a notion of information about partial objects, in terms of which the semantics of pronouns can be spelled out in a compositional way.[5] Furthermore, the notion of information at issue is spelled out entirely in terms of possible worlds and their inhabitants, respectable semantic entities, one may say.

Interlude: on Demonstrable Objects In this paper it is assumed that demonstratives get linked to the representatives of demonstrable individuals, which are conceived of as *partial objects*, like subjects. This, really, follows from the epistemic stance we adopt. It is generally acknowledged that there is a latitude between the epistemological and the ontological aspects of demonstrative reference, and this contrast obviously affects the (intended) interpretation of statements using demonstratives. (Cf., the work of Frege, Kripke, Kaplan, Lewis, Perry and Stalnaker.) We think people may simply fail to grasp singular propositions, propositions expressed using directly referential expressions, without thereby disqualifying as epistemic agents.

We may have to emphasize that the issue is not whether it is possible that two different individuals pointed at can be one and the same. It is not, we believe. The point is whether we can make two successful demonstrations and legitimately raise the question whether or not we have pointed at one and the same individual twice, or whether we have pointed at two distinct individuals. We believe we can do so. The point is that if we acknowledge such a possibility—that, for all we know, it might be just as well be the case that d_1 and d_2 are identical, or non-identical—, we are committed to the possibility that we don't have full information about the identity of d_1 or d_2, also in case these are

[5] The system is close in spirit to the one presented in Zeevat 1996.

demonstrative terms. This is what our analysis of demonstrative terms allows for.

By way of illustration, think of two individuals called 'Robin' and 'Joris' which are *documents* in the *world wide web* and inhabitants of a directory called 'Ei'. They are individuals which have indexical information. Robin is linked to herself and to her Ei-partner Joris, and conveys the information that the first, Robin herself, is inspired by the second, Joris. Suppose that the system manager makes a copy of directory Ei, which we, for proper reference, label 'Ei′'. In Ei′ there are two documents Robin′ and Joris′. In Ei′ Robin′ conveys the information that Robin′ is inspired by Joris′. Robin′ is an exact copy of Robin, so Robin′ has the same information as Robin. Thus, precisely the same information is about Robin and Joris in Ei, and about Robin′ and Joris′ in Ei′. More fancy examples can be found in Kaplan 1989; Lewis 1979; Perry 1979.[6]

3 Grades of Representationalism

We now turn to the question to what extent systems of interpretation which deal with singular anaphora are representational. In order to answer this question, we of course first have to answer the question what it means for a theory to be representational, and to this question various answers can be given. In this section we will address three definitions of representationalism which range from what we call 'strong' to what is called 'weak' representationalism. The first, strongest, notion of representationalism is originally presented by Hans Kamp at a workshop on context dependence in Prague, february 1995, and the last, weakest, one was inspired by suggestions made by Ede Zimmermann at the very same workshop. The intermediate notion, we think, naturally suggests itself. For each notion we will find a typical representative. *DRT* will be classified as strongly representational, *DPL* as midweakly representational, *EPL* as weakly representational. E-type pronoun theories can be called non-representational, although this classification, like the others, comes with its qualifications.

Strong representationalism All three notions of representationalism hinge upon a conception of semantic values, or of semantic valuation functions, which are or are not seen to depend on aspects of representations. Hans Kamp presents his definition of (what we call 'strong') representationalism relative to some syntactic and to some semantic theory. The idea is that one first has to make explicit what the language is and what it is about, loosely speaking, and then one can ask whether some specific mapping from the first to the second is representational. When the semantic values of specific syntactic structures are specified, the question is whether this is done in terms of the domain of seman-

[6] Of course there is an apparent qualitative difference between demonstrable objects and subjects construed from hear-say, but this difference can be analyzed in qualitative terms, viz., in terms of the characteristic properties attributed to these objects by information states, and of the highly specific relations in which they are assumed to stand with the information state's Self. At least for the purposes of this paper, there is no need to render it as a structural difference.

tic values only, or whether it is essential to employ (aspects of) the syntactic structures.

Let us be more specific.[7] Let us assume we have a theory $\mathcal{L}$ of sentence structure which assigns categories to the expressions of a (natural) language, and a semantic theory $\mathcal{S}$ which specifies the notion of a model M and the set of possible values of the expressions of a certain category in that model M.

(9) A theory of interpretation is strongly representational with respect to $\mathcal{L}$ and $\mathcal{S}$ iff it makes use of properties of descriptions of semantic values which are not recoverable from the values they describe.

Clearly, the phrase "makes use of" should not be taken to include the specific expressions which are assigned meanings, because every theory of interpretation would qualify as representational then. Maybe the idea can be put more clearly the other way around. If any two syntactically distinct expressions are associated with the same (independently specified) meaning, and, yet, play a different role in the explanation of certain semantic facts, then the explanation of these facts is (strongly) representational. Consider the following two sentences:

(10) A man is walking through the park.

(11) It is not so that no man is walking through the park.

If we make the not unusual assumption that the syntactic category of sentences is associated with a semantic type of truth-conditions, and that these two sentences have the same truth-conditions, then the two should play precisely the same semantic role in a non-representational theory. If, however, our theory wants a semantic explanation for the apparent fact that the sentence 10 licenses a continuation with an anaphoric expression, whereas example 11 does not, then that theory is bound to explain the difference in terms of a difference in syntactic structure. A theory of interpretation which, on the given assumptions, assigns the two sentences a different role, thus qualifies as strongly representational.

It may be clear that discourse representation theory DRT qualifies as strongly representational. The two sentences 10 and 11 are associated with the same meanings in DRT, because they have the same truth conditions, and, yet, they have a different semantic impact, for 10 triggers the introduction of a discourse referent, while 11 does not. Still, we can immediately qualify this claim. For DRT can only be said to be strongly representational relative to the semantic theory according to which the possible semantic values of sentences or discourse representation structures are truth values, or truth conditions. However, alternative theories are conceivable. Disregarding the question of its viability, one can imagine a semantic theory according to which the domain of semantic values consists of *updates of discourse representation structures*, and relative to such a theory DRT can be characterized as not (strongly) representational. Apparently, this possibility points at a conceptual 'leak' in the definition of (strongly) representational. For not only the philosophy underlying original DRT is proclaimed representational, also intuitively DRT seems to qualify thus.

[7] Our rendering of Kamp's definition is made up from our recollection and unreliable transcripts of the slides he showed at the workshop, so it may deviate from the original. We do hope we have preserved the spirit of his proposal.

E-type pronoun theories, too, can be characterized both ways. According to Evans 1977's proposal, pronouns stand proxy for definite descriptions which can be reconstructed on the basis of the syntactic contexts of the pronouns. Clearly, such syntactic contexts cannot in general be reconstructed from the semantic values which they describe, and, thus, this type of pronoun interpretation qualifies as strongly representational. However, one can also argue that the reconstruction of the descriptions is of a pragmatic nature, and then one might say that not the pronouns, but the reconstructed descriptions make up the input for semantic processing (cf., e.g., Cooper 1979). With such a conception of the general interpretive architecture, an E-type pronoun theory does not qualify as representational, not even in a weak sense.[8]

Other theories like *DPL*, *EDPL*, *MDPL* (as well as dynamic Montague grammar, *DMG*, Groenendijk and Stokhof 1990, and compositional *DRT*, *CDRT*, Muskens 1996) can be seen to be not strongly representational. In each of these theories some richer notion of sentence meaning is adopted, and with respect to domains of enriched semantic values interpretation is defined in a compositional and not strongly representational way. But again there is reason to be cautious. As easy as we have given *DRT* a not strongly representational outlook above, as easy we may have done the same to *DPL* and its offspring here. For it is at least conceivable that these enriched domains of semantic values are not genuinely semantic, or that they themselves have to be qualified as representational, in some yet to be specified sense. In order to test these doubts, and to stop the leak alluded to above, we turn to a weaker notion of representationalism which implies a stronger notion of non-representationalism.

Medium, or midweak representationalism With the second notion of representationalism we want to give a principled answer to the question what are respectable semantic values with respect to which theories of interpretation can be qualified as representational. The idea which we want to flesh out here is that a theory can only be called non-representational if its semantic objects can be independently specified, without (explicit) reference to the language which can be used to speak about them.

The second notion of representationalism is defined relative to an architecture $\mathcal{A}$, which specifies what a possible language $\mathcal{A}_{\mathcal{L}}$ is and what are its linguistic categories, and what are possible domains $\mathcal{A}_{(\mathcal{L})_{\mathcal{S}}}$ of semantic values for the categories of $\mathcal{A}_{\mathcal{L}}$.

(12) A theory of interpretation is midweakly representational with respect to $\mathcal{A}$ iff it is strongly representational or its specifications of $\mathcal{A}_{(\mathcal{L})_{\mathcal{S}}}$ depend on specifications of $\mathcal{L}$.

A semantic theory does not qualify as midweakly representational if the languages for whiahc it provides an interpretation are interpreted relative to a pre-given domain of possible semantic values. Notice that the usual *models* of formal languages qualify as midweakly representational, because they include

[8] Although we have to point out that the theory may exhibit gaps, which can only be filled by a resort to (midweak) representationalism, cf., the discussion on uniqueness in the next section.

valuation functions for the constants of a language and, hence, depend on that language. However, what are called 'frames' in modal logic do not qualify thus.

If we adopt the alternative conception of *DRT*'s semantic values, *DRT* is seen to be midweakly representational. If the semantic values of some specific language $\mathcal{A}_{\mathcal{L}}$ are conceived of as updates of *DRS*s, they cannot be specified without referring to $\mathcal{A}_{\mathcal{L}}$, because the updates include those of adding conditions which employ the predicate and relation expressions of $\mathcal{A}_{\mathcal{L}}$. (Notice, however, that this conclusion can be avoided again if we think of *DRS*s as expressions of a language of *thought* in its own right, cf. below.)

All systems of the *DPL*-variety have to be qualified as midweakly representational according to the definition above. In each of these systems the domain of sentence meanings is built up from a domain of variable assignments, i.e., from a domain of functions assigning individuals to the variables of the language which is interpreted.[9] Clearly, no specification of variable assignments can be given if the variables are not given.

It may be noticed that even ordinary first order predicate logic (*PL*) can be characterized as midweakly representational. The meanings of *PL*-formulas are specified in terms of sets of satisfying variable assignments in order to achieve a truly compositional treatment of quantified expressions (cf., e.g., Janssen 1986). Notice, however, that the resort to variable assignments is only relevant for the interpretation of formulas with free variables. When sentences are conceived of as formulas without free variables, the *PL*-combinatorics at the sentential level can be characterized as non-representational. Notice, too, that a corresponding notion of a sentence does not make much sense in *DPL*-style systems, because the raison d'être of these systems precisely lies in their treatment of sentences with free variables (read: pronouns).[10]

For other semantic theories of anaphora like that of *EPL* (presented here), E-type pronoun theories (mind the qualification made above), and choice function theories (Peregrin and von Heusinger 1997) the same things hold as for *PL*. For as far as these theories deal with quantifiers by quantification over variable assignments, they can be judged midweakly representational at the sub-sentential level. The respective notions of sentence meanings, however, do not qualify thus. E-type pronoun theories employ sets of possible worlds, *EPL* employs products from worlds and individuals, and choice function theories use choice functions over sets of individuals. Each of these three types values can be assumed to be given without reference to a specific language to be interpreted.

[9] For the systems of *DMG* and *CDRT* this need not be immediately clear, because the domains of interpretation employed there comprise *primitive* entities ('states'), which are made to *behave like* variable assignments by means of postulates. However, in these two systems the cardinality of the set of variables or discourse markers of a specific language $\mathcal{A}_{\mathcal{L}}$ constrains the cardinality of possible domains of states. The two are therefore seen to be midweakly representational in a more indirect way.

[10] Besides, we notice that the *PL* interpretation can also can be specified without reference to variable assignments, and with reference to sequences of individuals instead (cf., e.g., Tarski 1956).

Weak representationalism The notion of midweak representationalism can still be judged to be too strong a notion, or, conversely, the notion of midweak *non*-representationalism a too weak one. According to this midweak notion a theory of interpretation might qualify as non-representational also if it claims to give an exhaustive specification of the semantics of one (spoken) language by translating it into another one (e.g. a language of thought). Now even if we pass over the methodological and philosophical questions about the viability of such a theory, we want to ascertain that such a theory qualifies as representational in some sense. The notion of (non-)representationalism which seems to be at issue here, is one which Ede Zimmermann has presented as the only sensible one.

For a theory of interpretation to qualify as non-representational, it is not only required that its domain of semantic values can be independently specified, but that it can be independently motivated. The assumption here is that a theory of interpretation always comes with some implicit or explicit idea of what linguistic communication is about, or what it is used for. If this kind of subject matter can be independently motivated (and if the meaning assignment is not stated in a strongly or midweakly representational way), then the theory qualifies as not weakly representational, or just non-representational at all. The requirement can be formulated as follows, in a deliberately somewhat loose way:

(13) A theory of interpretation is weakly representational iff it is midweakly representational or its domains of semantic values cannot be motivated without reference to linguistic facts.

Notice that this definition is mainly concerned with the various conceptions of sentence meanings, or with the various notions of information conveyed by sentences. The mentioned linguistic facts are meant to include all aspects of (linguistic) information exchange. Notice, moreover, that there is some latitude in the characterization of a theory as weakly representational, because, of course, it is always hard to ascertain that something *cannot* be done (viz., that a certain kind of motivation cannot be given). However, this weak notion does allow for a determinate qualification of theories as non-representational, for independent motivation may as a matter of fact have been given for certain domains of semantic values.

Stalnaker, for instance, has given an in-depth explanation of and motivation for the use of sets of possible worlds to characterize the states and actions of living beings (Stalnaker 1984). To our conviction one need not concur with Stalnaker in all details to agree that it is key to our understanding of the behaviour of such agents that they impose some kind of restrictions on what the world is or is like, restrictions which can at least partly be fleshed out in terms of sets of (conceived possible) worlds. We conclude that sets of possible worlds qualify as respectable entities for a non-representational theory.[11]

On the basis of the preceding discussion we can immediately characterize E-type pronoun theories as not weakly representational, that is, as non-represen-

[11] Which is not to say that sets of possible worlds themselves are not representational in some sense. We decide not to launch the profound and confusing discussion we here feel to be on the lurk.

tational. At least under the assumption discussed above, about the division of labour between a semantic and a pragmatic component of an E-type theory, its (sentential) semantics is or can be properly spelled out in terms of sets of possible worlds. (But., cf., the qualification mentioned in footnote 8.) Choice function theories of anaphora may be granted the benefit of the doubt which the loose notion of weak representationalism allows for. For although we don't know, at present, of any independent motivation for the use of choice functions by living agents, yet we also know of no reason for excluding there to be any such possible motivation. This leaves us with the system of *EPL*. Is this system (weakly) representational?

One may be tempted not to qualify *EPL* as (weakly) representational. Its information states can be thought of as 'structured propositions', a kind of entities the existence of which has been motivated in the semantics literature. (We may also think of *EPL*'s information states as states 'inhabited' by subjects.) The real question is whether the organization of information around subjects can be independently motivated. After we have accepted sets of possible worlds as semantic entities which are acceptable in a non-representational theory of interpretation, we now have to look for (independent) motivation for the *additional* structure which characterizes *EPL*-information states.

4 On The Notion of a Subject

In all of the more or less dynamic semantic theories of (singular) anaphora, some notion of information about discourse referents is developed. These discourse referents are the kinds of things which can be introduced by indefinite noun phrases in discourse and they are the kinds of things which anaphoric pronouns can be taken to refer back to. Although such discourse referents have proved useful in the analysis of inter-sentential anaphoric relationships, their nature is still a matter of ongoing debate, a debate which already started in Heim 1982. The issue is whether these discourse referents really are what their name suggests—things which only live in discourse—or whether they are something more. In our terminology the question is what sense we can make of these subjects. Are they only a convenient means to store discourse information, or do they really render the kinds of things actually dealt with by agents who manipulate information? In this section we hope to contribute to a discussion about these issues, which we might label a donkey's worries: can *it* be analyzed away, and, if it can, should it be analyzed away? The discussion heavily draws from (Zimmermann 1998).

What is the added value of having information states which are *inhabited* by subjects? What exactly does such information offer on and above mere information about the world? A first and least ambitious answer runs as follows. Subjects present discourse information. The presence of a subject in an information state represents the information that the state results from interpreting a discourse in which a certain kind of noun phrase has been used, viz., one which can be referred back to by means of pronouns. We call this the 'de dicto' explanation of the notion of a subject. (Zimmermann called it the 'meta discourse' account.) This explanation of the notion of a subject we classify as midweakly representational.

Under the de dicto explanation, subjects are taken to function in the processing of discourse only, they are used to interpret certain anaphorical structures in discourse in a compositional way and they are not attributed any function besides that. A defense of such an account can be found in (Stalnaker 1998), where it is argued that the modeled information about discourse referents can in fact be explained away in terms of the (worldly) information that certain kinds of noun phrases have been used (in the actual world). A similar conception seems to speak from (Groenendijk and Stokhof 1996), in which the authors have it that "Once the discourse has ended, discourse information can be discarded, and the items [read: subjects, PD] can be deleted" (p. 110). However, at the same page Groenendijk and Stokhof acknowledge that their 'discourse items' may also come to life by "the salient presence of an object in the visual field shared by two or more agents". Thus, the de dicto account is seen to provide an incomplete explanation of subjects (or discourse items) in general. Subjects are not only invoked by the utterance of certain noun phrases.

We believe that subjects have more substance than that of mere devices to facilitate the interpretation of (anaphoric) structures in discourse. The last quote from (Groenendijk and Stokhof 1996) already seems to suggest that subjects present a focus of interest common to different information exchanging agents. The key to understanding subjects, we believe, is that they are addressable, and even re-addressable, from different perspectives.

Zimmermann has proposed a notion of subjects as representing *sources* of information, an account of subjects which he has dubbed the 'de origine' account. A demonstrably present subject can be taken to represent the real individual which constitutes the causal source of the information the subject is dressed with. Likewise, a subject which results from a hearer's update with a piece of text can be taken to represent the speaker's source of the subject. Thus, both the subjects grown by discourse, as well as those deriving from real life, are captured under one heading. A subject is re-addressable as another subject, if the two can be taken to be generated from one and the same source. Precisely by relying on such an assumption different agents are justified in referring to a subject, or raising it, and refer back to it afterwards. Thus it also can be seen to make intuitive sense to exchange information about them.

We think that Zimmermann's explanation that the information about a subject is assumed to derive from a determinate source is basically correct, but, also, that one essential aspect still fails. Subjects represent the sources of the information we have about them for a definite reason, viz., that we can gain *more* information about them. Because subjects represent their sources it makes sense for an agent to find out more about them himself, and to let others tell him more about them. Subjects thus should be thought of as representing 'live' sources of information. This account we dub the 'de fonte' account of subjects. Central to the notion of a subject is an agent's assumption that the information he has about a subject derives from a source which hasn't run dry.

The notion of a 'live' source of information has to be understood in quite an abstract sense. An epistemic agent whose information is centered around a subject is taken to assume that his information on the subject is about a real and definite individual (the source), about which in principle more information can be gained. The assumption that such a definite source exists can be substantiated in a variety of ways. The source may be the individual he actually is confronted with, or one that he actually has been acquainted with. In such a case the source can be seen to be the causal root of the information the agents has about the subject. The source of a subject can also be determined by means of an ordinary definite description, as the one and only individual that satisfies certain properties. If we believe there is one and only center of mass of the solar system, or precisely one present king of Belgium or exactly one being which is the first being born after epistemic agents have become extinct, then the center of mass of the solar system, the present king of Belgium and the first being born after epistemic agents have become extinct are all legitimate (possible) subjects.

Moreover, on the assumption that speakers introduce subjects to hearers which are legitimate subjects for the speakers themselves, these newly introduced subjects are legitimate to the hearer, too. For the hearer they may be determined as 'the entity which speaker such and so talked to me about then and then' and its source is (assumed to be) the source of the speaker's subject.[12] It is precisely this assumption which makes it sensible to ask after the identity of a subject introduced by a speaker.

Finally there can be determinate subjects which only live in stories. If a couple of protagonists have a definite role in a certain story and if the story is re-told, then the subjects in the reproduction of the story can be argued to have the protagonists of the original as their sources. Notice, however, that after such a story has been re-told completely, the subjects it has introduced are 'dead'. So, although the actual interpretation of a fictitious story may involve the introduction of legitimate subjects, viz., the subjects which the interpreter is told more about in the process of interpretation, at the end they all become useless. The de fonte account of subjects thus combines intuitions from both de dicto and the de origine account.[13]

Digression The de fonte account of subjects is closely related to but also significantly different from Nirit Kadmon's uniqueness theory of definites (Kadmon 1990). Building forth on the work of Evans, Kadmon defends the view "that

[12]It may need to be emphasized here that things of course can go wrong, e.g., if the hearer is mistaken in assuming that a speaker talks about a legitimate subject. For one thing, indefinite noun phrases do not always introduce subjects. The idea must be, roughly, that if an indefinite is read in something like a topical way, then its semantics can be rendered like that of our dynamic existential quantifier; an indefinite need not always be read that way.

[13]Two things may have to be added here. In the first place, fictional stories of course yield what people may call fictive characters. When properly qualified as fictional, these may serve as subjects, too. In the second place, fictional stories can be (correctly understood to be) about real individuals. In that case it is true, of course, that the individuals are not literally the sources of the information in the story that such and so happened to them, but they can be described, correctly, as those individuals about which the story-teller has (misleadingly) stated that such and so happened.

definite NPs are systematically accompanied by uniqueness implications", also anaphoric definite descriptions, including anaphoric pronouns. Something similar holds in our theory. In our theory pronouns may refer back to antecedents in virtue of the subjects which these antecedents set up, and subjects have been taken to come with the assumption that they derive from definite (i.e., unique) sources. Hence, in our theory, pronouns are associated with uniqueness assumptions, too. Furthermore, both Kadmon's and our theory agree that all kinds of pragmatic factors may be involved in determining the unique sources of the subjects involved (or of the referents of the pronouns).

The major difference between Kadmon's theory and our's lies in the identification of the source of the uniqueness condition (a subject introducing noun phrase versus an anaphoric pronoun), and more importantly, in its status. In our view it holds as a general (default) assumption that subjects derive from definite sources, but we don't think that all subjects must meet this requirement, and we think a semantic theory can leave it open which uniqueness conditions are associated with specific subjects. Notably, we think that the fact that an indefinite noun phrase introduces a subject with an identifiable source does not belong to its truth-conditions. In contrast, in Kadmon's theory, like in that of Evans, the uniqueness requirements do show up in the truth conditions (hence, the term 'uniqueness *implications*').

Consider the following two examples from Heim, also discussed in Kadmon's paper:

(14) A wine glass broke last night. It had been very expensive.

(15) A wine glass which had been very expensive broke last night.

For the pronoun *it* in example 14 to relate to the indefinite *a wine glass*, this indefinite must set up a subject, and for this subject to be legitimate, the speaker must have had some definite wine glass in mind. (The wine-glass must at least be a subject to him.) But, in our theory, this is a *pragmatic* condition on the legitimate introduction of discourse referents. In Kadmon's theory, on the other hand, the requirement that a unique wine glass broke last night is conceived of as a *semantic* requirement. We think this to be problematic. Kadmon concedes that the glass may not have been the only wine glass which broke last night, without this rendering 14 false, but she adds: "[h]owever, it **does** have to be unique in **some** way" Kadmon 1990, p. 282. The problem here is that one simply cannot see from the two sentences in example 14 in **which** way the wine glass is unique.

We think Kadmon's theory here faces a severe complication. Let us agree that the pronoun *it* in example 14 refers to the wine-glass mentioned in the sentence before, and also that the example does not by itself supply more information on the subject than what is said, viz., that a very expensive wine glass broke last night. This is precisely what is modeled in *DPL*, *EPL*, etc. Let us also agree, with Kadmon, that an utterance of example 14 must concern a particular wine glass, which the speaker must have in mind. We take it that this assumption is not part of the meaning of 14, and likewise, in *DPL*, *EPL*, etc., we simply pass by the question which specific uniqueness requirement is involved here, and leave that question to pragmatics. Now, in Kadmon's theory, this question cannot be, thus, delegated to pragmatics. For, for the pronoun to receive any interpretation at all, the specific way in which the glass is unique must be made explicit. But then, for the semantic analysis of example 14 to be completed, it seems we can do no better that interpreting the pronoun as "the glass which the speaker has in mind when he utters the noun phrase "a glass", a typically (midweakly) representational description.

Despite the dissimilarities, our theory does help in explaining a difference Kadmon observes between the two examples 14 and 15. Since the indefinite noun phrase *a wine glass* in example 15 does not serve as an antecedent for an anaphoric pronoun, it need not be read as setting up a subject (although it can of course be read that way). This sentence can therefore be uttered felicitously, also if the speaker does not have a specific wine glass in mind. Thus we agree with Kadmon that if a speaker does not have a specific glass in mind, then an utterance of 14 would be infelicitous, and one of 15 felicitous. *End of digression*

Now that we have laid out and motivated our notion of subjects, let us return to the main question of this section, whether this constitutes an independently motivated notion of information. Again, there is no clear-cut answer to the question. It seems sensible enough to organize information around subjects, if subjects represent sources of information which an agent may gain more information from. Thus, the package of information which constitutes a subject can be extended with additional information concerning the same source. Notice that such a use of subjects doesn't need to involve communication or information exchange.

However, and now we are turning to our second question, viz., whether subjects can be analyzed away, it seems that the type of update of information mentioned can as well be spelled out in terms of updates of information states which are sets of possible worlds. On the de fonte account an agent may extend his information about a subject with additional information if he is justified in assuming that the subject and the additional information derive from the same source. However, this presupposes that the agent knows enough of the subject to judge whether the additional information derives from the same source, which is to say that he has identifying information. A subject δ, thus, is known as the unique individual or source which has a certain package π of properties. For this reason, any addition of information that δ also has some property ρ can equally well be spelled out as an update on sets of possible worlds, simply by taking the restriction of such a set to those worlds in which the unique individual which has all the properties in π, also has the property ρ. In short, if the subjects of an agent come with unique source assumptions, then the agent can refer to each of them using definite descriptions.

The above line of argument seems to lead to the conclusion that the structure of our inhabited information states is not independently motivated after all, and, hence, that the *EPL* system must be characterized as weakly representational. But again we have to qualify this characterization. In the first place, we belief that one subject, the Self, cannot be eliminated along the above lines; in the second place, we believe that a general elimination of subjects would bring in explicit representationalism again.

David Lewis distinguishes information about the world (which he conceives of as information about the location of epistemic agents in *logical space*) from self-locating information—information about the location of agents in *ordinary time and space* (Lewis 1979). Building on the work of John Perry (Perry 1979), Lewis argues that even if somebody has complete information about the world, he may still fail to know who or where he himself in fact is. For this reason, the second kind of information cannot be appropriately dealt with in terms of the first, and Lewis therefore proposes to render self-locating information, not as functions from possible worlds to truth values, but as functions from worlds to sets of individuals. The idea is that if an agent's information is characterized by such a property π, and if individual d is in the extension $\pi(w)$ of π in a world w, then the agent considers it possible that w is the actual world and that he himself is the individual d in w (or that his location is that of d in w). Notice that Lewis' space of information states is isomorphic to the set of states of information about one subject (viz., the Self).

If, as we think, Lewis is right about the indispensability of the Self, then we have rescued at least one subject. Although the Self of any concrete information state may have a unique source (its owner), and although it may refer to itself with the *relative* description "the owner of this state", it need not be associated with any *absolutely* identifying information. In one possible world two distinct individuals may have all the properties which an information state ascribes to itself (e.g., the individuals Robin and Robin$'$ mentioned in section 2). So, reference to the Self may essentially involve the kind of self-locating information which, as Lewis argues, is not retrievable from sets of possible worlds.

Now what about subjects other than the Self? Although the subjects of an information state are partial objects, we have argued they are assumed to be in some or other way uniquely related to the state's Self. Demonstrably present subjects stand in for the real individuals which are assumed to cause our information about them, and subjects constructed from hear-say stand in for the (assumed) sources of those who have introduced them to us. If our assumptions are correct, then in both cases the sources are determinable relative to our Selves, and, hence, the corresponding subjects are eliminable. However, as we argued above, the determination of subjects constructed from hear-say will generally be of a particular representational kind. The subjects which are introduced to us by others will often only be known as "the subject which such and such introduced to me as a 'so and so', and such that etc". In such a case the sources are only known under a typically (midweakly) representational description.

We therefore conclude this section by locating the realm of inhabited information states at the borderlines of representationalism and non-representationalism. Compared to respectable entities like sets of possible worlds, they are characterized by additional structure which we judge useful to agents who collect and exchange information about their world and its inhabitants. Following Lewis, the subject called the Self has been presented as an irreducible subject. The other subjects have been argued to be eliminable. However, the eventual elimination of subjects which derive from discourse in its turn invokes explicit (midweak) representationalism. In this sense *EPL* constitutes a system for the interpretation of pronouns in the least (but weakly) representational way.

5 Conclusion

Starting from the metaphor of information states as (the interpretations of) discourse representations, we have proposed to specify the information which is conveyed by a discourse, and which is also relevant for the proper interpretation of further discourse, as information about the numbers of individuals the discourse is about. Formally this can be fleshed out in terms of sets of sequences e consisting of a world w and n individuals, such that in a pre-given model, the sequences of individuals in the sequence stand in w in the relation attributed to them by the discourse. These information states have been argued to model two aspects of information: (i) information about the world, and (ii) information about subjects. The first aspect corresponds to what one may called the standard, model-theoretic notion of information. The second aspect is relatively new, and it includes what Lewis has called 'self locating' information, or what is also known as information 'de se'. This self-locating aspect of information has been argued to be independently required.

Information about subjects has been modeled in various ways in the literature, in terms of discourse representations (DRT), files or sets of variable assignments (FCS, DPL), and sets of cases (EPL). The first way of modeling them is classified as strongly, the second as midweakly and the third as weakly representational. As an explanation of the concept of a subject we have proposed a 'de fonte' account, which has it that the subjects of an information state represent what are assumed to be 'live' sources of information. Subjects are not strongly or midweakly representational because they need not be related only to the discourses they evolve from, or with the noun phrases that introduce them. Still, the motivation for having subjects does appear to lie in securing an efficient update and exchange of information about retrievable subjects under discussion.

A final issue which must be addressed here is that of the *order* of subjects of an *EPL* information state. It seems that such an order can only be understood as the order in which these subjects have been introduced in an information state. However, if it is not for the purpose of constructing the interpretation of a discourse in a compositional way, this order of introduction can hardly be seen to constitute an independently useful kind of information. This aspect of information thus seems to be particularly weakly representational. The (puzzling) issue is that there appears to be no way of removing this structure without making these states midweakly representational again.

As we already mentioned above, *EPL* information states can be thought of as the intensions of λ-terms, and the order of subjects, thus, can be seen to derive from the order in which the various free variables are bound by λ-operators. We can not do away with this order in ordinary λ-calculus, precisely because the various λ-operators must be added one by one. Recently, however, calculi for simultaneous abstraction over *sets of* variables have been developed in (Ruhrberg 1996). Simultaneous abstraction can be used to remove the order from the subjects in information states, which seems to be what we want. However, for a proper formulation of the semantics of these system of simultaneous abstraction, the simultaneous abstracts have to be applied to *indexed* argument terms. Thus, if we abolish the order of subjects of *EPL* information states by simultaneous abstraction, we are forced to *name* them again, and we seem bound to get a midweakly representational system like *DPL* or *EDPL* in return.

In this paper we have gone quite a long way to end up with a model of information exchange in which the real entities we deal with in the actual world are rendered to us as partial objects. This has enabled us to deal both with the (epistemic) puzzles presented by Frege, as well as with the linguistic puzzle posed by (singular) pronouns. Still, the reader may be worried about the implications of our picture for (his information about) the things he is acquainted with, and for ordinary first order theories, which, after all, are understood to deal with domains of ordinary individuals. It turns out that the implications are not so dramatic.

If the information of an agent concerns a fixed domain of things which is properly individuated to him (not with other things or with information others may have about the domain), then the classical and our epistemic picture coincide. In such a case all identity issues are settled, and it won't make any difference if the things at issue are conceived of as partial objects (subjects) or as total ones (individuals). One may observe that agents in actual life situations most of the times find themselves concerned with such well-individuated, ordinarily behaving domains. And also in classical first order logic, the total number of subjects of discussion gets introduced as a discrete domain of fully individuated individuals. Thus, in these real situations, and in the first order models, nothing unexpected can happen if what seemed to be individuals, turn out to be subjects after all.

Appendix, the *EPL* system

The *EPL language* is that of predicate logic with two additional categories of terms, viz., anaphoric pronouns ($\mathtt{p}_0, \mathtt{p}_1, \ldots$) and demonstratives ($\mathtt{d}_1, \mathtt{d}_2, \ldots$). *EPL interpretation* is defined as a (partial) update function on information states. The update of information state σ with formula ϕ is defined relative to a model M and a variable assignment g, the result of which is an information state indicated as $[\![\phi]\!]_{M,g}(\sigma)$. An *EPL model* M is a triple $M = \langle W, D, F\rangle$, where W is a set of possible worlds, D a domain function which assigns a domain of individuals D_w to each world $w \in W$, and F an interpretation function, which assigns in each possible world w extensions $F_w(c) \in D_w$ and $F_w(R) \subseteq D_w^n$ to individual constants c and n-ary relational constants R.

Definition 1 (Information States)

$$\sigma^n \subseteq \bigcup_{w \in W}(\{w\} \times D_w^n) \text{ is an information state about } n \text{ subjects}$$

For $s \in S^n$ and $e = \langle w, \langle d_1, \ldots, d_n \rangle\rangle \in s$, we will also write e as $\langle w, d_1, \ldots, d_n\rangle$, we indicate its length n as N_e, and for $i \leq n$ we let e_i indicate d_i and w_e indicate w. Generally we assume that the first subjects of an information state are representatives of individuals which are demonstratively present, and that the last ones have been introduced by indefinite noun phrases in discourse. The first subjects are addressed by demonstratives, the last by anaphoric pronouns.

Definition 2 (Interpretation of Terms)

$$\begin{array}{ll}
[c]_{M,g,e} & = F_{w_e}(c) \text{ for all constants } c \\
[x]_{M,g,e} & = g(x) \text{ for all variables } x \\
[\mathtt{d}_i]_{M,g,e} & = e_i \text{ for all demonstratives } \mathtt{d}_i \\
[\mathtt{p}_j]_{M,g,e} & = e_{N_e - j} \text{ for all pronouns } \mathtt{p}_j
\end{array}$$

Since a demonstrative $\mathtt{d}_i$ fixes the i-th individual of any case e in σ, it actually refers to σ's i-th subject. The silent assumption here is that σ's i-th subject as a matter of fact is a representative of a demonstratively present individual. A less silent assumption is, of course, that σ has at least i subjects. If these two assumption are not met, the demonstrative is undefined (or its index infelicitous).

Something similar holds of the interpretation of pronouns. A pronoun $\mathtt{p}_j$ fixes the $n-j$-th individual of any case e in σ, where n is the number of subjects in σ. The silent assumption now is that σ has more than j subjects, and that the selected one is introduced by preceding text. If these two assumption are not met, the pronoun is undefined (or its index infelicitous).

Definition 3 (Interpretation of Formulas)

$$\begin{array}{ll}
[\![Rt_1 \ldots t_n]\!]_{M,g}(\sigma) & = \{e \in \sigma \mid \langle [t_1]_{M,g,e}, \ldots, [t_n]_{M,g,e}\rangle \in F_{w_e}(R)\} \\
[\![t_1 = t_2]\!]_{M,g}(\sigma) & = \{e \in \sigma \mid [t_1]_{M,g,e} = [t_2]_{M,g,e}\} \\
[\![\neg\phi]\!]_{M,g}(\sigma) & = \{e \in \sigma \mid \neg\exists e' \colon e \cdot e' \in [\![\phi]\!]_{M,g}(\sigma)\} \\
[\![\exists x\phi]\!]_{M,g}(\sigma) & = \{e' \cdot \langle d\rangle \mid e' \in [\![\phi]\!]_{M,g[x/d]}(\{e \in \sigma \mid d \in D_{w_e}\})\} \\
[\![\phi \wedge \psi]\!]_{M,g}(\sigma) & = [\![\psi]\!]_{M,g}([\![\phi]\!]_{M,g}(\sigma))
\end{array}$$

where $e \cdot e'$ indicates the concatenation of e and e'

These clauses are pretty straightforward, except, maybe, the one dealing with existential quantifiers. The interpretation of an existentially quantified formula $\exists x\phi$ in a state σ involves, for any individual d, the interpretation of ϕ under the assignment of d to x in the information state $\{e \in \sigma \mid d \in D_{w_e}\}$, which is the state σ updated with the information that d exists. All possibilities that can be found in the output of this update are next extended with d as the witness for the existential quantifier.

Bibliography

Cooper, R.: 1979, The interpretation of pronouns, in F. Heny and H. Schnelle (eds.), *Syntax and Semantics 10*, Academic Press, New York

Dekker, P.: 1994, Predicate logic with anaphora (seven inch version), in L. Santelmann and M. Harvey (eds.), *Proceedings of SALT IV*, pp 79–95, DMLL Publications, Cornell University

Dekker, P.: 1996, The values of variables in dynamic semantics, *Linguistics and Philosophy* 19(3), 211–257

Dekker, P.: 1997, On context and identity, in H. Kamp and B. Partee (eds.), *Context-dependence in the Analysis of Linguistic Meaning. Volume I: Papers*, pp 87–115, IMS, Stuttgart

Evans, G.: 1977, Pronouns, quantifiers and relative clauses (1), *The Canadian Journal of Philosophy* 7, 467–536, reprinted in: Evans, G., 1985, *Collected papers*, Dordrecht: Foris, pp 76–152

Frege, G.: 1892, Über Sinn und Bedeutung, *Zeitschrift für Philosophie und philosophische Kritik* NF 100, 25–50

Gawron, J. M. and Peters, S.: 1990, *Anaphora and Quantification in Situation Semantics*, CSLI, Stanford

Groenendijk, J. and Stokhof, M.: 1990, Dynamic Montague grammar, in L. Kálmán and L. Pólos (eds.), *Papers from The Second Symposium on Logic and Language*, pp 3–48, Akadémiai Kiadó, Budapest

Groenendijk, J. and Stokhof, M.: 1991, Dynamic predicate logic, *Linguistics and Philosophy* 14(1), 39–100

Groenendijk, J. and Stokhof, M.: 1996, Changing the context. Dynamic semantics and discourse, in E. Doron and S. Winter (eds.), *Proceedings of the 11-th Annual Conference Tel Aviv University 1995 and of the Workshop on Discourse The Technion 1996*, IATL

Groenendijk, J., Stokhof, M., and Veltman, F.: 1996, Coreference and modality, in S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, pp 179–213, Blackwell, Oxford

Heim, I.: 1982, *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. thesis, University of Massachusetts, Amherst, published in 1988 by Garland, New York.

Heim, I.: 1983, File change semantics and the familiarity theory of definiteness, in R. Bäuerle, C. Schwarze, and A. von Stechow (eds.), *Meaning, Use, and Interpretation of Language*, de Gruyter, Berlin

Heim, I.: 1990, E-type pronouns and donkey anaphora, *Linguistics and Philosophy* 13(2), 137–78

Janssen, T. M.: 1986, *Foundations and Application of Montague Grammar. Part 1: Philosophy, Framework, Computer Science*, CWI Tract 19, CWI (Centre for Mathematics and Computer Science), Amsterdam

Kadmon, N.: 1990, Uniqueness, *Linguistics and Philosophy* 13, 273–324

Kamp, H.: 1981, A theory of truth and semantic representation, in J. Groenendijk, T. Janssen, and M. Stokhof (eds.), *Formal Methods in the Study of Language*, Mathematical Centre, Amsterdam, Reprinted in J. Groenendijk, T. Janssen, and M. Stokhof (eds.), *Truth, Interpretation and Information.*

Foris, Dordrecht, 1984

Kamp, H. and Reyle, U.: 1993, *From Discourse to Logic*, Kluwer, Dordrecht

Kaplan, D.: 1989, Demonstratives. An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals, in J. Almog, J. Perry, and H. K. Wettstein (eds.), *Themes from Kaplan*, pp 481–563, Oxford University Press, New York

Landman, F.: 1986, *Towards a Theory of Information*, Foris, Dordrecht

Lewis, D.: 1979, Attitudes de dicto and de se, *The Philosophical Review* 88(4), 513–45

Milward, D.: 1996, Integrating situations into a theory of discourse anaphora, in P. Dekker and M. Stokhof (eds.), *Proceedings of the Tenth Amsterdam Colloquium*, pp 519–538, ILLC, University of Amsterdam

Muskens, R.: 1996, Combining Montague semantics and discourse representation, *Linguistics and Philosophy* 19(2), 143–186

Peregrin, J. and von Heusinger, K.: 1997, Dynamic semantics with choice functions, in H. Kamp and B. Partee (eds.), *Context-dependence in the Analysis of Linguistic Meaning. Volume I: Papers*, pp 329–353, IMS, Stuttgart

Perry, J.: 1979, The problem of the essential indexical, *Noûs* 13, 3–21

Ruhrberg, P.: 1996, *Simultaneous Abstraction and Semantic Theories*, Ph.D. thesis, University of Edinburgh

Stalnaker, R.: 1984, *Inquiry*, MIT Press, Cambridge, Mass.

Stalnaker, R.: 1998, On the representation of context, *Journal of Logic, Language and Information* 7, 3–19

Tarski, A.: 1956, The concept of truth in formalized languages, in *Logic, Semantics, Metamathematics*, Clarendon Press, Oxford

Vermeulen, C. F.: 1996, Merging without mystery or: Variables in dynamic semantics, *Journal of Philosophical Logic* 24(4), 405–450

Zeevat, H.: 1996, A neoclassical analysis of belief sentences, in P. Dekker and M. Stokhof (eds.), *Proceedings of the Tenth Amsterdam Colloquium*, pp 723–741, ILLC, Amsterdam

Zimmermann, E.: 1998, Remarks on the epistemic rôle of discourse referents, in L. Moss (ed.), *Proceedings of ITALLC 96*